

# Genetics Lecture Notes

7.03 2006

Lectures 1 - 2

## Lecture 1

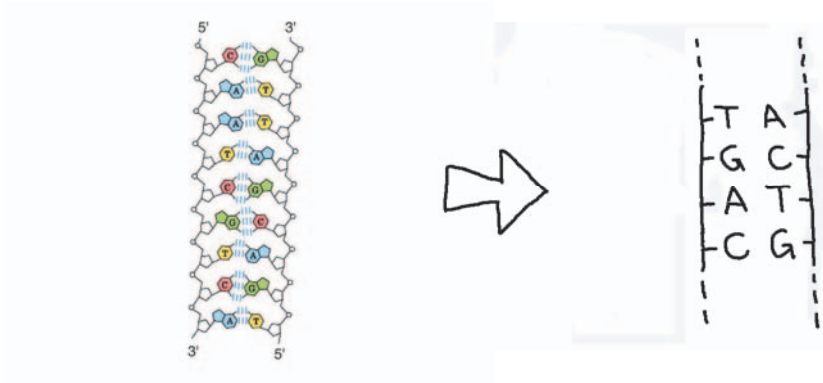
We will begin this course with the question: What is a gene?

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.

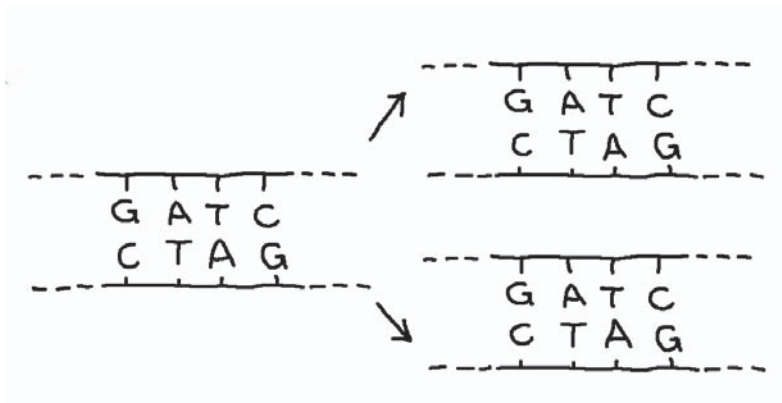
Genes are made of DNA

For this course we will mostly think of DNA as an information molecule rather than a chemical substance.



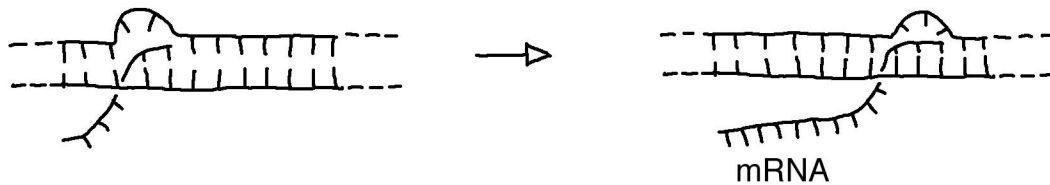
In 1953, Watson and Crick deduced that the structure of DNA was a double helix. It was not the helical structure per se, but the discovery of complementary base pairing that revealed how information could be encoded in a molecule and how this information could be exactly duplicated each cell division.

