

7.013 Recitation 9 – Spring 2018

(**Note:** The recitation summary should NOT be regarded as the substitute for lectures)

Summary of Lectures 15 (3/9) and 16 (3/12):

Shuttle vector: This is a plasmid that can self replicate in two different species i.e. shuttle vector with bacterial and yeast ori should be able to replicate in bacteria and yeast. This can be of great significance since it allows you to clone and amplify a eukaryotic gene in bacteria and then express it in eukaryotic system.

DNA sequencing: DNA sequencing is the technique by which the nucleotide sequence of DNA is determined. DNA sequencing is simply the process of doing DNA replication in the presence of all four nucleotides (dNTPs), and a small percentage of some “dideoxynucleotides (2', 3' di-deoxynucleotides or ddNTPs) that act as chain terminators and do not allow for replication to proceed any further. This technique simply requires you to combine in a test tube: the original template DNA, the four deoxynucleotides, DNA polymerase, a primer, and then a small concentration of “ddA” (colored red), “ddT” (colored blue), “ddC” (colored yellow), and “ddG” (colored green). Most of the time, the deoxynucleotides (dNTPs) will be added and the template will be extended normally. Occasionally, a ddC (for example) will be put in across from a G in the template. This will stop replication and label the fragment yellow. You allow the reaction to proceed and then load all of the replication products onto a gel that separates them by size. The sequence can then be read from the result of that gel (in earlier years) or by a machine now.

SANGERS Method of dna sequencing: <https://www.dnalc.org/view/15479-Sanger-method-of-DNA-sequencing-3D-animation-with-narration.html>

Next generation DNA sequencing: <https://www.youtube.com/watch?v=womKfikWlxM>

Genome sequencing: The way that genomes are sequenced is that many copies of the genome are randomly broken into pieces, which are sequenced by sequencing machines. A computer then analyzes those sequences and reassembles them into one long continuous piece (for a bacterium with a single circular chromosome) or several linear pieces (for a eukaryote with multiple chromosomes). Once the genome is sequenced, one can predict where the genes are in the genome by looking for the signatures of genes, like start codons and stop codons, long open reading frames, promoters, and splice sites (if applicable). Each gene in the genome can then be compared to all other sequences using programs such as BLAST. If a new gene shows a large amount of homology to a previously studied gene, the most likely possibility is that the new gene encodes a protein with a similar function to the protein encoded by the previously studied gene.

Find out what your DNA Says (23 and me): <https://www.23andme.com/>

Polymerase chain reaction (PCR): This is the technique of Polymerase Chain Reaction, which is used to make multiple copies of a chosen piece of DNA. PCR is simply the process of doing about 30 rounds of DNA replication in a row. This technique requires you to combine in a test tube: the original DNA, nucleotides, thermostable DNA polymerase (isolated from the *Thermus aquaticus* found in hot springs), and a pair of primers. ***This polymerase lacks the proof reading ability unlike the regular DNA polymerase.*** Then, one puts the reaction mix in a machine that performs multiple cycles of replication, each in a row of three temperatures. The first temperature (95°C approximately) denatures the double-stranded DNA molecule, the second (55°C approximately) allows the primers to base-pair with the template, and the third (70°C approximately) allows the DNA polymerase to generate the newly synthesized strand 5' → 3' starting from the 3'OH of the primer. You can also use PCR to identify the recombinant clone, from the ligation mix, that contains the gene of your interest and this is often called cloning by PCR. It is important to note this can be used to select a gene that has not been conventionally cloned already.

<https://www.dnalc.org/view/15475-the-cycles-of-the-polymerase-chain-reaction-pcr-3d-animation.html>

Single nucleotide polymorphisms (SNPs) and SNP microarrays: In order to understand the genetic bases of different diseases, biologists are now using rapid genotyping technologies to create haplotype maps, which are used to identify the single nucleotide polymorphisms (SNPs) that are linked to genes involved in diseases. The SNPs that differ between individuals are not inherited as independent alleles. Rather, a set of SNPs that are present on a segment of chromosome are usually inherited as a unit. This linked piece of chromosome is called a haplotype. SNPs can be anywhere in the genome - in between the coding regions of the gene, in introns or in intergenic regions between two genes. SNP defined haplotypes can often be used as markers to predict disease-associated phenotypes of an individual. These associations have revealed particular haplotypes or alleles that are associated with modestly increased risks for breast cancers, diabetes, arthritis, obesity and coronary heart disease.

