Genetics Lecture 1

We will begin this course with the question: What is a gene?

This question will take us four lectures to answer because there are actually several different definitions that are appropriate in different contexts.

We will start with a physical definition of the gene. Conceptually this is the simplest and it will give me an excuse to briefly review some of the molecular biology that you probably already know.

**Genes are made of DNA**

For this course we will mostly think of DNA as an information molecule not as a chemical substance.

In 1953, Watson and Crick deduced that the structure of DNA was a double helix. It was not the helical structure per se, but the discovery of complementary base pairing that revealed how information could be encoded in a molecule and how this information could be exactly duplicated each cell division. **Replication**.

In order to extract information from the DNA, the cell again uses the complementary base-pairing to make a copy of the information copied onto an RNA molecule. This is known as **Transcription**. RNA is chemically less stable than DNA and mRNA can be thought of as a temporary copy of DNA’s information.
Transcription

Translation

Folded proteins:
enzymes
structural proteins
membrane channels
hormones

Gene: DNA segment needed to make a protein

Genes are typically $10^3 - 10^4$ base pairs in size although they can be much larger. For example, the human dystrophin gene is $2 \times 10^6$ base pairs.

*E. coli* has about 4,200 genes which isn't very many considering that at least 1,000 different enzymes are needed carry out just the basic biochemical reactions in a cell. The smallest genome for a free-living organism (i.e. a cell, not a virus) is that of the bacterium *Mycoplasma genitalium* which encodes only 467 genes. Humans are at the other end of the spectrum of complexity and have about 35,000 genes.
In the demonstration in class you see how a mutation in the *Shibire* gene in the fly *Drosophila* gives a heat sensitive protein that is required for synaptic transmission. When the flies that carry this mutation are warmed by the projector lamp they become paralyzed.

Gene — Protein — Cell Process — Organism “disease”

(Shibire) (Dynamin) (Synaptic Signaling) (Paralyzed Fly)

This example illustrates two powerful aspects of genetic analysis. First, we can follow microscopic changes in the DNA such as the *Shibire* mutation as they are revealed by the macroscopic consequences of the mutation such as a paralyzed fly. Second, we have a very precise way of studying the function of individual proteins by examining the consequences of eliminating just that one protein function in an otherwise normal organism.

**Alleles:** different versions of the same gene

Often alleles are referred to as mutants but actually this usage is often incorrect particularly when we discuss naturally occurring variants in a population.

**Mutation:** an altered version of a gene when we have “witnessed” the alteration but not when it is preexisting in the population.

**Genotype:** all alleles of an individual

**Wild type:** defined standard genotype

The concept of wild-type is used as a defined reference for organisms where we can do breeding experiments. Of course there is no realistic way to define a standard genotype for humans, therefore “wild type” has no meaning when we discuss human genetics.

The physical definition of the gene is a very good one but there are many instances where we wish to study genes whose DNA sequences are not known. For example, say we have isolated a new mutant fly that is also paralyzed and we want to know whether this mutation is also in the *Shibire* gene. We will see in the next several lectures that we can answer this question without knowledge of the DNA sequence either by a test for gene function known as a complementation test or by a test of the chromosomal position of the mutation by recombinational mapping. In practice, these other ways of defining genes by function or by position are often much more useful than a definition based on the DNA sequence.