I. A) Tell whether β-galactosidase and permease activity would be regulated, constitutive or absent for each of the lac operon strains or partial diploids shown below.

<table>
<thead>
<tr>
<th>GENOTYPE</th>
<th>β-GAL'ASE</th>
<th>PERMEASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 P I P O Z Y A</td>
<td>Abs</td>
<td>reg</td>
</tr>
<tr>
<td>2 P I^s P O Z Y A</td>
<td>abs</td>
<td>abs</td>
</tr>
<tr>
<td>3 P I P O Z Y A</td>
<td>const</td>
<td>abs</td>
</tr>
<tr>
<td>4 P I P O^- Z Y A</td>
<td>const</td>
<td>const</td>
</tr>
<tr>
<td>5 P I P O Z Y A/ P I P O Z Y A</td>
<td>reg</td>
<td>reg</td>
</tr>
<tr>
<td>6 P I^s P O Z Y A/ P I P O^- Z Y A</td>
<td>const</td>
<td>const</td>
</tr>
</tbody>
</table>

B) Which, if any, of the strains would respond differently if the tests were made using lactose rather than IPTG as the inducer? Why?

# 1; permease would be absent since there is no β-galactosidase to make allolactose, the actual inducer.

II. You would like to clone a gene that is expressed only after the "host" organism is exposed to a virus.

a) In some cases, it may be easier if the host is a eukaryote than a prokaryote. How can this be?

The induced messages would have poly-A tails which can make isolation of mRNA simpler

b) Outline a series of steps, including enzymes and components that would allow you to clone potential "virus induced" genes into a plasmid vector such as pUC-18.

1) isolate mRBNA using oligo T (or U) columns or beads
2) Use reverse transcriptase to make the first cDNA strand, from oligo dT primer
3) add RNase H, DNA pol, and DNA ligase to make the second strand
4) add tails using TdT, Eco RI (etc.) linkers or Eco-5' tailed primers using DNA ligase
5) digest linkers with Eco RI or amplify from primers with TAQ polymerase
6) digest pUC with Eco RI; treat with CIAP to prevent self religation
7) Mix cDNA and vector with DNA ligase
8) transform into competent E. coli cells.

c) List various features of the plasmid vector that make it useful.

a) origin of replication allows multiple copies per cell
b) antibiotic resistance gene (amp) allows selection of transformed E. coli
c) Multiple cloning site with unique RE sites
d) Lac Z gene for blue-white screening
e) M13 forward and reverse sites flanking MCS for sequencing

d) How will you determine if your clones came from the virus or the host?

Southern blot should work; sequencing might tell

III. Define the following:

"alu" sequence: A SINE family of retro-elements found in human DNA that is about 300 bp repeated 500,000 times and scattered through the genome. It is defined because it has an alu restriction site. It contains 2 copies of a 7s-snRNA gene, only one of which can function.

kinetochore: centromere associated protein that serves as attachment for spindle fibers for chromatid separation

telomere: The end of a linear chromosome that includes multiple copies of a short repeat that can be extended by telomerase

Cot curve: a measure of reannealing rates of melted DNA; the curve indicates the complexity- repeated versus unique sequence DNA
nucleosome a histone ball made of 2 copies each of H2A, H2B, H3 and H4 wound by 2 turns of dsDNA and "tied" with H1.

heterochromatin (both types) condensed, inactive chromatin, constitutive heterochromatin is 1' highly repeated sequences near the centromere or telomeres that is "permanently" condensed. Facultative heterochromatin is condensed in particular situations, such as the extra X in female cells.

satellite DNA short, highly repeated tandem sequences found in the genome originally defined because they produced a satellite band when DNA was purified using density gradient centrifugation

degenerate primer: a PCR primer made with a mixture of bases at specific positions, usually to account for codon degeneracy when using amino acid sequence to clone the respective gene.

LINE element L1: a mammalian retrotransposon about 6Kb in length that is found in 100,000 copies/genome

IV. Show two cells of a diploid organism (2N=4) in anaphase of the first division of meiosis that illustrate Mendel's laws of segregation and independent assortment.

