Problem set questions from Exam 3 – Eukaryotic Gene Regulation, Genome Modifications in Eukaryotes, Population Genetics

Characterizing novel pathways that control the expression of yeast genes

1. You are studying regulation of the yeast enzyme glutamine synthetase (GS), which is encoded by the GLN1 gene. You have isolated two mutants, designated gln2$^-$ and gln3$^-$, that give decreased GS activity. Mating of either gln2$^-$ or gln3$^-$ haploids to wild-type haploids produce heterozygous diploids that show normal amounts of GS expression. When you cross either a gln2$^-$ or gln3$^-$ haploid strain to a gln1$^-$ haploid strain, the resulting diploids show normal expression of GS.

(a) From these experiments, classify the gln2$^-$ and gln3$^-$ mutations in terms of their basic genetic properties (dominant vs recessive, cis vs trans, uninducible vs constitutive), and explain the rationale behind your conclusions. Based on these properties, make a proposal for the nature of the wild-type regulatory functions of the GLN2 and GLN3 genes.

(b) Diagram two different linear models and one parallel model that could illustrate how the GLN1 gene is regulated by the wild-type GLN2 and GLN3 genes.

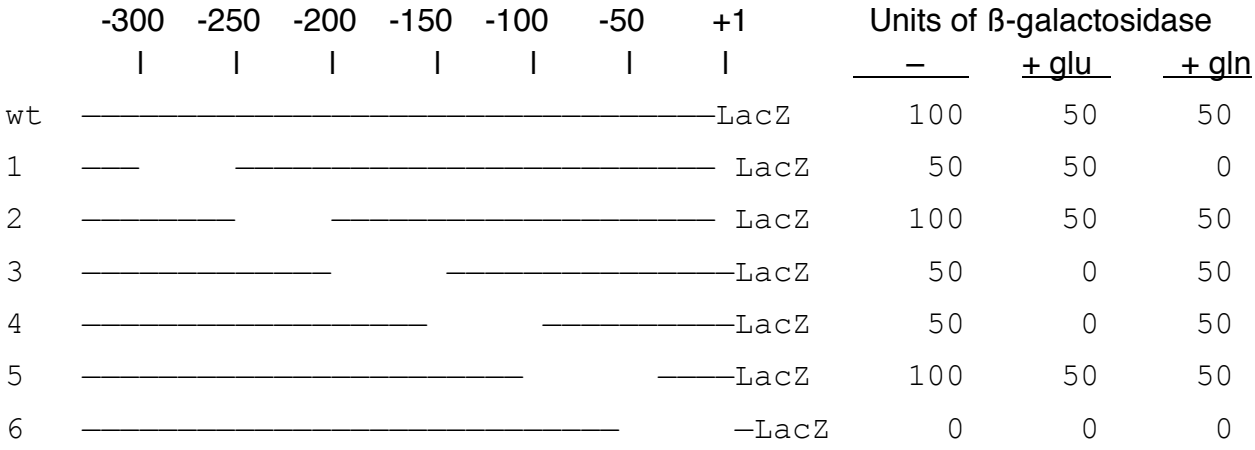
The GLN1 gene shows a rather complex regulation in response to two different amino acids. When either glutamate (glu) or glutamine (gln) is added to the medium, the amount of GS expression diminishes; when both glutamate *and* glutamine are added to the medium, GS expression is shut off completely. The effects of different mutants on the response to glu and gln are shown below.

	Units of GS activity			
	–	+ glu	+ gln	+glu& gln
wild type	100	50	50	0
gln1$^-$	0	0	0	0
gln2$^-$	50	50	0	0
gln3$^-$	50	0	50	0

(c) Which of the models from part **(b)** best fits these experimental results? Diagram a complete model for the regulation of GLN1 that includes the effects of glu, gln, wild-type GLN2, and wild-type GLN3.

(d) Based on your model for part **(c),** how would you expect a gln2$^-$ gln3$^-$ double mutant to behave?

Next, you decide to evaluate the cis regulatory DNA sequences found in front of the GLN1 open reading frame. To do this, you first fuse these regulatory DNA sequences to the LacZ coding sequence, and then place this hybrid gene on an appropriate yeast plasmid. You find that cells carrying this reporter gene express LacZ activity under the same conditions that GS is expressed in wild-type cells. This tells you that the cis regulatory region you have selected contains all of the necessary cis-acting sequences for normal regulation. The figure below shows the effect of six different 50-basepair-long deletions in the cis regulatory region on the amount of ß-galactosidase activity expressed by the reporter gene.



	Units of ß-galactosidase		
	–	+ glu	+ gln
wt	100	50	50
1	50	50	0
2	100	50	50
3	50	0	50
4	50	0	50
5	100	50	50
6	0	0	0

(e) Describe the cis-acting elements in the GLN1 cis regulatory region that are evident from these experiments, giving both their positions and as much of their wild-type function as you can deduce.

(f) How many units of ß-galactiosidase would you expect to be expressed in gln2$^-$ mutant yeast from a reporter gene construct carrying deletion #1…
with neither amino acid added?
with only glu added?
with only gln added?

(g) How many units of ß-galactiosidase would you expect to be expressed in gln2$^-$ mutant yeast from a reporter gene construct carrying deletion #4…
with neither amino acid added?
with only glu added?
with only gln added?