### Replication.

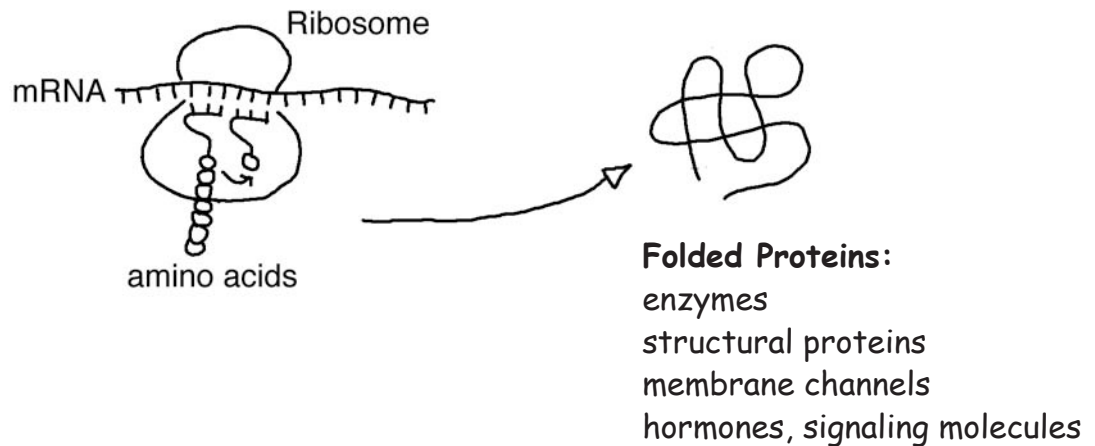


In order to extract information from the DNA, the cell again uses the complementary base-pairing to make a copy of the information copied onto an RNA molecule. This is known as Transcription. RNA is chemically less stable than DNA and mRNA can be thought of as a temporary copy of DNA's information.

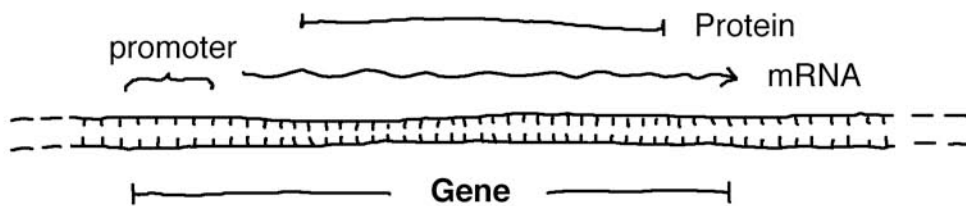
### Transcription



### Translation



**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 to 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 20,000 - 25,000 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	—	Cellular function	—	Organismal function
( <i>Shibire</i> )		(Dynamamin)		(Vesicle recycling)		(Neuromotor activity)

This example illustrates two powerful aspects of genetic analysis. First, we can follow microscopic changes in the DNA such as the temperature sensitive mutation in the *Shibire* gene 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

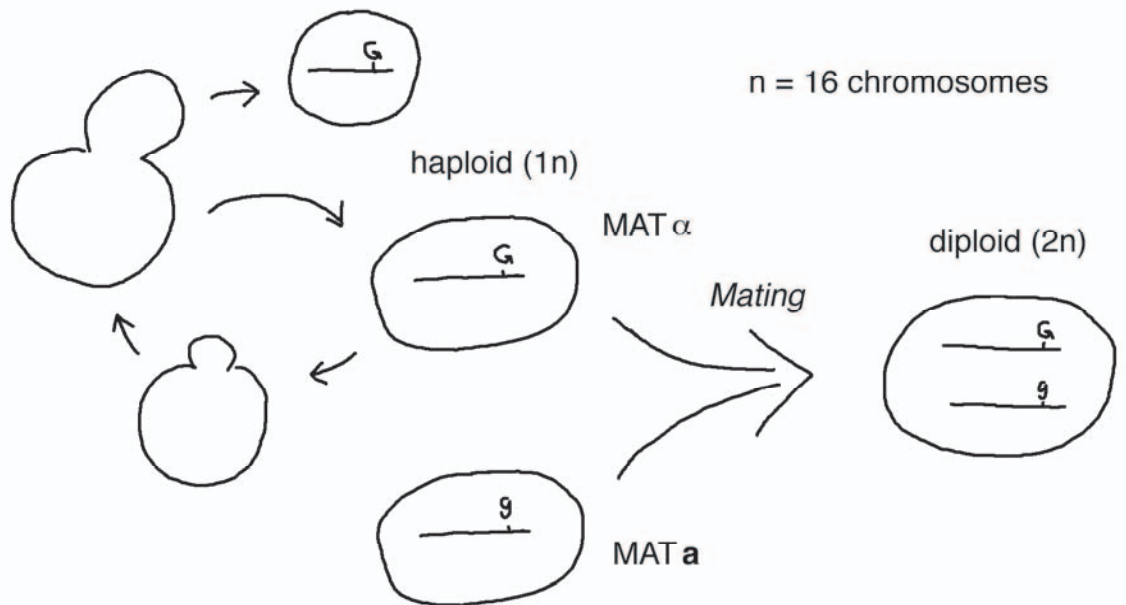
**Wild-type:** defined standard genotype

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.

