

NAME _____ KEY _____

Midterm Exam

Applied Molecular Genetics BIOC 571a

October 2, 2001

Please limit your answers to the space provided.

1. (5 pts) Does the OD_{260} value (ultraviolet light absorbance at a wavelength of 260 nm) of a DNA solution increase or decrease when the temperature is raised from 50 degrees to 90 degrees? Explain.

The nucleotide bases are ring structures that absorb light at 260 nm and the amount of absorbance is inversely proportional to the degree of double helix formation in the DNA molecules. At 50 degrees most of the DNA is double stranded and OD_{260} values are low, however, at 90 degrees, most of the DNA is single-stranded and the bases absorb more light so the OD_{260} values are high.

2. (5 pts) Other than the components of a standard hybridization buffer (salt, formamide, phosphate, etc.), what *two* parameters of a solution hybridization reaction could be optimized to ensure that the reaction goes to completion (maximum amount of double strand hybrids)?

The most important parameters of a hybridization reaction that could be optimized are the nucleic acid concentration and time of the hybridization. The Cot value of a hybridization reaction is a measure of $Cot = [DNA/RNA] \times \text{time}$, and therefore, by increasing the nucleic acid concentration *and* by increasing the reaction time, the maximum yield of double strand hybrids will be obtained.

3. (10 pts) The following information can be used to answer the questions below:

BamHI recognition sequence:	G ' GATCC
BglII recognition sequence:	A ' GATCT
MboI recognition sequence:	X ' GATCX
Sau3AI recognition sequence:	X ' GATCX

What is it about the structure-function relationship of most restriction enzymes that explains why they *cannot* cleave single strand DNA containing the correct recognition sequence?

Restriction enzymes bind to DNA as homodimers and cleave double strand DNA. Single strand DNA contains only one binding site and the surface contacts between either a monomer or homodimer with DNA is not sufficient to promote strand cleavage.

Which enzymes, if any, will cleave the recognition sequence G ' GATCT ?

This sequence can be cleaved by MboI and Sau3AI.

What is the predicted frequency of an Sau3AI recognition site in genomic DNA?

The Sau3AI recognition site will occur once in every 256 (4^4) nucleotides.

4. (5 pts.) Explain why the E. coli K-12 strain used in recombinant DNA methods is not a dangerous pathogen, even though it is a close cousin of the highly virulent E. coli strain 0157:H7.

The E.coli K-12 strain is non-pathogenic, primarily because it does not produce toxin proteins like E. coli strain 0157:H7. Other differences include the inability of E. coli K-12 to live outside of the lab because of genetic deficiencies in host defenses (restriction systems), DNA repair and dependence on nutrient rich culture media.

5. (5 pts) Explain why IPTG (isopropylthiogalactoside) is a more potent inducer of the lac operon than is lactose.

IPTG is a nonmetabolizable allolactose analogue that has a high affinity for lac repressor protein. Since IPTG cannot be metabolized by beta-galactosidase, the concentration of IPTG in the media remains constant and thus maintains repressor inhibition.

6. (10pts) Two bacteriophage strains commonly used in applied molecular genetics are lambda phage and M13 phage. Compare and contrast the following attributes of these viruses:

Genome organization and size:

Lambda phage has a ~40kb linear double strand DNA genome. M13 phage contain a small ~5kb single strand DNA circular genome.

Functional role of each virus in phagemid rescue from a lambda ZapII cDNA library:

Lambda phage DNA provides the vector backbone for cDNA inserts which are packaged by in vitro protein packaging extracts. Following E. coli infection, the library is screened by conventional plaque assays. M13 helper phage are used to co-infect a special strain of E. coli with the recombinant lambda phage. Rolling circle replication of the phagemid sequence within the lambda ZapII vector, results in the production of single strand M13 phage which are used to infect E.coli K-12 to generate double-strand plasmid molecules conferring antibiotic resistance.

7. (5 pts) What two key breakthroughs made it possible to generate useful DNA libraries based on lambda phage vectors?

The two key breakthroughs in lambda DNA library applications were the development of phage DNA packaging extracts that permitted in vitro assembly of infectious particles. The other advance was the development of genetic strategies to increase the number of recombinant phage in the library, for example, insert-dependent disruption (or replacement) of phage genes required for lysogeny.

8. (5 pts) What is the difference between a representative cDNA library and a normalized cDNA library? Which one is easier to construct in the lab, and which one is easier to screen?

A representative cDNA library contains all of the transcript sequences converted to cDNA from a particular RNA preparation; abundant transcripts in the RNA prep are represented by a large number of recombinant phage in the library. A normalized cDNA library is made by removing excess cDNA sequences derived from abundant sequences (using Cot hybridization methods). The combined cDNA library contains roughly equivalent numbers of sequences from expressed genes.

Representative cDNA libraries are easier to construct (no additional steps are required), and normalized cDNA libraries are easier to screen (less recombinants need to be screened).

9. (10 pts) Site-directed mutagenesis can be done using dUTP incorporation into an M13 phagemid vector. The underlying principle of this method is to increase the number of bacterial colonies containing mutagenized plasmids. The key steps in this strategy are listed below. Place a number in the blank corresponding to most correct sequential steps 1 through 6; one of the steps is already numbered.

