Name __________________________

Gene 5102-034
Examination I
September 29, 2004

For each item, your tasks are highlighted in **bold face**. Please ask for clarification if you do not understand a question! *Sample answers are given in italics.*

I. Basis for principles (35 points)
   A. Principle:
      Some transposable elements have very strict requirements for the sequence at which they insert themselves in the genome. Others have few requirements for the nature of the target sequence.

   Assignment (20 points):
      Explain how, experimentally, the insertion specificity of transposable elements is usually determined OR give an example of how it was determined for a particular transposable element.
      
      The sequences immediately surrounding the transposable element in multiple independent insertions are determined. The sequences are aligned and the alignment is examined to determine whether a consensus sequence exists. If one is found it is assumed to represent the site preference or “insertion specificity” of the transposable element.

   B. Principle:
      Before molecular techniques for examining chromosomes were developed, *Drosophila* researchers were able to establish that the microscopically visible order of “genes” (actually regions containing specific genes) on salivary gland chromosomes was the same as the order of those genes on genetic maps determined by recombination frequencies.

   Assignment (15 points):
      Describe the process of assigning a genetically identified (i.e. mapped by recombination frequency) gene to a region of a salivary gland chromosome.
      
      The gene is mapped genetically to a position on a *Drosophila* chromosome. *Drosophila* lines with recombination breakpoints (as in inversions, translocations) near the genetically mapped position are examined by genetic mapping to determine on which side of the breakpoint the gene is located. The salivary gland chromosomes of these lines are also examined to locate the breakpoints to band locations. The process results in the assignment of the gene to a region between two breakpoints.
Interpretation (35 points total)

A. Facts and observations:

Many ORF (open reading frame) maps show the distribution of translation termination codons in all six reading frames. A large region without termination codons is considered an ORF.

Assignment:
Assuming a 50% G+C content of a DNA and a random nucleotide sequence, calculate the number of termination codons expected in 1,000 nucleotides of one reading frame in an organism using the universal genetic code. Show calculation for partial credit in case of arithmetic error. (5 points)

1000 nt have 333 codons. Since 3 of 64 codons are stop codons, there should be \((3 \times 333)/64\) stop codons in one frame of 1000 nt, assuming all codons occur with equal frequency.

Polypeptides synthesized from prokaryotic messenger RNAs are usually shorter than expected from the number of nucleotides in the ORF (as defined above).

Assignment:
What is the likely explanation for this observation? (5 points).

The region coding for a polypeptide is defined by a start codon AND the first following stop codon that is in frame, while an ORF is defined only as the distance between stop codons. The former is thus a shorter region.
B. Facts and observations:

In the first experimental use of *Drosophila melanogaster* P transposable elements as a vector to create transgenic flies, the xanthine dehydrogenase gene encoded by the wild type *rosy* locus was used as test gene. A homozygous *rosy* mutant line, recognizable by abnormal pigmentation of the eye was used as recipient of the transgene. Appropriate progeny were scored for eye color and for banding pattern in Southern blotting of a *Sal* I digest of the fly DNA and hybridization with a *rosy* gene probe. *Sal* I does not have a recognition site within the gene whose transfer was attempted. Two pattern types were found. One, with a single *Sal* I band at 7.2 kbp was also present in the parent recipient line. The other contained one or more bands additional to the 7.2 kbp band. Flies with the former pattern type had mutant eye color, while flies with the latter had wild type eye color.

**Assignment (15 points):**

What important conclusion could the experimenters draw from this result?

*That the presence of the introduced wild type *rosy* gene correlated with the restoration of the wild type phenotype*

C. Facts and observations:

An interesting mutation occurred in an organism under study. The mutation appeared to be the result of insertion of a 0.5 kbp fragment of DNA in the affected gene. Sequences of the wild type and mutant genes were obtained and are depicted below. The “…..” indicates additional nucleotides.

Wild type sequence
CGACCGCTTTGGCCGCGCCCAGTCCTGCTCGCTTCGCTACTTGGAGCC

Mutant sequence
CGACCGCTTTGGCCGCGCCCAGTCCGACATCACCG……….CGGTGATGTCAGTCCTGCTCGCTTCGCTACTTGGAGCC

**Assignments:**

1. Based on this information, present a reasonable hypothesis for how the mutation arose. (2 points)

   *Transposable element insertion*

2. What observations about the sequence lead you to this hypothesis? (5 points)

   Size consistent with sizes of transposable elements; the presence of an inverted terminal repeat; the duplication of a target sequence (CAGTCC)

3. What additional tests are needed to support the hypothesis? (3 points)

   *Demonstration that other such mutations exhibit similar structures.*
III. Experimentation (30 points)

A. In the papers we studied, three recombinant DNA vectors were used: sep6, pBR322 and M13mp8. Each was used for a different purpose.

Assignment:
1. Match the purpose with the vector below

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Answer</th>
<th>Vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create a library of DNA fragments</td>
<td>C</td>
<td>A. pBR322</td>
</tr>
<tr>
<td>Subclone DNA fragments</td>
<td>A</td>
<td>B. M13mp8</td>
</tr>
<tr>
<td>Sequence DNA fragments</td>
<td>B</td>
<td>C. sep6</td>
</tr>
</tbody>
</table>

2. Briefly explain why each match represented the best choice of vector for that application at the time (early 1980’s) that the research was done.

Library: largest insert size; most efficient cloning to get a large library size
Subclone: easy to produce large quantities of ds DNA with inserts of manageable size.
Sequence: Single stranded DNA was preferred template for dideoxynucleotide sequencing.
B. Each method of gene isolation listed in the left column below requires circumstances listed in the right-hand column.

**Assignment:**
*For each left-hand column item, choose the best right hand item. Right-hand items may be used 0 to 5 times.*

<table>
<thead>
<tr>
<th>Method</th>
<th>Circumstances</th>
</tr>
</thead>
</table>
| Design a degenerate oligonucleotide hybridization probe from a single amino acid sequence. | **D**  
A. Molecular markers mapped genetically to flank the target gene and be close to it |
| Screening of a lambda library expressing introduced genes as polypeptides with an antibody against the protein encoded by the target gene. | **D**  
B. A similar gene has been isolated from a very closely related organism |
| Hybridization of a DNA probe to a genomic library                      | **B**  
C. Similar genes have been isolated from several distantly related organisms. |
| Tagging mutagenesis                                                    | **F**  
D. No related gene is available, but large quantities of purified protein are available |
| Chromosome walking                                                     | **A**  
E. The complete nucleotide sequence of the gene is available through a genome project |
| Direct PCR amplification                                               | **E**  
F. Mutation of the gene creates a readily identifiable phenotype. |