2. Consider a eukaryotic gene regulatory pathway where a small molecule X activates the expression of a reporter gene. You have isolated loss-of-function mutations in two different genes, **A** and **B**, both of which give uninducible expression of the reporter. Genes **A** and **B** are not linked to each other and neither gene is linked to the reporter.

(a) Assuming that the regulatory factors encoded by **A** and **B** act in series, there are two possible orders in which these two regulatory factors can act. Draw out these two models showing the relationships between the wild-type regulatory functions of **A** and **B,** and the reporter. Also be sure to indicate where and how the inducer X acts.

(b) In order to distinguish between the two models from part **(a)**, an epistasis test would be useful. Because the mutations that have been isolated in the **A** and **B** genes have the same phenotype (uninducible), it is not possible to perform an epistasis test using these alleles. Fortunately, you are able to isolate an allele of gene **A** that gives constitutive expression of the reporter. This allele, called $\mathbf{A^S}$, causes a dominant phenotype of constitutive reporter expression. Describe in molecular terms how the allele $\mathbf{A^S}$ affect the normal regulatory function of **A**, given each of your models from part **(a)**.

(c) Assume that you are studying this regulatory pathway in yeast and you wish to perform an epistasis test by constructing the $\mathbf{A^S\ B^-}$ double mutant. To do this, you cross a MAT**a** $\mathbf{A^S\ B^+}$ haploid strain to a MATα $\mathbf{A^+\ B^-}$ haploid strain, and induce sporulation of the resulting diploid. You examine the resulting tetrads. For each of the two models from part **(a)**, give the types of tetrads that you would expect and their relative frequencies. The tetrad types should be described by the phenotypes (constitutive, uninducible, or regulated) of the four spores in each tetrad.

(d) Now assume that you are studying this regulatory pathway in *Drosophila*. To perform the epistasis test, you cross a $\mathbf{A^S/A^S\ B^+/B^+}$ male to a $\mathbf{A^+/A^+\ B^-/B^-}$ female. For each of the two models from part **(a)**, give the expected phenotype of the F_1 flies from this cross. Now you cross the F_1 flies among themselves to produce F_2 flies. For each of the two models, give the expected ratio of constitutive, uninducible, or regulated phenotypes among the F_2 flies.

3. You are studying the regulation of an enzyme in yeast. To begin your analysis of this regulation, you first fuse the cis regulatory region found upstream of the coding region of the gene encoding this enzyme to the LacZ coding sequence. You then place this hybrid gene on an appropriate yeast plasmid. You are relieved to find that cells carrying the hybrid gene do not express ß-galactosidase activity unless the known inducer for the enzyme synthesis is present, meaning that the regulatory region you have selected contains all of the necessary cis-acting sequences for normal regulation. You next identify two different mutants that show abnormal regulation of your reporter gene construct. Mut1$^-$ gives constitutive expression, whereas Mut2$^-$ shows uninducible expression.

(a) Are Mut1$^-$ or Mut2$^-$ cis- or trans-acting? Explain.

Next you cross a Mut1$^-$ haploid yeast strain to a Mut2$^-$ haploid yeast strain. You induce sporulation of the resulting diploid (which has the phenotype of having regulated expression of the reporter gene). Three different tetrad types are obtained. The phenotypes of the four spores found in each tetrad are listed below:

Type 1	Type 2	Type3
regulated (wt)	regulated (wt)	constitutive
regulated (wt)	constitutive	constitutive
uninducible	uninducible	uninducible
uninducible	uninducible	uninducible

There are twenty tetrads of Tetrad Type 2, five of Type 1, and two of Type 3.

(b) Next to the phenotype of each of the twelve spores listed above, write the genotype at the Mut1 locus and at the Mut2 locus of each spore whose genotype you are absolutely sure of. Use the symbols 1+ or 1$^-$ and 2+ or 2$^-$.

(c) Are the Mut1 and Mut2 loci linked? If so, what is the distance between them?

(d) What is the phenotype of a Mut1$^-$ Mut2$^-$ double mutant?

(e) Diagram a model to explain the regulation of the enzyme you are studying; your model should be consistent with all of the data you have, and should contain wild-type Mut1, Mut2, and the gene encoding your enzyme. Your model should also contain the small molecule inducer that induces expression of the enzyme you are interested in.

4. You have discovered a gene in yeast that is involved in repairing damaged DNA. Mutations in this gene make yeast more sensitive to DNA-damaging agents such as UV radiation. You designate your new gene Rad66. To study the regulation of Rad66, you fuse the cis regulatory region upstream of the Rad66 open reading frame to the LacZ coding sequence. You then place this hybrid gene (designated P_{rad66}–LacZ) on a yeast plasmid. As hoped, yeast cells carrying the P_{rad66}–LacZ plasmid do not express ß-galactosidase activity unless exposed to UV light, showing that the hybrid gene is regulated in the same way that the Rad66 gene is normally regulated.

(a) You next identify a mutant that you call Reg1⁻, which causes expression of the P_{rad66}–LacZ reporter gene construct, regardless of whether or not the cells have been exposed to UV light. By mating a Reg1⁻ haploid mutant strain to a wild-type haploid strain, you find that the resulting diploid only expresses the reporter gene in the presence of UV light. Cassify the Reg1⁻ mutation in terms of its basic genetic properties (constitutive vs uninducible, dominant vs recessive, cis- vs trans-acting). Explain the rationale behind your conclusions. Based on these properties, make a proposal for the wild-type regulatory function of the Reg1 gene.