All about SNPs: <https://learn.genetics.utah.edu/content/precision/snips/>

SNPs of an individual can be detected by SNP microarrays. The basic principles of SNP array are the same as a DNA microarray. The three mandatory components of a SNP array are:

- Immobilized allele-specific oligonucleotide probes.
- Fragmented DNA sequences of target, labeled with fluorescent dyes.
- A detection system that records and interprets the base pairing or hybridization signal.

A SNP array is a useful tool for studying slight variations between whole genome of individuals. The most important applications of SNP arrays are for determining disease susceptibility and for measuring the efficacy of drug therapies designed specifically for individuals. Each individual has many SNPs. SNP-based genetic linkage analysis can be used to map disease loci, and determine disease susceptibility genes in individuals.

DNA expression microarray: This is a collection of microscopic DNA spots attached to a solid surface. Scientists use DNA microarrays to measure the expression of large numbers of genes simultaneously or to genotype regions of a genome. Each DNA spot contains multiple copies of a specific DNA sequence or probe. A probe can be a short section of a gene or other DNA element that hybridizes with the cDNA sample (called *target*) under high-stringency conditions. Probe-target hybridization is then detected and quantified to determine relative abundance of nucleic acid sequences in the target.

Creating a chimera: In order to create a **chimera** you take the embryonic stem cells (ES cells) from a 3 -5 day old embryo. These cells are pluripotent and can differentiate into any cell type of the body. You can then put these cells into a recipient embryo. The recipient embryo can then be implanted in a pseudo-pregnant female. The newborn produced is a chimera since it has cells of different genetic make-up.

Genes can also be inactivated by homologous recombination: One way to study a gene or protein in order to understand its function is to inactivate the gene so that it is not transcribed or translated into a functional protein. Such manipulation is called a knockout experiment. Homologous recombination is a targeted way to produce knockout mutants. In this case, one uses the recombinant DNA technology to replace the existing allele of a gene in the genome with a mutant allele and then studies the effect of this perturbation under various conditions.

- MIT Tech News gene therapy tested as a way to stop blindness:
<http://www.technologyreview.com/news/519426/gene-therapy-tested-as-a-way-to-stop-blindness/>

- Gene therapy and its success stories:
<https://www.youtube.com/watch?v=H0RvTOF1fEc>
- Mice more powerful than a locomotive: <https://vimeo.com/6541290>

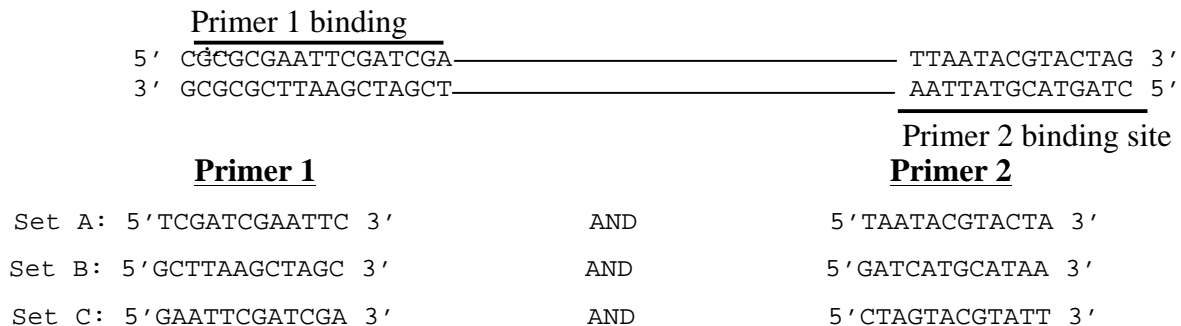
Gene editing by Clustered regularly interspaced short palindromic repeats (CRISPR): CRISPR systems are adaptable immune mechanisms used by many bacteria to protect themselves from foreign nucleic acids, such as those of viruses or plasmids.