V. Check the following sequences that could be targets for restriction endonucleases. (all are shown in the 5' to 3' direction)

______AAAAAA  ____ACTTCA    X____GAATTG    X____GACGTC
VI. How can pseudogenes that a) include introns and b) those that do not originate? Which is most likely to be found in a gene cluster and why?

Those with introns arise by gene puplication followed by inactivating mutations; those without introns are made via reverse transcriptase of mRNA. The latter can be reinserted almost anywhere so the duplicated copies are most likely to be in gene families.

VI. A cross of two truebreeding white-flowered peas gave F1 progeny that all had purple flowers. When these were self pollinated there were 87 F2 plants with purple flowers and 63 with white flowers.

A) Do a "goodness of fit" test to determine if the F2 data support a hypothesis of 1:1 segregation? Show your work and the basis for your conclusion.

<table>
<thead>
<tr>
<th>Observed</th>
<th>Expected/(1:1)</th>
<th>d</th>
<th>$d^2/e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>87</td>
<td>75</td>
<td>12</td>
<td>1.92</td>
</tr>
<tr>
<td>63</td>
<td>75</td>
<td>12</td>
<td>1.92</td>
</tr>
</tbody>
</table>

Chi squared with 1 df = 3.84, meaning there is almost exactly a 5% chance the data could arise by chance under the hypothesis. This is not grounds for rejecting under normal

B) Do the same for a 9:7 F2 ratio.

<table>
<thead>
<tr>
<th>Observed</th>
<th>Expected/(1:1)</th>
<th>d</th>
<th>$d^2/e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>87</td>
<td>84.375</td>
<td>2.625</td>
<td>0.082</td>
</tr>
<tr>
<td>63</td>
<td>65.625</td>
<td>2.625</td>
<td>0.105</td>
</tr>
</tbody>
</table>

Chi-squared is 0.187 meaning there is between a 0.5 and 0.7 chance of deviations this large or larger under the hypothesis.

C) Make a model that shows how flower color inheritance can be explained in this cross. Include: i) Legends for the gene or genes involved, ii) genotypes of parents and F1s according to your legend, iii) a pathway that would account for the ratio you have predicted.

Since this is an F2, 9:7 is more logical than 1:1;

i) A_ B_ purple; aa B_ white. A_ bb white; aa bb white
ii) AA bb X aa BB parents gives Aa Bb F1

\[ \frac{3}{4} \]

iii  White $\rightarrow$ White $\rightarrow$ purple

VII. When Mendel tested tall F2 plants that could be TT or Tt, he grew out 10 F3 progeny per self pollinated F2 plant. What fraction of the Tt plants would he miss-classify as homozygous?

those that gave all 10 Tall: $\left(\frac{3}{4}\right)^{10}$ or about 5.6 %

VIII. An achondroplastic dwarf male marries a female with the same genotype. If they have 4 children:

a) What is the probability that 3 will be dwarfs and 1 normal?

Since the trait is lethal when homozygous:

\[ \frac{4!}{3!1!}(\frac{2}{3})^3(\frac{1}{3})^1 \]

b) What is the probability that at least one will be a dwarf?

\[ 1 - P(\text{all normal}) = 1 - (\frac{1}{3})^4 = 0.998 \]

IX. In Ayrshire cattle, when purebred mahogany bulls are mated to purebred red cows, the sons are mahogany and the daughters are red. However, when these F1s are inter-mated, 3/4th of the bull calves are mahogany and 1/4th are red, while in the heifers, 3/4th are red and 1/4th are mahogany. What is the simplest explanation for the inheritance of coat color in this breed of cattle? (Include a legend in your answer)

<table>
<thead>
<tr>
<th>Sex influenced inheritance:</th>
<th>Genotype</th>
<th>males</th>
<th>females</th>
</tr>
</thead>
<tbody>
<tr>
<td>R'R'</td>
<td>mahogany</td>
<td>mahogany</td>
<td></td>
</tr>
<tr>
<td>R'R</td>
<td>mahogany</td>
<td>red</td>
<td></td>
</tr>
<tr>
<td>RR</td>
<td>red</td>
<td>red</td>
<td></td>
</tr>
</tbody>
</table>

X. A female Drosophila is heterozygous for a recessive sex-linked lethal that has no associated phenotype. How can this allele be detected? How can it be maintained? only 1/3 of her progeny will be male; can maintain
as heterozygous females, but will have to test several each generation and keep female progeny of those that produce few sons.