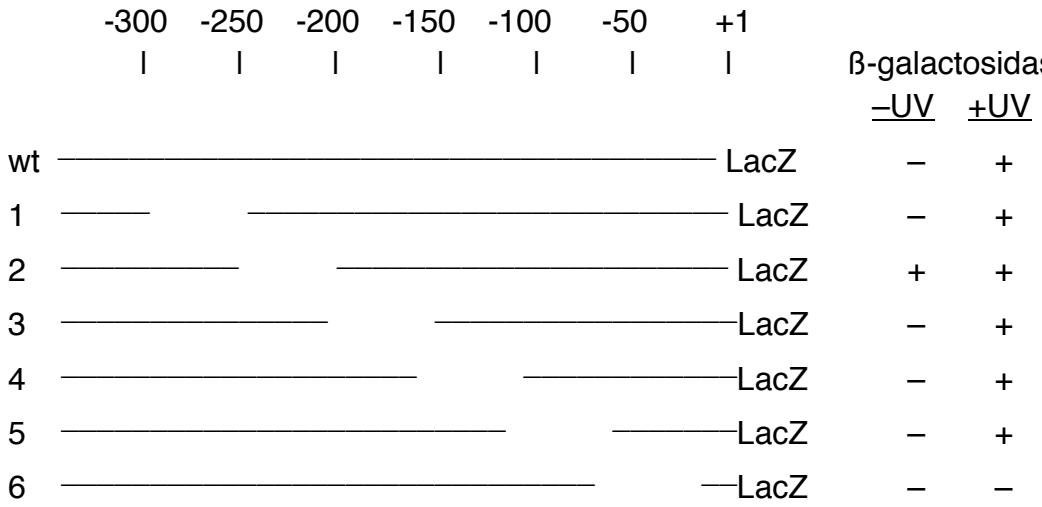
(b) Give a model for the regulatory pathway for Rad66 that is consistent with the data presented in this problem so far. In the model you diagram, include the wild-type Rad66 and Reg1 genes. Also be sure to include in your model a role for UV radiation.

(c) Next, you isolate a mutant that you call Reg2⁻, which will not express the P_{rad66}–LacZ reporter, even after cells have been exposed to UV radiation. Mating a Reg2⁻ haploid mutant strain to a wild-type haploid strain gives a diploid that can expresses the reporter only in the presence of UV radiation. Given all of the available data on the Reg1⁻ and Reg2⁻ mutants, diagram **two** different linear pathways that can explain the roles of the Reg1 and Reg2 gene products and UV radiation in the regulation of Rad66.

(d) You cross a haploid strain that carries the Reg1⁻ mutation to a haploid strain that carries the Reg2⁻ mutation. You then induce sporulation of the resulting diploid. You analyze only two tetrads, and they both show the same pattern of expression of P_{rad66}–LacZ in response to UV light: one spore shows normal induction in response to UV radiation, two spores express the reporter even in the absence of UV radiation, and one does not express the reporter. What is the phenotype of a Reg1⁻ Reg2⁻ double mutant haploid strain?

(e) Select the model from part **(c)** that is consistent with the double mutant data above.

Next you evaluate the cis regulatory sequences necessary for expression of the P_{rad66}–LacZ reporter gene construct. The figure below shows the effect of six different 50 basepair long deletions (#1 - #6) in the regulatory region on the amount of ß-galactosidase activity expressed by the reporter gene.



(f) In light of the experiments from parts **(a)** – **(e)**, propose a specific function for any cis-acting segment that has a clear regulatory role as defined by these deletion constructs.

(g) What two conclusions can you draw from the fact that deletions #3 - #5 show normal regulation of P_{rad66}–LacZ?

5. In the examples of gene expression that we have covered in class, the genes were regulated in some way or another. Some eukaryotic genes are expressed constitutively regardless of environmental conditions. Nevertheless, transcription of these constitutive genes depends on the same kind of transcriptional activator proteins that are employed in regulated transcription. The difference is that, for regulated genes, at least one activator (or repressor) must itself be regulated, whereas for constitutive genes, the activator(s) are always active. In the next parts of this problem, we will analyze the expression of a hypothetical constitutive gene in yeast.

The first step in your analysis is to fuse the regulatory region of your gene of interest to the LacZ coding sequence, and to place the hybrid gene on a yeast plasmid. Say that cells carrying the reporter gene construct express 100 units of ß-galactosidase activity under all conditions that you test. You next identify two different mutations that show decreased ß-galactosidase activity: either Con1$^-$ or Con2$^-$ single mutants express about 50 units of ß-galactosidase activity under all conditions that you test.

Next you cross a Con1$^-$ haploid yeast strain to a Con2$^-$ haploid yeast strain. You induce sporulation of the resulting diploid (which shows regulated expression of the reporter). Three different tetrad types are obtained. The amount of activity of ß-galactosidase found in each of the four spores found in each tetrad are listed below. There are three tetrads of Tetrad Type 1, forty of Type 2, and twelve of Type 3.

