

The Concept of Genetic Point Mutations¹

Please be sure to read the textbooks treatment of repair mechanisms (pp 305-307) and point mutations (pp. 328 – 330) – there are questions dealing with this reading attached to these notes.

Review: We have seen that structural genes contain information necessary to specify the primary structure of a protein². Furthermore, we have learned that to the extent that the environment is predictable, a given primary structure will interact with the environment and produce predictable higher-order protein structure and therefore predictable a protein with predictable function. Thus individual genes have evolved to produce a specific protein structure -- the information in those genes, through the transcription, translation, and post-translational modification processes will give a predictable result provided that the "expected" environment exists when the protein is formed and used. This has also led us to the realization that if the environment changes in certain ways, the protein produced will not be exactly what the gene "intended". One of the reasons that homeostasis evolved was to reduce the frequency at which proteins would be produced that did not function correctly because the genetic information interacted with the wrong ("unexpected") environmental conditions. Thus, homeostasis implies increased genetic control over protein characteristics.

Mutation: When genetic information is altered, we say that a **mutation** has occurred. In the days before molecular genetics, mutation simply referred to some heritable change in phenotype. In this class we will look at mutations that occur at the nucleotide level. These are termed **point mutations**. In some cases they will cause very noticeable changes in the phenotype while in others the mutations will be invisible except to a molecular biologist who compares the sequence of nucleotides for the same locus but from different individuals. Other types of genetic changes that typically involve large amounts of DNA (instead of just a few nucleotides), such as duplications or deletions of parts of chromosomes, have historically been called mutations but now are more generally referred to as "chromosomal aberrations".

Mutation Rates: Typically, noticeable mutations occur in about 1 out of a million copies of a gene. The actual rate is somewhat higher since some mutations are "silent" (see below). The rate could be considerably higher than it is if was not for the presence of mechanisms to correct mistakes during replication (primarily mist-match repair) and to correct mistakes that crop up as cells are exposed to

¹ © 2006 by W. R. Healy and K.N. Prestwich, Dept. Biology, College of the Holy Cross, Worcester, MA 01610

² Recall that we have seen that the information actually used to produce a primary structure is modified by splicing.

various physical and chemical mutagenic agents. See your textbook readings for more details.

Types of Point Mutations and How They Cause Phenotypic Effects:

Substitutions: In this case, a single base pair within a gene becomes altered. For instance, the sequence **ATTGCT** becomes **ACTGCT**. In many cases, this makes no difference at all to the final protein product. The reason traces back to the "**degeneracy**" or **redundancy** of the genetic code. For instance, let's say that a particular part of a structural gene has the sequence **GAA** that in turn corresponds to the codon **CUU**. This codon stands for the amino acid leucine. If the third nucleotide in this sequence, **A**, mutates to a **G**, the new mRNA codon will be **CUC**. This still codes for leucine. In fact the following sequences all code for leucine: **GAA, GAG, GAT, GAC** (corresponding to the codons **CUU, CUC, CUA, CUG**). Thus, starting from the original **GAA**, the final nucleotide of this triplet could mutate to either a **G, T, or C** and the final protein would still be the same. Such mutations are only detectable by molecular means where sequences are determined and compared. Since these mutations have no effect on the phenotype³, we call them **silent mutations**.

However, not all substitutions are silent. Let's continue with the example of the DNA sequence **GAA**. Let's substitute for the first nucleotide in the codon. If we flip from **G** to **A** so that the sequence is now **AAA**, we call for phenylalanine -- an amino acid with different properties from leucine. Instead, let's change the **G** to **C** (**CAA**) -- now we call for valine. **G** to **T** (**AUU**) -- now we call for isoleucine. What if we alter the second nucleotide instead of the first? Let's change the **A** in **GAA** to a **T** (**GTA**) -- instead of calling for leucine, we now call for histidine, an amino acid with strikingly different properties. OK, so you get the idea. Many point substitutions will cause a change in the amino acid that is specified somewhere in the protein.