XI. In tomato, tall is dominant to dwarf and round fruit is dominant to pear-shape. A cross between tall, round tomato (1) to a dwarf, pear tomato produces 82 tall, round : 78 dwarf, pear : 22 tall, pear : 18 dwarf, round. Use of a different tall, round tomato (2) in the same cross gave 43 tall, pear : 37 dwarf, round : 11 tall, round : 9 dwarf, pear. A) Show the genotypes of tall, round parents 1 and 2.

These are test-crosses that do not give a 1:1:1:1 ratio, indicating linkage; 1)T\textsuperscript{20}R/t r 2) T r/t R

B. Predict the phenotypic ratio in the progeny of a cross between the two tall, round tomatoes. The genes are 20 map units apart therefore:

<table>
<thead>
<tr>
<th>gametes from coupling parent</th>
<th>gametes from repulsion parent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.1 T R 0.4 T r 0.4t R 0.1 t r</td>
</tr>
<tr>
<td>0.4 T R</td>
<td>0.04 0.16 0.16 0.04</td>
</tr>
<tr>
<td>0.1 Tr</td>
<td>0.01 0.04 0.04 0.01</td>
</tr>
<tr>
<td>0.1 t R</td>
<td>0.01 0.04 0.04 0.01</td>
</tr>
<tr>
<td>0.4 tr</td>
<td>0.04 0.16 0.16 0.04</td>
</tr>
</tbody>
</table>

0.54 tall, round; .21 tall, pear; 0.21 dwarf, round; 0.04 tan = dwarf, pear

XII. Predict the frequency of: 1) abc gametes that will be produced from the following tri-hybrid parent if I= 0.333.

<table>
<thead>
<tr>
<th>A 20  b 30 C</th>
</tr>
</thead>
<tbody>
<tr>
<td>a B c</td>
</tr>
</tbody>
</table>

1) if I = 1/3 there are only 2/3 as many DXO gametes as predicted; predicted is 0.2 times 0.3 = 0.06 so we will actually get 4 per 100. 1/2 of these will be ABC and 1/2 will be abc so the answer is 2%

2) f(ABc) gametes these are "parental" if there are 4% DXO gametes, 16% will have SXO in region one and 26% will have SXO in region II. Thus 54% of the gametes have no XO at all, half of these (27%) will be A b C.
XIII   Compare the Lyon hypothesis and genomic imprinting; tell how they are alike and how they differ.

Both deal with gene/chromatin inactivation, probably involving DNA methylation and histone acetylation.  In the Lyon hypothesis, every X chromosome > 1 in cells is inactivated during embryogenesis, with the same X inactive in all subsequent daughter cells.

Genomic imprinting occurs during gametogenesis and inactivates different regions of specific chromosomes in males (sperm) or females (eggs).  It is reprogrammed every generation.

BONUS: Assume a genome has $3 \times 10^9$ bp of DNA and for simplicity that BACs can be made that include random fragments of the genome that average 300 kb.  a) What is the probability that "gene a" which is present as a single copy in the genome will be found on any BAC selected?  
b) If I want to be 95% certain that "gene a" will be included in the library at least once, how big should the library be.  Show your answer as a formula only.

a) $300,000/3,000,000,000 = 1 \text{ in } 10,000$

b) $P(\text{at least one}) = 0.95 = 1-(0.9999)^n$ thus $(0.9999)^n = 0.05$ 

taking the ln of both sides, $n(ln0.9999) = ln(0.05)$ and $n = ln(0.05)/ln(0.9999). = -2.995/-0.0001 = 29,950$ BAC clones