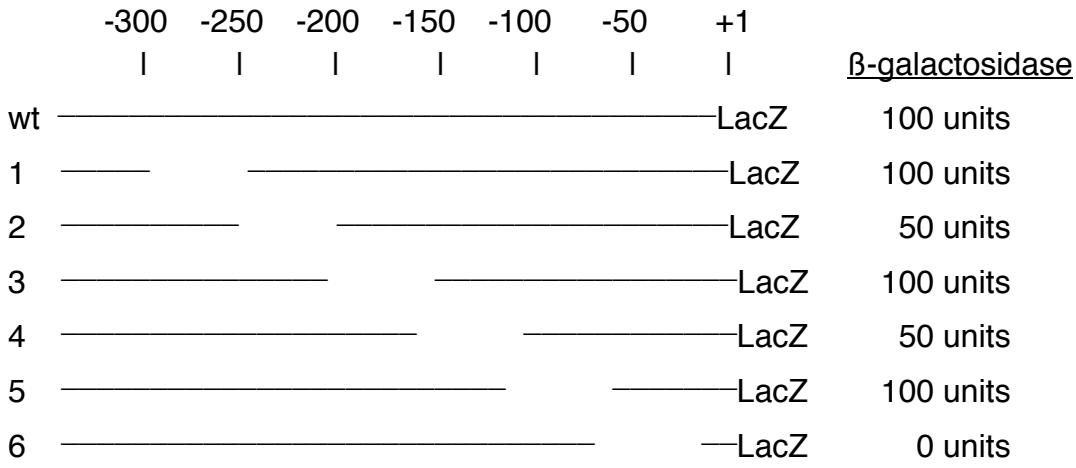
Type 1	Type 2	Type3
100 units	50 units	100 units
100 units	50 units	50 units
none	50 units	50 units
none	50 units	none

(a) Are the Con1 and Con2 loci linked? If so, what is the distance between them?

(b) How many units of ß-galactosidase activity would a Con1$^-$ Con2 $^-$ double mutant display?

(c) Diagram a model to explain the regulation of the enzyme you are studying; your model should be consistent with all of the data you have, and should contain wild-type Con1, Con2, and the gene encoding your enzyme.

Next you evaluate the cis regulatory sequences necessary for expression of the reporter gene construct you have made. The figure below shows the effect of different 50 base pair long deletions in the cis regulatory region on the amount of ß-galactosidase activity expressed by the reporter gene construct.



When deletion #2 is placed in the Con1^{-} strain, no ß-galactosidase activity is expressed, whereas deletion #2 in the Con2^{-} strain expresses 50 units of ß-galactosidase activity. Conversely, deletion #4 in the Con1^{-} strain expresses 50 units of ß-galactosidase activity, whereas deletion #4 in the Con2^{-} strain expresses no ß-galactosidase activity.

(d) Based on this new information, specify how the Con1 and Con2 gene products interact with the regulatory sequences found upstream of the open reading frame that encodes your enzyme of interest.

Altering the genomes of mice -- Transgenics and Gene targeting

1. You hypothesize that a loss of function of the *Pindrop* gene is the cause of the recessive phenotype of deafness in a strain of mice called the *Ard* strain. You decide to test this hypothesis using a variety of genetically modified mice. You have available pieces of genomic DNA that contain either the entirety of the human *Pindrop* gene or the entirety of the mouse *Pindrop* gene. When creating engineered mice, the following 8 steps need to be considered. **For each genetically engineered mouse you make**, please state:

i) whether you are using pronuclear injection or gene targeting techniques
ii) what DNA you would introduce into the mouse cells (also draw the DNA)
iii) whether you would put the DNA into a fertilized egg or ES cells
iv) what is the genotype of the fertilized egg or the ES cells you would start with
v) where in the mouse genome the DNA you introduced would integrate
vi) which additional breeding steps you would do to make the mouse you wanted
vii) two possible phenotypic results you could get from the newly made mice, and the corresponding conclusions you would make based on each result

(a) What modification to the mouse genome would allow you to test the hypothesis that *Pindrop* function is required for hearing in mice (regardless of whether *Pindrop* and *Ard* are the same gene)?

(b) What if any additional experiments would be required to test the hypothesis that *Pindrop* (a molecularly defined gene) is mutated in the *Ard* strain of mice?

(c) What additional experiments would allow you to test the hypothesis that the human and mouse *Pindrop* genes are functionally interchangeable?

2. In mammals, including humans and mice, growth hormone (a protein) is speculated to play a prominent role in determining adult size. You decide to test this hypothesis in mice using genetically modified mice. Growth hormone is encoded by a single gene (the GH gene) in humans and in mice; the DNA sequences of the human and mouse GH genes are very similar but not identical. You have available pieces of genomic DNA containing either the human or the mouse GH genes. When creating engineered mice, the following 8 steps need to be considered. **For each genetically engineered mouse you make**, please state:

i) whether you are using pronuclear injection or gene targeting techniques
ii) what DNA you would introduce into the mouse cells (also draw the DNA)
iii) whether you would put the DNA into a fertilized egg or ES cells
iv) what is the genotype of the fertilized egg or the ES cells you would start with
v) where in the mouse genome the DNA you introduced would integrate
vi) whether creating the mouse should involve the generation of a chimera or not
vii) which additional breeding steps you would do to make the mouse you wanted
viii) two possible phenotypic results you could get from the newly made mice, and the corresponding conclusions you would make based on each result

You first decide to test the specific hypothesis that additional copies of the mouse GH gene would yield mice larger than wild-type mice (which, of course, have two copies of the GH gene).