 4 Perform in vitro DNA synthesis

 6 Purify double strand DNA

 1 Transform ung⁻, dut⁻ E. coli with DNA

 5 Transform ung⁺, dut⁺ E. coli with DNA

 2 Purify single strand DNA containing uracil incorporation

 3 Anneal mutagenic oligo to single strand template

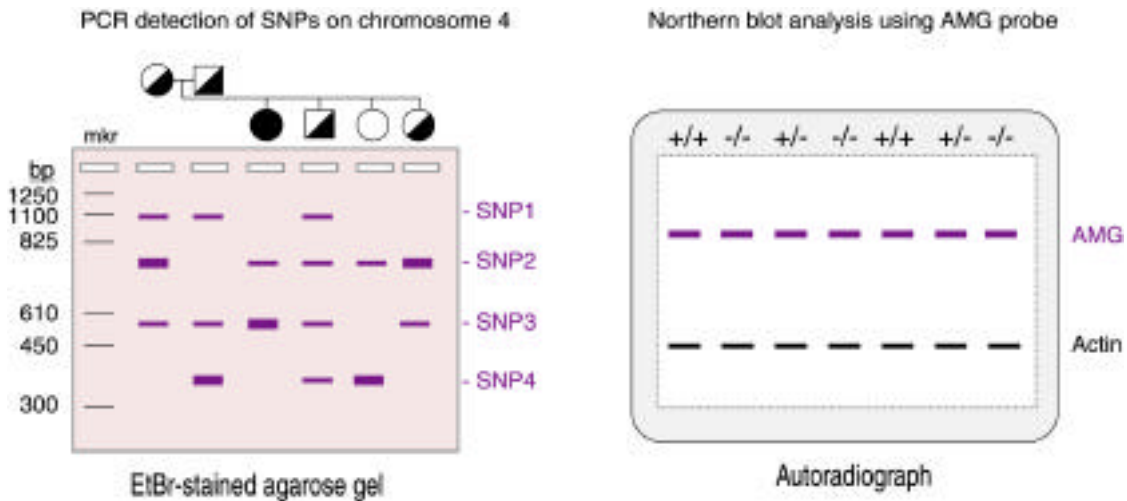
10. What is the difference between using classical genetics and "reverse" genetics in the context of yeast-based molecular genetic studies?

Classical genetics is based on an observed [mutant] phenotype that is then used to isolate the causative gene using linkage maps and physical mapping. Reverse genetics is based on first identifying a segment of DNA thought to contain a gene of unknown function. A mutation is introduced into the gene and then the researcher looks for an observable mutant phenotype.

11. (5pts) Describe the main difference in strategy between how the public effort (NIH directed consortium) and the private sector (Celera) compiled a draft of the human genome sequence.

The public effort was a "top down" strategy using physical and genetic maps to determine which regions need to be sequence, and then overlaying the data to form a contiguous sequence. The private sector used a "bottom up" approach in which the genome was fragmented into small pieces, sequenced as random bits of information and assembled into contigs using sequence overlaps.

12. (10 pts) Use the pedigree analysis and Northern blot data shown below to answer the following questions about an autosomal recessive mutation in the AMG gene resulting in loss of enzyme function in homozygous individuals:



Which single nucleotide polymorphism (SNP) marker is associated with this autosomal recessive disease? Justify your answer.

The SNP3 marker is associated with the disease phenotype because it is the only polymorphism that shows the expected pattern; absent in homozygous normal, one copy in heterozygous and two copies in homozygous disease.

What is the most likely explanation for lack of enzyme function in the diseased individuals and how would you test this directly?

The most likely explanation is a mutation in the protein coding sequence resulting in production of a non-functional or unstable protein. The mutation has no effect on AMG transcript levels and is probably not due to the SNP polymorphism, although it is physically nearby based on linkage. To test this directly, you would need to isolate the gene and show by recombinant protein expression that a point mutation in the coding sequence resulted in synthesis of a defective protein.

13. (5 pts) Why is it difficult to isolate intact mRNA from some tissues such as the pancreas?

Some tissues contain high levels of ribonucleases and it is difficult to inhibit the activity completely.

14. (5 pts) List three *functional* cDNA screening strategies that have been developed to identify cDNA-encoded proteins?

- 1. Protein activity assays (e.g., calcium release by capsaicin receptor)**
- 2. Yeast two-hybrid system is based on protein-protein interactions.**
- 3. cDNA phage display is a method utilizing solid support interactions.**

15. (10pts) A researcher purified large quantities of a specific protein from mesquite bean extracts and asked for help to develop **two independent** cDNA screening strategies using structure based approaches (independent on protein function). Briefly describe two screening strategies beginning with the purified protein. Assume the appropriate cDNA libraries are available.

Strategy #1

Identify as much of the protein amino acid sequence as possible using standard biochemical methods (e.g., Edman degradation or NMR-based methods through a proteomic core service). Based on this information, use a computer algorithm to design degenerate oligonucleotides that cover more than one region of the polypeptide sequence. Screen an appropriate cDNA library (derived from a tissue source known to express the protein) using a sequential probe strategy. Sequence candidate cDNA clones and determine the predicted ORF.

Strategy #2

Use the purified protein as an antigen to raise antibodies in rabbits or mice that can be shown to identify the correct protein in Western blots using extracts from the tissue source. With this characterized antibody, screen an appropriate cDNA expression library (λ ZapII or λ gt11) by antibody cross-reactivity to lacZ fusion proteins contained on membrane filters. Confirm protein cross-reactivity with the antibody using Western blots containing the recombinant protein.