PROBLEM 1

You isolate loss-of-function mutations in three insect genes, *caprine*, *homarine*, and *porcine*, that control body hardness. Loss-of-function mutations in each cause Hard bodies. Gain of function mutations in each cause Soft bodies.

A. Construct a formal genetic pathway for the role of these three genes in the control of body hardness. Explain your reasoning.

B. Predict the phenotypes, if possible, of the other three double mutants listed below.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type</td>
<td>Normal</td>
</tr>
<tr>
<td>caprine(lf)</td>
<td>Hard</td>
</tr>
<tr>
<td>porcine(lf)</td>
<td>Hard</td>
</tr>
<tr>
<td>homarine(lf)</td>
<td>Hard</td>
</tr>
<tr>
<td>caprine (gf)</td>
<td>Soft</td>
</tr>
<tr>
<td>porcine(gf)</td>
<td>Soft</td>
</tr>
<tr>
<td>homarine(gf)</td>
<td>Soft</td>
</tr>
<tr>
<td>porcine(lf); caprine (gf)</td>
<td>Hard</td>
</tr>
<tr>
<td><strong>caprine (lf); homarine(gf)</strong></td>
<td>Hard</td>
</tr>
<tr>
<td>caprine (lf); porcine(gf)</td>
<td>Soft</td>
</tr>
<tr>
<td>homarine(lf); caprine(gf)</td>
<td>___</td>
</tr>
<tr>
<td>homarine(lf); porcine(gf)</td>
<td>___</td>
</tr>
<tr>
<td>porcine(lf); homarine(gf)</td>
<td>___</td>
</tr>
</tbody>
</table>

C. You perform yeast-two-hybrid analysis with the MANATINE, HOMARINE and PORCINE proteins, and find that

*HOMARINE* binds *CAPRINE*  
and  
*CAPRINE* binds *PORCINE*

Propose a model for the function of these three proteins. Explain your reasoning, and predict the outcome of an immunoprecipitation experiment in which you pull down a complex containing the PORCINE protein.
PROBLEM 2

Construct a diagram by placing the following genes into modules and indicate their interactions. Rank (>>:, >, ≥, =) your confidence of the integrity of each module?

Subcellular localization:

C = cytoplasm
M = membrane
N = nucleus

bio..... M
cit...... N
comp. N
elec... C
geo ... C
mech. N
phys.. C
sci..... M
star ... N

Interactions:

G = Genetic
P = Physical

<table>
<thead>
<tr>
<th></th>
<th>bio</th>
<th>cit</th>
<th>comp</th>
<th>elec</th>
<th>geo</th>
<th>mech</th>
<th>phys</th>
<th>sci</th>
<th>star</th>
</tr>
</thead>
<tbody>
<tr>
<td>bio</td>
<td></td>
<td>G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cit</td>
<td>G</td>
<td>P</td>
<td>G,P</td>
<td></td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>comp</td>
<td></td>
<td>G,P</td>
<td></td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>elec</td>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>geo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mech</td>
<td></td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G</td>
</tr>
<tr>
<td>phys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sci</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>star</td>
<td>G,P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Each experiment was done reciprocally.
PROBLEM 3
You isolate a cell type from a C. elegans larvae and analyze its transcripts by DNA sequencing. Your sequence analysis identifies the following genes that are most highly expressed and most specific:

*cdk-1*
*eye-1*
*myo-2*
*hlh-2*
*mlc-1*
*mlc-2*
*mup-2*
*myo-3*
*pat-3*
*pat-4*
*pat-6*
*pat-10*
*tmd-2*
*tni-3*
*tnt-2*
*tnt-3*
*unc-112*

**a.** What type of cell do you hypothesize this is? State how you reached this conclusion.

**b.** Imagine you are able to culture this cell type. You add a small molecule kinase inhibitor, BF100666, that your collaborator has discovered, and profile the cells after 12 hours of exposure to BF100666. You observe the following set of genes most highly expressed. What do you infer?

*eql-2*
*trp-4*
*eql-2*
*eql-10*
*gpa-14*
*goa-1*
*nlp-22*
*flp-7*
*nhr-33*
*itr-1*
*ipp-5*
*dgk-1*
*unc-13*
*ric-3*
*ham-1*
eat-16
ceh-10
lev-1
egl-43
rab-3
PROBLEM 4

You are interested in identifying the protein-binding partners of a little-studied human gene, and turn to yeast-two-hybrid assays for information. A private company provides a service where they will probe a bait of your choice (your gene, fused at the 5’ end to a DNA binding domain) with prey that includes the entire human proteome (fused at the 5’ end to a transcriptional activator domain).

a. Explain at least one possible source of false negatives (in what situation might this assay say there is no interaction, when the two proteins do interact in the human body).

b. Explain at least one possible source of false positives (in what situation might this assay say there is protein-interaction, when the two proteins do not interact in the human body).

c. In the so-called “split ubiquitin” assay, the bait-prey interaction can occur in the nucleus, cytoplasm, or along any membrane. Explain why this assay may give different results from the classic yeast-two-hybrid assay described above (you should be able to answer this from basic knowledge of biology, and you do not need to know specific details of the split ubiquitin assay).

d. To learn more about how your gene of interest is regulated, you create a GFP-fusion transgene in which the promoter of your favorite gene is used to drive expression of a fluorescent protein. You create multiple clonal lines of transgenic human cells in which this transgene has been inserted into a single random position in the genome. You expose each line to an array of different drug treatments and observe the change in GFP expression, which does not match between different clonal lines. Although your lines all contain the same transgene, some treatments have different effects on different lines (for example, drug treatment A may boost GFP expression in line 1, but diminish GFP expression in line 2). Give and explain two reasons this might happen.
PROBLEM 5

You want to identify genetic factors involved in waspophobia, heightened avoidance of any type of wasp. After struggling with many families that purportedly had congenital waspophobia, but later realizing that most of these were clustered environmental effects, you decided to do a large scale case-control study with a candidate gene, WSPA1. You chose this gene because it encodes a Wiskott-Aldrich Syndrome (WAS) related protein, and it sounds a lot like GWAS. (Admittedly, in retrospect this was not a particularly compelling rationale for a focus on this locus.)

You test 1000 individuals of each of two genotypes at the WSPA1 locus and find the following results. You develop a qualitative assay for wasp avoidance using exposure to jars of wasps in a quiet room and measure several aspects of discomfort. You bin volunteers into three groups: low, medium and high wasp avoidance.

<table>
<thead>
<tr>
<th>Allele</th>
<th>Discomfort level (#individuals)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>high</td>
</tr>
<tr>
<td>A</td>
<td>450</td>
</tr>
<tr>
<td>C</td>
<td>350</td>
</tr>
</tbody>
</table>

Which allele of WSPA1, if any is associated with waspophobia.

Explain your statistical reasoning (There are a few ways to set this up).