(a) What modification to the mouse genome would allow you to generate a mouse with three copies of the GH gene?

You then decide to test the specific hypothesis that mice with zero or one copy of the mouse GH gene would be smaller than wild-type mice.

(b) What modification to the mouse genome would allow you to generate a mouse with only one copy of the GH gene?

(c) How would you generate mice with zero copies of the GH gene?

(d) How would you generate mice with four copies of the GH gene?

(e) How would you generate mice with five copies of the GH gene?

(f) How would you generate mice with six copies of the GH gene?

3. As we will study later in the semester, there are genes in the human and mouse genomes that control cell proliferation. Mutations in such genes can result in uncontrolled proliferation and, as a result, cancer. For some such genes, a single mutant allele is sufficient to cause a diploid cell to proliferate wildly and cause cancer; such mutant alleles cause the **dominant** phenotype of promoting proliferation.

An example of such a gene is RAS. A mutation in RAS results in the dominant phenotype of uncontrolled proliferation and thus a predisposition to cancer (even in the presence of wild-type alleles of the RAS gene). We want to make a mouse model of cancer using this mutation in RAS.

(a) Would it be possible to generate a suitable mouse model using pronuclear injection (transgenes)? Briefly explain why or why not.

(b) Would it be possible to generate a suitable mouse model using gene-targeting techniques? Briefly explain why or why not.

(c) If a suitable mouse model could be generated using either of the two techniques, which one would you choose? Briefly explain your reasoning.

(d) What exact type of modification to the mouse genome would you make to create a mouse model for cancer? Explain your choice.

(e) Draw the DNA construct that you would use to modify the mouse genome, and explain how your construct would integrate into the mouse genome.

(f) What cell type would you put your DNA construct into? Include your choice of fertilized egg or ES cell, and what the genotype of the cells you would use should be.

(g) Explain what (if any) steps you will need to do to proceed from part **(f)** above to obtaining the final modified mouse you actually want.

For other such genes described in the introduction to this question, mutant alleles cause the **recessive** phenotype of promoting proliferation; in these cases, both copies of the gene (both alleles) must be defective to make the cell cancerous.

An example of such a gene is RB. We want to make a mouse model of cancer using this mutation in RB. This gene's normal function is to keep cells from proliferating out of control and one wild-type copy of the RB gene is sufficient to do so. However, if both RB alleles are defective, the cell will likely become cancerous as a result of uncontrolled proliferation.

(h) Would it be possible to generate a suitable mouse model using pronuclear injection (transgenes)? Briefly explain why or why not.

(i) Would it be possible to generate a suitable mouse model using gene-targeting techniques? Briefly explain why or why not.

(j) If a suitable mouse model could be generated using either of the two techniques, which one would you choose? Briefly explain your reasoning.

(k) What exact type of modification to the mouse genome would you make to create a mouse model for cancer? Explain your choice.

(l) Draw the DNA construct that you would use to modify the mouse genome, and explain how your construct would integrate into the mouse genome.

(m) What cell type would you put your DNA construct into? Include your choice of fertilized egg or ES cell, and what the genotype of the cells you would use should be.

(n) Explain what (if any) steps you will need to do to proceed from part **(m)** above to obtaining the final modified mouse you actually want.

4. Many mouse genes are expressed in a "tissue-specific" manner; that is, the genes themselves are present in all cells in the body, but the genes are expressed (transcribed and translated) in only one of the animal's many tissue types. Geneticists can study the regulation of a mouse gene by fusing the gene's cis regulatory region to the LacZ coding sequence and injecting the fusion construct into a fertilized egg to create a transgenic mouse. For example, if a geneticist wanted to know where the mouse amylase gene was expressed, she would fuse the cis regulatory region of the mouse amylase gene to LacZ to yield a $P_{amylase}$-LacZ fusion construct.

(a) Would microinjection of the $P_{amylase}$-LacZ construct into the male pronucleus of a fertilized egg likely result in integration of the construct into the copy of the amylase gene present in the male pronucleus? Briefly explain your answer.

(b) You make one strain of mice that are heterozygous for the resulting $P_{amylase}$-LacZ transgene. These mice display LacZ expression exclusively in the pancreas. Would you expect homozygotes for the transgene to also display LacZ expression in the pancreas? Elsewhere? Briefly explain your answer.

(c) You are surprised to observe that the mice that are homozygous for your transgene insertion display a serious heart defect. (Heterozygotes have normal hearts.) Suggest a possible explanation.

For the last two parts of this problem, you will be asked to design genetically engineered mice. When creating engineered mice, the following 8 steps need to be considered. **For each mouse you make**, please state:

i) whether you are using pronuclear injection or gene targeting techniques
ii) what DNA you would introduce into the mouse cells (also draw the DNA)
iii) whether you would put the DNA into a fertilized egg or ES cells
iv) what is the genotype of the fertilized egg or the ES cells you would start with
v) where in the mouse genome the DNA you introduced would integrate
vi) whether creating the mouse should involve the generation of a chimera or not
vii) which additional breeding steps you would do to make the mouse you wanted
viii) two possible phenotypic results you could get from the newly made mice, and the corresponding conclusions you would make based on each result

(d) Propose an experimental test of your hypothesis from part **(c)** using a genetically modified mouse.

(e) Propose how you might use LacZ in a gene targeting experiment in mice to test whether the amylase gene is expressed exclusively in the pancreas.