## Lecture 2

In this lecture we are going to consider experiments on yeast, a very useful organism for genetic study. Yeast is more properly known as *Saccharomyces cerevisiae*, which is the single-celled microbe used to make bread and beer. Yeast can exist as haploids of either mating type **a** (**MAT<sub>a</sub>**) or mating type **α** (**MAT<sub>α</sub>**). Haploid cells of different mating type when mixed together will mate to make a diploid cell.



Haploids and diploids are isomorphic - meaning that a given mutation will cause essentially the same change in haploid and diploid cells. This allows us to look at the effect of having two different alleles in the same (diploid) cell.

All yeast needs to grow are salts, minerals, and glucose (minimal medium). From these compounds, yeast cells can synthesize all of the molecules such as amino acids and nucleotides that are needed to construct a cell. The synthesis of complicated molecules requires many enzymatic steps. When combined, these enzymatic reactions constitute a **biochemical pathway**

Consider the pathway for the synthesis of the amino acid histidine.



Enzyme:      (1)      (2)      (3)      (4)

Each intermediate compound in the pathway is converted to the next by an enzyme. For example, if there is a mutation in the gene for enzyme 3 then intermediate **C** can not be converted to **D** and the cell can not make histidine. Such a mutant will only grow if histidine is provided in the growth medium.

This type of mutation is known as an **auxotrophic mutation** and is very useful for genetic analysis.

	growth on minimal	growth on minimal + histidine
His <sup>+</sup> (wild-type)	+	+
His <sup>-</sup>	-	+

**Phenotype:** All traits of an organism (with an emphasis on trait under investigation)

**Homozygote:** diploid with two like alleles of same gene

**Heterozygote:** diploid with two different alleles of same gene

**Recessive Allele:** trait not expressed in heterozygote

genotype	phenotype	Mate to :	diploid genotype	diploid phenotype
MAT <sub>a</sub> His3 <sup>-</sup>	His <sup>-</sup>	MAT <sub>α</sub> His3 <sup>-</sup>	His3 <sup>-</sup> /His3 <sup>-</sup>	His <sup>-</sup>
MAT <sub>a</sub> His3 <sup>-</sup>	His <sup>-</sup>	MAT <sub>α</sub> His3 <sup>+</sup>	His3 <sup>-</sup> /His3 <sup>+</sup>	His <sup>+</sup>

Based on the His<sup>-</sup> phenotype of the His3<sup>-</sup>/His3<sup>+</sup> heterozygote, we deduce that His3<sup>-</sup> is recessive to wild type.

Let's consider a different kind of mutation giving resistance to copper that occurs in a gene known as CUP1.

genotype	phenotype	Mate to :	diploid genotype	diploid phenotype
MAT <sub>a</sub> Cup1 <sup>r</sup>	copper resistant	MAT <sub>α</sub> Cup1 <sup>+</sup>	Cup1 <sup>r</sup> /Cup1 <sup>+</sup>	copper resistant

**Dominant Allele:** trait is expressed in heterozygote

Cup1<sup>r</sup> is dominant to wild-type (Cup1<sup>+</sup>).

The terms dominant and recessive are simply shorthand expressions for the results of particular experiments. If someone says a particular allele is dominant that means that at some point they constructed a heterozygous diploid and found that the trait was expressed in that diploid.

Note: Sometimes an allele will have more than one phenotype and may be recessive for one and dominant for another. In such cases, the phenotype must be specified when one is making statements about whether the allele is dominant or recessive. Consider for example, the allele for sickle cell hemoglobin in humans designated Hbs. Heterozygous individuals ( $Hb^S/Hb^A$ ) are more resistant to malaria, thus Hbs is dominant for the trait of malaria resistance. On the other hand,  $Hb^S/Hb^A$  heterozygotes do not have the debilitating sickle cell disease, but  $Hb^S/Hb^S$  homozygous individuals do. Therefore,  $Hb^S$  is recessive for the trait of sickle cell disease.

Once we find out whether an allele is dominant or recessive, we can already infer important information about the nature of the allele. The following conclusions will usually be true.