Here you introduce a guide RNA (gRNA) that is complementary to and can therefore hybridize with the target gene of interest. You can introduce the gRNA into a cell or an embryo. The target sequence has a Cas-9 nuclease site. Once the gRNA hybridizes with the target DNA, the Cas-9 nuclease makes a double stranded break in the target DNA. The imprecise repair of this double stranded break results in addition or deletion of nucleotides within the target gene thus changing the reading frame.

Genome editing mediated by these nucleases has been used to rapidly, easily and efficiently modify endogenous genes in a wide variety of biomedically important cell types and in organisms that have traditionally been challenging to manipulate genetically. Although the genome-wide specificities of CRISPR-Cas9 systems remain to be fully defined, the power of these systems to perform targeted, highly efficient alterations of genomic sequences and gene expression will undoubtedly transform biological research and spur the development of novel molecular therapeutics for human disease.

Questions

1. You are interested in making many copies of a specific DNA sequence. The sequence to be amplified is flanked by regions whose sequence is given below:



Select the primer pair from above (**Set A/ Set B/ Set C**) that you can use to PCR amplify the gene.

2. You decide to sequence the following piece of DNA.

5'...GTATAGCCTCCGT...-3'

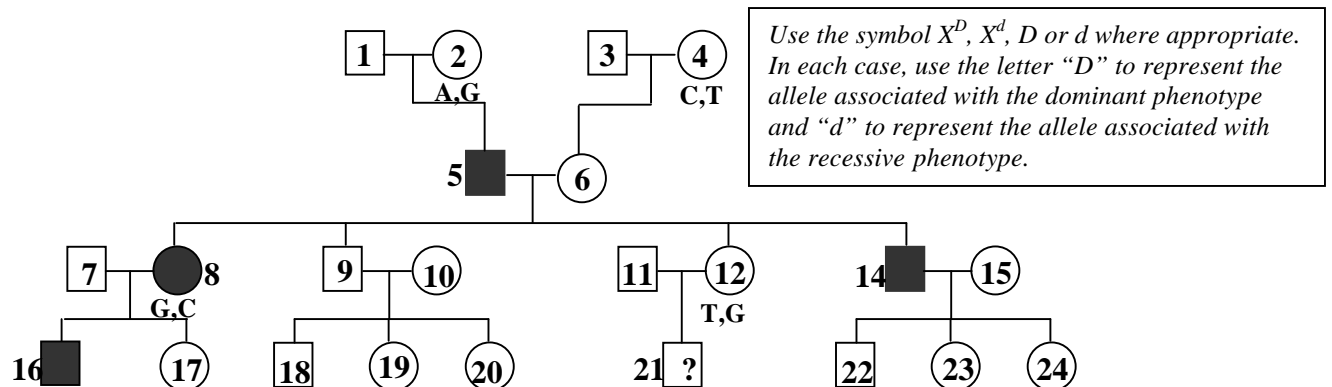
3'...CATATCGGAGGCA...-5'

a) The primer you use for sequencing is 5'-ACG-3'. **Note:** real primers are usually 18-22 nts long). You set up a DNA sequencing reaction with this DNA, this primer, DNA polymerase, dATP, dTTP, dGTP, dCTP, and ddATP. How many different products would form from this reaction?

b) How many nucleotides long would each different product of this reaction be?

c) Write the sequence of DNA that you read from this gel below. Label the 5' and 3' ends of the molecule.

3. The following human pedigree shows the inheritance of a specific disease. **Please note:** The filled squares or circles represent the abnormal phenotype. The individuals marrying into the family do not have the disease-associated allele. Assume that no other mutation arises within the pedigree. Also assume complete penetrance.



a) What is the **most likely** mode of inheritance of this disease? _____

b) The disease shown by the pedigree above is caused by a mutation in Gene D that encodes Protein D. You identify a SNP that is **tightly linked** to Gene D and may be used **as a marker for the disease**. The alleles (A, G, T, C) of this SNP for some individuals are given in the pedigree above.

i. Identify the SNP(s) that is/are tightly linked with the mutant allele of Gene D.

ii. Write the SNP genotypes of the following individuals.

#5: _____

#14: _____

MIT OpenCourseWare
<https://ocw.mit.edu/>

7.013 Introductory Biology
Spring 2018

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.