Population Genetics – populations at Hardy-Weinberg Equilibrium

1. Consider a rare autosomal recessive trait that is possessed by one in every 3,000 children in the U.S. Assume that the population is in Hardy-Weinberg equilibrium.

(a) What is the frequency (in the U.S. population) of the allele that is associated with the trait?

(b) What is the frequency of heterozygotes in the U.S. population?

What is the probability that a child born in the U.S. will have the trait if:

(c) Neither the mother nor the father has the trait, and neither have family members with the trait.

(d) The mother has the trait. The father does not have the trait and has no family members with the trait.

(e) The mother does not have the trait but has a brother with the trait. The father does not have the trait and has no family members with the trait.

(f) The mother has the trait. The father does not have the trait and has no family members with the trait, but he is an immigrant from another population (itself in Hardy-Weinberg equilibrium) where the disease affects one in every 11,000 children.

2. Consider the gene that determines which blood type a human is (A, B, AB, or O). This single gene has three alleles called I^A, I^B, and i. There are four "blood types" (phenotypic classes), as follows:

Blood type	Genotype(s)
A	$I^A I^A$ or I^A i
B	$I^B I^B$ or I^B i
AB	$I^A I^B$
O	i i

Assume that the human populations that you are working with are at Hardy-Weinberg equilibrium.

In Norway, the frequencies of the I^A and I^B alleles are 0.26 and 0.07, respectively. Based on this, estimate the following (showing your calculations):

(a) The frequencies of the six genotypes ($I^A I^A$, I^A i, $I^B I^B$, I^B i, $I^A I^B$, and i i) in Norway.

(b) The frequencies of the four blood types (A, B, AB, or O) in Norway.

You now examine the ABO blood types in a new population. You observe that the AB blood type has a frequency of 0.50 in this population.

(c) What are the frequencies of the I^A and i alleles in this population? (Show your calculation or otherwise explain your answer.)

3. In the 1950's, a screening of Europeans revealed that 30% were unable to taste the chemical compound phenylthiocarbamide (PTC). This is an autosomal recessive trait; the ability to taste PTC is dominant. Assume that, in the 1950's, the European population was in Hardy-Weinberg equilibrium with respect to the gene that is critical for PTC tasting.

(a) In Europe in the 1950's, what was the frequency of the allele (call it allele NT) associated with the inability to taste PTC?

(b) In Europe in the 1950's, what fraction of NT alleles were found in individuals who could not taste PTC?

What was the probability that a child born in Europe in the 1950's would be a PTC taster if:

(c) Both parents were PTC tasters.

(d) One parent was a PTC taster, but the other parent was not.

Also in the 1950's, a screening of West Africans revealed that 1% were unable to taste PTC. Assume that, in the 1950's, the West African population was in Hardy-Weinberg equilibrium with respect to the gene that is critical to PTC tasting

(e) In West Africa in the 1950's, what was the frequency of the NT allele?

(f) In West Africa in the 1950's, what fraction of NT alleles were found in individuals who could not taste PTC?

4. Consider two large but completely isolated populations of rabbits: population X (consisting of 100,000 randomly mating rabbits) and population Y (consisting of 50,000 randomly mating rabbits). In both populations there are two alleles for tail color: the B allele (associated with brown tails), and the b allele (associated with white tails). Brown tail color is dominant to white tail color. In population X, 1000 rabbits have white tails. In population Y, 5000 rabbits have white tails. Assume Hardy-Weinberg equilibrium. Estimate the following (showing your calculations):

(a) The frequency of the b allele in population X.

(b) The number of heterozygous rabbits in population X.

(c) The probability that a rabbit will have a white tail if its mother is a randomly selected white-tail rabbit from population X and its father is a randomly selected brown-tail rabbit from population Y.

(d) The probability that a rabbit will have a brown tail if its mother is a randomly selected brown-tail rabbit from population X and its father is a randomly selected brown-tail rabbit from population Y.

Populations 1 and 2 have been separated for many years by a deep, raging river. Suppose that the river is now dammed upstream of the two populations, allowing for easy crossing of the riverbed and random mating between the two formerly isolated populations.

(e) Once Hardy-Weinberg equilibrium is established in the new, joint population (call it population XY), what will be the frequency of the b allele in population XY? Show your calculations.

(f) Once Hardy-Weinberg equilibrium is established in population XY, what fraction of the rabbits will have brown tails? Show your calculations.

(g) After the river is dammed, how many generations of random mating will be required for genotype frequencies to match Hardy-Weinberg expectations?

5. In humans, albinism (unpigmented skin, hair, and eyes) is due to an enzymatic deficiency, and it is an autosomal recessive trait. Suppose that in a small country of one million people ("Generation 1"), there are 500 *aa* albinos and 9000 *Aa* heterozygous carriers.

(a) Estimate q, the frequency of allele *a*, in Generation 1. Show your calculations.

(b) Estimate p, the frequency of allele *A*, in Generation 1. Show your calculations.

(c) In the next generation of 1 million individuals (Generation 2), what are the expected numbers of *aa* albinos and *Aa* carriers? (Assume random mating and all other Hardy-Weinberg conditions.) Show your calculations.

(d) Has the frequency of allele *a* changed between Generations 1 and 2? Briefly justify your answer.

(e) Was Generation 1 in Hardy-Weinberg equilibrium? Briefly justify your answer.

