Gene 5102-06
Final Examination
December 14, 2006

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

I. Basis for principles (35 points)

A. Principle:
   In addition to the information specifying the order of addition of amino acids into a polypeptide, the protein encoding gene region contains many other signals that function in the correct production of the encoded polypeptide. For eukaryotic RNA polymerase II transcribed genes these include (in alphabetical order):
   - Branch point signal
   - Coding sequence (2)
   - Initiation codon
   - InR
   - Insulator element (2)
   - Polyadenylation signal
   - Silencer
   - Splice acceptor
   - Splice donor
   - Splicing enhancer
   - TATA box
   - Termination codon
   - Transcription enhancer

Assignment (9 points):
Arrange these elements linearly in a way that this particular gene will be highly expressed (produce large quantities of the encoded polypeptide) while a neighboring gene will not. (2) means that two of these should be used in your arrangement.

B. Principle:
   Many, but not all, signal elements function in an orientation-specific way.

Assignment (8 points):
From the list in A, identify the signal elements whose function is orientation independent.
C. Principle:
Single nucleotide insertion or deletion mutations often cause loss of gene function.

Assignment (9 points):
Identify where in your gene diagram single nucleotide insertion or deletion mutations are most likely to cause loss of function because they disrupt the relative spacing of one element to another.
D. Hypothesis:
Mitochondria arose by engulfment of bacteria related to *Agrobacterium* and *Rhizobium* followed by extensive loss of genes from the engulfed genome.

Assignment (9 points):
What observations support the above hypothesis?

II. Interpretation (35 points total)

A T-DNA insertion line of Arabidopsis thaliana was identified that when homozygous gave viable plants but with severely altered appearance (Fig. 1). A single insertion of T-DNA was present. It was located in the *fas1* gene. This gene encodes a component of CAF-1 (chromatin assembly factor 1), a factor that is bound to PCNA, a protein at replication forks. The *fas1*-4 allele is recessive. The following further observations were made (Kirik et al. 2006 Plant Cell 18:2431):

A. Facts and observations:

Homozygous *fas1*-4 plants were crossed with a line N1IC4 651 that contains two partial copies of the beta-glucuronidase gene (GUS) in inverted orientation (Figure 2). Progeny of the resulting heterozygous plants were tested for the numbers of spots or sectors exhibiting beta-glucuronidase activity.

Assignment (8 points):
Provide a hypothesis regarding the mechanism accounting for the tissue specificity of P element transcript splicing.
B. Facts and observations:
   • *Tc1* is a 1.6 kbp transposable element of the nematode *Caenorhabditis elegans*.
   • Molecular methods can recover nematodes whose genomes contain *Tc1* insertions in a gene of interest without selecting for a mutant phenotype.
   • Most such insertion mutants appear phenotypically normal when homozygous, even though homozygous null mutations (such as deletions) in the same gene have severe phenotypes.
   • Northern blots probed with gene specific probes show, in addition to a band that is 1.6 knt larger than the wild type transcript, several other bands, including one that has the same apparent size as the wild type transcript.

Assignment (9 points):
Provide a mechanistic hypothesis explaining the absence of a phenotypic effect of these *Tc1* insertions.

C. Facts and observations:
   • Comparisons of substitutions between sequences of mitochondrial genes of different *Drosophila* species show that transitions are about equal to transversions in frequency.
   • In similar comparisons between strains of *D. yakuba*, transitions considerably outnumber transversions.

Assignment (9 points):
Provide a hypothesis explaining the difference between the transition/transversion ratios in the two situations.
D. Facts and observations:

Protein-coding genes in the mitochondrial genome of *D. melanogaster* are interspersed with tRNA genes with few extragenic nucleotides between genes.

Assignment (9 points):
Provide a hypothesis about how such an arrangement was selected for in evolution of these genomes.

III. Experimentation (30 points)

A. Gene identification
- Your have isolated several *D. melanogaster* mutants that have indistinguishable phenotypes.
- The mutations map to the same region of the *D. melanogaster* chromosome 2.
- In the DNA sequence of this region, you identify a candidate gene likely to have been the site of the mutations.

Assignment (7.5 points):
Describe two separate approaches that can provide convincing support that mutations in the candidate gene are indeed responsible for the mutant phenotypes. List the experimental steps required to carry out the approaches.
B. Genetic engineering

- Linear DNA molecules introduced into host cells integrate at random in host chromosomes.
- The plasmid depicted below, when linearized at the unique restriction site B, stably transforms a small proportion of target cells. Different clones of these cells have widely differing expression levels of the green fluorescent protein (GFP).

(E = enhancer, P= promoter, GFP = gene for GFP, HygR = gene for hygromycin resistance, A B C D = unique restriction sites in multiple cloning site, ori = origin of replication in *E. coli*, ampR = ampicillin resistance gene.

**Assignment (7.5 points):**
Describe how you would modify the plasmid DNA to reduce the variation in expression levels on transformation.
C. Applying methods

Three methods (among others) for relatively precise location of certain nucleotides on DNA sequences use, for detection, separation of polynucleotide chains on nucleotide sequencing gels (gels that allow one nucleotide length differences to be detected). These methods are:

- Primer extension
- S1 nuclease protection
- DNase I footprinting.

Assignment:

For each each of the three purposes below, choose the most appropriate method. Briefly describe the method (using labeled diagrams may save time). Sketch what the result will look like. Describe how the result you sketch should be interpeted (7.5 points). Each method should be chosen just once.

1. Identification of the site of cleavage of a DNA sequence by a transposase.

2. Location of splice junctions.

3. Identification of the nucleotides to which a repressor protein binds.
D. Protein function

Imagine a protein, PSI, (and its gene) identified as interacting with sequences (F1) whose mutation allows splicing of the intron between ORFs 2 and 3 of the P element transposase gene. The hypothesis is that PSI binding to F1 prevents splicing in somatic cells.

You have available:
- Cultures of somatic cells of Drosophila
- A PSI cDNA clone with two identical tandem inserts, except that one insert is inverted with respect to the other. The clone is in a plasmid vector that has promoters for T7 and SP6 RNA polymerases flanking the insertion point.
- The RNA polymerases.
- A good assay for production of active P transposase,
- Common tissue culture and molecular biological tools

Assignment (7.5 points):
- How would you use the tools to test the hypothesis?
- What result would you expect if the hypothesis is correct?
- What result would you expect if the hypothesis is wrong?