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:
      Genes in microscopically observed metaphase chromosomes are in the same linear order as they are on the genetic map of one linkage group.

   **Assignment (10 points):**
      Describe how the linear order of genes can be determined experimentally for the microscopically observed chromosomes.

      Describe how their order can be determined experimentally for the genetic map.

   B. Principle:
      Chromatin is organized in domains. The genes in each domain are uniformly repressed or uniformly not repressed

   **Assignment (5 points):**
      Give one experimental observation that supports this principle.
C. Principle:
    The sequence of A, C, G and T nucleotides in DNA is not the only carrier of heritable genetic information

Assignment (5 points):
    Describe one example that supports this principle.

D. Principle:
    Current methods for human DNA typing using simple tandem repeats (STR) have several essential features.

Assignment (10 points):
    Explain why each one of the following is important:

1. Each STR used is present at only one locus in the human genome.

2. Multiple loci are used.

3. The length of the simple sequence that is repeated is 3 nucleotides or more.

4. For best results, the multiple loci are on different chromosomes.
E. Principle:
For high-level protein production in one organism from a cloned cDNA from another organism, it may be necessary to alter many of the codons in the DNA.

Assignment (5 points):
Explain why such alteration of codons may be necessary.

II. Interpretation (35 points total)
A. Facts and observations (Mutation):
In experiments meant to discover the function of a specific gene, two approaches were used.

- In one, the knock-out strategy, homologous recombination at regions flanking the two ends of the gene was used to replace the gene with a selectable marker. Breeding resulted in a transgenic organism homozygous for the replacement.
- In the other, the RNA silencing strategy, non-homologous recombination resulted in the introduction into the chromosome of a DNA construct in which a constitutive promoter controls synthesis of an RNA containing an inverted repeat of the specific gene sequence.
- The knock-out organisms showed only a mild growth deficit, while the RNA silencing organisms were severely stunted.
- The same investigators also tried both approaches with a different specific target gene. In this case, the results were opposite to those above. The knock-out had severe defects, while the RNA silencing was almost wild type in appearance.

Assignment (10 points):
Provide a hypothesis about the targeted genes that can explain the difference in result between the knock-out and RNA silencing approaches and the difference in result between the two genes.
B. Facts and observations (Bioinformatics):

Bioinformatics can provide information useful in constructing hypotheses about the structure and function of genes.

Assignment (15 points):

In the left-hand column find a description of several bioinformatic operations. For each, choose the most appropriate potential result (only one) from the list in the right column. If you are unsure of your answer, give an explanation of your choice. Note more potential results are provided than operations.

<table>
<thead>
<tr>
<th>Operation</th>
<th>Choice</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Search of an EST insert sequence library for ATAA sequences near the 3' end of library inserts.</td>
<td>____</td>
<td>A. Predicted positions of introns</td>
</tr>
<tr>
<td>2. Examination of cDNA sequences for DNA sequence regions lacking TGA, TAG or TAA sequences offset from one another by a multiple of three.</td>
<td>____</td>
<td>B. Predicted eukaryotic 3' RNA processing site</td>
</tr>
<tr>
<td>3. Search bacterial DNA sequences for short inverted repeats followed 3' by T-rich stretches.</td>
<td>____</td>
<td>C. Possible $\rho$-dependent terminator</td>
</tr>
<tr>
<td>4. BLAST search of a database of protein sequences predicted for one organism for sequences with significant similarity to a single sequence of a protein enzyme from the same organism.</td>
<td>____</td>
<td>D. Open reading frames</td>
</tr>
<tr>
<td>5. Comparison of sequences in an EST insert sequence library with sequences from the genome of the same organism from which the EST library was constructed.</td>
<td>____</td>
<td>E. Possible $\rho$-independent terminator</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F. Probable products of a multigene family</td>
</tr>
</tbody>
</table>
C. Facts and observations (Gene isolation and identification)

A cosmid clone of a segment of the genome of MFO has been isolated and its sequence determined. It is suspected to contain the gene for the enzyme examinase. There is a simple in vitro assay for the activity of examinase. The following list contains some of the items of evidence that led to confirmation that the segment indeed contains the gene for the examinase enzyme. However, each piece of evidence by itself was not sufficient evidence that the gene was present in the segment.

Assignment (10 points):
For each item of evidence, explain why that evidence, by itself, is not sufficient. Your explanation should consist of a feasible alternative explanation for the finding (independent of the other results). Note there are 5 items, 3 on the following page.

1. A BLASTx search of the general protein database using the segment sequence as query identified, as top hit, an examinase enzyme gene from a distantly related organism with an E-value of 0.0001.

2. By screening a cDNA library with a probe obtained from the cosmid segment, a cDNA clone was isolated. The sequence of the insert reveals that it consists of a concatenation of several parts of the sequence of the cosmid sequence.
3. A Southern blot of a gel electrophoretic separation of restricted MFO genomic DNA was probed separately with the cosmid insert and the cDNA insert. In both cases, a single band of intensity consistent with its being in single copy in the genome was found. The band was of the same size in both probings.

4. The cDNA insert was transferred to an *Escherichia coli* expression vector to produce an expression construct. Homogenates of the host *E. coli* cells (without introduction of the construct) did not have measurable examinase enzyme activity. Homogenates of *E. coli* cells transformed with the construct exhibited measurable examinase enzyme activity.

5. A knock-out MFO cell line was constructed in which the putative gene was replaced with a selectable marker by homologous recombination at regions flanking the gene. The knock-out line did not have active examinase enzyme while the control parental line exhibited substantial activity of the enzyme.
III. Experimentation (30 points)

A. Gene regulation

Assignment (15 points):
In the left column are some common objectives in gene regulation investigations. For each, choose the most appropriate method (only one) from the list in the right column. If you are unsure of your answer, give an explanation of your choice. Note more methods are provided than tasks.

<table>
<thead>
<tr>
<th>Objective</th>
<th>Choice</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Confirm that an isolated auxiliary transcription factor binds specifically to a given short target DNA sequence.</td>
<td>_____</td>
<td>A. Bioinformatics</td>
</tr>
<tr>
<td>2. Test whether an increase in expression level of a gene (in one physiological state relative to another) is due to an increase in transcription or to greater mRNA stability</td>
<td>_____</td>
<td>B. Microarray analysis of gene expression</td>
</tr>
<tr>
<td>3. Determine which genes of a well-studied organism appear to be coregulated.</td>
<td>_____</td>
<td>C. DNA mobility shift analysis</td>
</tr>
<tr>
<td>4. Given a list of genes from a well-studied organism that appear to be coregulated, identify possible regulatory elements in their DNA sequences.</td>
<td>_____</td>
<td>D. Chromatin immunoprecipitation</td>
</tr>
<tr>
<td>5. Given an antibody to an auxiliary transcription factor, produce a DNA clone library containing sequences to which it binds.</td>
<td>_____</td>
<td>E. Nuclear run-off analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F. Southern blot hybridization</td>
</tr>
</tbody>
</table>
B. Gene fusion

“Gene fusions”, where parts of two distinct genes are joined together by recombinant DNA techniques, are widely used in molecular genetics.

Assignment (15 points):

For each of the following experimental purposes, describe a gene fusion that is used in achieving the purpose (well-labeled diagrams are acceptable) and describe its role.

1. Recombinant protein (a protein made from an in vitro constructed DNA) purification.

2. Yeast two-hybrid screening

3. Transcription initiation regulatory region analysis

4. Identifying protein targeting signals

5. Co-immunoprecipitation binding analysis in absence of antibody to the protein of interest.