What will be the effect of these mutations? In some cases, there will be virtually no effect. If one amino acid is replaced by another with very similar properties (example alanine for glycine or valine for leucine -- see fig 5-17 on page 79 in Campbell) it is likely that there will be very little if any noticeable effect on the protein's function. The mutation may only be noticeable in the sense that if someone sequenced the amino acids in the protein or did certain other tests of its properties, they might find differences. So, the mutation could be essentially silent even though they did cause slight differences in protein structure.

However, in most cases, this sort of mutation will not be silent. Most amino acids have very different properties from each other. In the example of changing the triplet **GAA** to **GTA** with a change in the protein from leucine to histidine (see p. 69 in Campbell), you are substituting a non-polar, relatively small amino acid (leucine) for a large, alkaline, charged amino acid. Great differences in higher order structure could be expected to result.

³ Silent mutations do affect the DNA structure itself and, after all, this is also part of the phenotype! But we usually ignore this and for most circumstances think of them as having no phenotypic effect.

One of the most famous examples of this is with the allele for sickle cell anemia. This allele contains a single base-pair substitution that result in a single amino acid difference in hemoglobin as compared to "normal" hemoglobin. The resulting hemoglobin has very different properties, especially when the amount of O₂ in a red blood cell drops.

Frameshift Mutations: Insertions and Deletions: Imagine that we have a gene that codes for a protein that is 200 amino acids long. Thus, the mRNA molecule that directed its synthesis contains 200 codons -- put another way, it contains 200 triplet reading frames. Since each codon is a triple, then the mRNA is 600 nucleotides long. The exons of the gene that coded for this mRNA must also contain 200 reading frames, one for each amino acid. It exons are 600 amino acids long.

In the previous examples of base substitution, there was no change in the length of the gene. However, sometimes, entire base pairs are deleted; in other cases an extra base pair gets added. Thus, the sequence **AGCATAC** becomes **ACATAC** (deletion of a guanine) or **AGTCATAC** (addition of a thymine). Notice what happens. Assume that in the sequence AGCATAC corresponded to the codons UCG and UAU. After the addition, we now have the codons UGU and AUG (a stop codon); after the addition, we now have the codons UCA and GUA. Notice what has happened is that starting with the first codon effected, every codon from then on potentially will be changed and many if not most of these changes will lead to new amino acid specifications. **Frameshift mutations are truly devastating and usually result in a total loss of function of the gene in question.**

Questions about point mutations:

1. Why don't frameshift mutations affect all genes downstream from the frameshift? -- Why do they only affect the gene where the mutation occurs? After all, it's all one big DNA molecule and there are no gaps in nucleotides between genes. *Hint* -- think about what you have learned about the organization of genes and the why that transcription happens.
2. Where is the worst place for a frameshift mutation -- should it end up showing up near the 5' or 3' end of a mRNA molecule to cause the most mischief? Would this matter with substitutions?
3. Do you think that "mutations" ever occur in mRNA -- *i.e.*, does mRNA ever end up with mistakes in its sequence? What are the consequences?
4. Why is it that most non-silent point mutations are deleterious? Think about your answer in terms of the evolutionary history behind each protein molecule.
5. What is the effect of a substitution that creates a stop triplet in the middle of a structural gene? Let's say the gene is for a protein that is 100 amino acids long and the mutation occurs at position #50. What will be the result?

6. What, if any, would be the effects of point mutations in regulatory sequences? Put another way, can mutations in regulatory sequences affect the organism?

Questions from Campbell, pages 305 – 307—repair and telomeres

1. What is mismatch repair? What type of point mutation does it help to decrease?
2. What is "proofreading"? What molecule accomplishes this function?
3. What is excision repair? When does this occur -- during synthesis or otherwise? You need not learn the exact details of excision repair -- wait and take Genetics.
4. Why are the 5' ends of linear DNA molecules a problem in replication? What structures exist on the ends of chromosomes to lessen this problem?
5. What is telomerase and what is its role in cell immortality?

Finally -- one to think about:

Can you select for a mutation rate? In other words, is it possible to evolve higher and lower mutation rates than those that we typically observe?