**Recessive** alleles usually cause the loss of something that is made in wild type

**Dominant** alleles usually cause increased activity or new activity

It turns out that the  $Cup^r$  allele actually carries more copies of the gene for a copper binding protein and therefore increases the activity of the gene.

Last lecture we defined the gene structurally as the DNA needed to encode a protein. We can now define a gene in a new way based on its function. Using the phenotypic difference between wild type and a recessive allele we can use a **Complementation test** to determine whether two different recessive alleles are in the same gene.

Say you isolate a new recessive histidine requiring mutation that we will call  $HisX^-$ . In principle, this mutation could be in  $His3$  or it could be in any of the other genes in the histidine biosynthetic pathway. In order to distinguish these possibilities we need a test to determine whether  $HisX$  is the same as  $His3$ .

To carry out a complementation test, one simply constructs a diploid carrying both the  $His3^-$  and  $HisX^-$  alleles.

An easy way to do this would be to mate a  $MATa$   $HisX^-$  strain to a  $MAT\alpha$   $His3^-$  strain.

possibility	genotype of diploid	phenotype of diploid	complementation
$HisX = His3$	$His3^-/His3^-$	$His^-$	<b>No</b>
$HisX \neq His3$	$His3^-/His3^+, HisX^-/HisX^+$	$His^+$	<b>Yes</b>

Having performed this test, if the two mutations don't complement we conclude that they are in the same gene. Conversely, if they do complement we conclude that they are in different genes.

This test only works for recessive mutations. Think about what the outcome would be if HisX<sup>-</sup> were dominant.

The complementation test can be thought of in the following way. If I have an allele with an observable phenotype whose function can be provided by a wild type genotype (i.e., the allele is recessive) — I can ask whether the function that was lost because of the recessive allele can be provided by another mutant genotype. If not, the two alleles must be defective in the same gene. The beauty of this test is that the trait can serve as a read-out of gene function even without knowledge of what the gene is doing at a molecular level.

**Gene:** The fundamental unit of heredity which can be defined in three ways. i) A gene can be defined in molecular terms as a segment of DNA carrying the information necessary to express a complete protein or RNA molecule including the promoter and coding sequence. ii) A gene can be defined by function with a group of recessive mutations that do not complement each other. iii) A gene can be defined by position with a single-locus segregation pattern in a cross between lines with different alleles. Examples are a 1:3 phenotypic ratio in the F<sub>2</sub> generation in a cross between diploid organisms or a 2:2 segregation pattern in yeast tetrad analysis.

**Alleles:** Different versions of the same gene.

**Locus:** The site on a chromosome where a gene is located. Usually defined by recombinational mapping relative to neighboring loci.

**Genotype:** The allelic constitution of an individual, usually with emphasis on the gene or genes under examination.

**Phenotype:** All of the traits or characteristics of an organism, usually with emphasis on traits controlled by the gene or genes under examination.

**Wild-type:** A standard genotype that is used as a reference in breeding experiments. Note that for human crosses there is no standard genotype and the concept of wild-type is therefore not meaningful.

**Haploid:** A cell or organism with one set of chromosomes (1n).

**Diploid:** A cell or organism with two sets of chromosomes (2n).

**Homozygous:** The condition of having two like alleles in a diploid.

**Heterozygous:** The condition of having two different alleles in a diploid.

**Dominant allele:** An allele that expresses its phenotypic effect or trait in the heterozygous state.

**Recessive allele:** An allele whose phenotypic effect or trait is not expressed in a heterozygous state.

**Incomplete dominance:** The case where a heterozygote expresses a phenotype intermediate between the corresponding homozygote phenotypes.

**Complementation test:** A test of gene function where two genotypes with recessive alleles are combined by a cross to test whether the genotype of one parent can supply the function absent in the genotype of the other parent.

**F<sub>1</sub>:** First generation produced by interbreeding of two lines.

**F<sub>2</sub>:** Generation produced by interbreeding of F<sub>1</sub> individuals.

**Incomplete penetrance:** Cases where certain alleles are not always expressed to give observable traits because of other environmental or genetic influences.

**True-breeding:** Refers to a line of individuals that on intercrossing always produce individuals of the same phenotype. This can almost always be taken to mean that the individuals are homozygous at all loci (The major exception is sex chromosome differences between males and females).