What is the probability that a child will be albino if:

(f) Both parents are non-albino members of Generation 2?

(g) One parent is a non-albino member of Generation 2, and the other parent is a non-albino member of Generation 1.

(h) Both parents are members of Generation 2, and one parent is albino and the other is non-albino.

Population Genetics – populations not at equilibrium (because of selection and/or mutation)

1. In answering the various parts of this question, assume that mating is random. State any additional simplifying assumptions that you employ, and show your calculations.

First, consider an autosomal recessive disease that is usually lethal in childhood, and that has an incidence among newborns of 1/3000.

(a) What mutation rate would be required to maintain this steady state frequency in the population?

(b) If your answer to part **(a)** seems very high, perhaps the explanation for the high incidence of the disease is heterozygote advantage. How large would this heterozygote advantage have to be (assuming the mutation rate is negligible) to maintain the same steady state frequency in the population?

Colorblindness is an X-linked recessive trait that is found in about 5 percent of males. Assume that colorblindness does not affect one's fitness at all and assume no new mutations.

(c) What fraction of females are heterozygous carriers of the allele that leads to colorblindness?

(d) What fraction of females are homozygous for the allele that leads to colorblindness?

2. Consider a heritable autosomal disease with an incidence in the population of 1 per thousand. On average, individuals with the disease have 80% as many children as the population average.

In answering the various parts of this question, assume that mating is random. State any additional simplifying assumptions that you employ, and show your calculations.

(a) What mutation rate would be required to maintain the observed incidence of the disease in the population if the disease is dominant?

(b) What mutation rate would be required to maintain the observed incidence of the disease in the population if the disease is recessive?

(c) Now assume that the mutation rate is zero, that the disease is recessive, and that the disease allele is maintained in the population by heterozygote advantage. Calculate the heterozygote advantage necessary to keep this population at steady state.

3. Cystic fibrosis is an autosomal recessive disease that currently affects about 1 in 1600 children in Europe. It has been hypothesized that heterozygotes for the cystic fibrosis allele may be less susceptible to some infectious disease (possibly plague) that was common in Europe hundreds of years ago, but that is seen rarely if ever today. Assume that today, and in the past, individuals with cystic fibrosis disease have about 5% as many offspring as average individuals in European populations. Assume that mating is random, and that the mutation rate is negligible.

(a) In Europe today, what is the frequency of the allele (call it CF) associated with cystic fibrosis?

(b) If the hypothesis stated above is correct, is the heterozygote advantage today higher, lower, or the same as it was hundreds of years ago? Explain your answer.

(c) If the hypothesis stated above is correct, is the frequency of CF heterozygotes today likely to be higher, lower, or the same as it was hundreds of years ago? Explain your answer.

4. In practice, it can be very difficult to detect subtle selection for or against the heterozygote for an allele that causes an obvious recessive phenotype. Consider a homozygous-lethal allele that has a steady-state frequency of 0.0004 in a case in which there would be no selection for or against the heterozygote. Assume mating is random.

(a) Calculate the mutation rate for this gene in this population.

(b) Now change one assumption: Assume that heterozygous carriers have a fitness of 0.99. (Assume no change in the mutation rate.) What would be the steady-state frequency of the homozygous-lethal allele under these conditions?

(c) Now reverse the assumption: Assume that heterozygotes experience an advantage of $h = 0.01$. (Again, assume no change in the mutation rate.) What would be the steady-state frequency of the homozygous-lethal allele under these conditions?

5. Suppose that body color in cockroaches is controlled by an autosomal gene "gene G." GG and Gg cockroaches are black, and gg cockroaches are white. Let us consider a population of cockroaches that lives in your apartment and that mates at random.

(a) You count a week's worth of newborn cockroaches in your apartment and find that they include 99,990 black cockroaches and 10 white cockroaches. Estimate the frequency of the **g** allele in this population.

(b) Assume that the cockroach population in your apartment has held steady for more than a year. Throughout this period, you have disliked cockroaches and have smashed them whenever you spotted them. White cockroaches are easy to spot on the black floor of your apartment, and thus white cockroaches (gg) have suffered a selective disadvantage. White cockroaches are only 20% as likely as black cockroaches to survive to reproductive age. What is the mutation rate of (G → g) per generation in this population?

(c) You now have less time to rid your apartment of these pests. This environmental change results in white cockroaches being 60% as likely as black cockroaches to survive to reproductive age. What would be the new frequency of the **g** allele at steady state?

(d) Now assume that the G → g mutation rate falls to zero. Simultaneously you apply a pesticide that kills many of the cockroaches. Unfortunately, the **g** allele confers partial resistance to this pesticide so that, in the presence of the pesticide, Gg heterozygotes have 20% more offspring than do GG cockroaches. White cockroaches continue to be 60% as likely as black cockroaches to avoid smashing prior to reproductive age. What would be the new frequency of the **g** allele at steady state?

Population Genetics – populations not at equilibrium (because of non-random mating)

1. In answering the various parts of this question, show your calculations, and state any additional simplifying assumptions that you employ.

The frequency of the autosomal recessive disorder PKU (phenylketonuria) among newborns is approximately 1/10,000 (when the parents are unrelated). Assume mutation rates and selection are negligible.

(a) What is the risk of having a child with PKU if the parents are siblings?

(b) What is the risk of having a child with PKU if the parents are uncle and niece?

(c) What is the risk of having a child with PKU if the parents are second cousins? (Note: Second cousins share a set of common great grandparents. First cousins share a set of common grandparents.)

PP, QQ, and RR are three different inbred, true-breeding strains of mice. You can mate two of these strains to each other to get a "hybrid" of the two strains.

A hybrid between PP and QQ is mated to an RR mouse, and then a hybrid between QQ and RR is mated to a PP mouse. A male resulting from the first mating is mated to a female resulting from the second mating.

(d) What is the inbreeding coefficient of their progeny?

2. "Double first cousins" are the result of either two brothers marrying two sisters, or of a brother/sister pair marrying another brother/sister pair. In some human populations where first-cousin marriages are common, double-first-cousin marriages also occur at a significant rate. Consider two large but completely isolated human populations (populations 1 and 2). In population 1, mating is random. In population 2, 15% of matings are between first cousins and 5% of matings are between double-first cousins. Assume that all other matings in population 2 are random. Consider two autosomal recessive diseases (diseases S and T). The incidence of disease S is the same (1 in 10,000 people) in both populations. In the case of disease T, the disease allele frequency is the same (0.002) in both populations. Assume that mutation rates and selection are negligible.

(a) Draw a pedigree reflecting the mating of double first cousins, and calculate the inbreeding coefficient for matings between double-first cousins.

(b) What is the incidence of disease T in population 1?

(c) What is the incidence of disease T in population 2?

(d) What is the frequency of the allele associated with disease S in population 1?

(e) What is the frequency of the allele associated with disease S in population 2?

(f) In population 2, what fraction of individuals with disease S are products of either first cousin or double-first-cousin matings?

(g) In population 2, what fraction of individuals with disease T are products of either first-cousin or double-first-cousin matings?

3. Consider two large but completely isolated human populations (populations M and N). A particular autosomal recessive disease affects 1 in 2500 people in both populations. Population M is characterized by random marriage. However, 10% of marriages in population N are between first cousins. (Assume that all other marriages in population N are random.)

(a) What is the frequency of the disease-associated allele in population M?

(b) What is the frequency of the disease-associated allele in population N?

Now assume that, instead, a particular autosomal dominant disease affects 1 in 2500 people in both populations. Population M is characterized by random marriage. However, 10% of marriages in population N are between first cousins. (Assume that all other marriages in population N are random.)

(c) What is the frequency of the disease-associated allele in population M?

(d) What is the frequency of the disease-associated allele in population N?

Now assume that, instead, a particular X-linked recessive disease affects 1 in 2500 males in both populations. Population M is characterized by random marriage. However, 10% of marriages in population N are between first cousins. (Assume that all other marriages in population N are random.)

(e) What is the frequency of the disease-associated allele in population M?

(f) What is the frequency of the disease-associated allele in population N?

4. Johnny Lunchbucket and Betty Juicebox were both raised by single mothers. They met in 7.03 and fell in love calculating LOD scores together. They have been married for several years and want to start a family. Recently, however, they discovered that they share the same father. Thus, Johnny and Betty are half-brother and half-sister.

(a) If Johnny and Betty decide to have a child, what would the inbreeding coefficient be?

(b) They seek the advice of a genetic counselor, who describes the increased risk of autosomal recessive traits among inbred children. To illustrate the risks, the counselor tells them about a rare autosomal recessive disease known as *dormus in lecturitis*. The counselor tells Johnny and Betty that the probability of their child being afflicted with *dormus in lecturitis* is 2×10^{-4}. What is the frequency of the mutant allele in the population? (Assume no selection and no new mutations.)

(c) The genetic counselor also tells them that 200 times as many children with *dormus in lecturitis* come from random matings as from half-sibling marriages. What is the frequency of half-sibling marriages in this population?

(d) Betty and Johnny are particularly concerned about genetic load and the risk of a stillbirth or neonatal death. The genetic counselor informs them that their father was likely to carry 2 alleles that cause recessive lethal disorders, and that the frequency of stillbirth or neonatal death in children of unrelated parents is 0.04. What is the likelihood that their child would be stillborn or die neonatally?

5. In this question we will consider the interaction of selection and inbreeding in determining the incidence of autosomal recessive diseases. Consider a gene in which mutations occur at a rate of 10^{-5}, and that these mutations all create alleles that cause the recessive disease. Assume a selective disadvantage S of 0.4 in homozygotes for the allele associated with the disease.

In answering the various parts of this question, show your calculations (unless none are required), and state any additional simplifying assumptions that you employ.

(a) Calculate q, the frequency of the allele associated with the disease. Also calculate the incidence of the disease. Assume random mating.

(b) Now assume that, for thousands of generations, 10% of all children have been products of first-cousin matings (the remaining 90% being products of random matings). Calculate the steady-state value of q. Also calculate the incidence of the disease at steady state. (Hint: first modify the equation $\Delta q_{sel} = - Sq^2$ to reflect inbreeding's effects on the incidence of homozygotes for the recessive allele.)

(c) Now assume that the population described in part **(b)** suddenly and completely ceases all inbreeding. Calculate the incidence of disease in the first generation conceived with no inbreeding.

(d) Would q be expected to rise, fall, or remain unchanged during the first 10 generations after the cessation of inbreeding described in part **(c)**? Briefly justify your answer. (No calculations needed.)

(e) What numerical value would q approach after thousands of generations with no inbreeding?