7.03 Problem Set 1 Answer Key

1. Imagine a natural compound called Spindlestop, which shows promising anti-tumor activity. The primary source of Spindlestop is the bark of an endangered species of tree, but suppose that the yeast *S. cerevisiae* is found to produce Spindlestop in minute, yet detectable, quantities. To increase production, you isolate 30 yeast mutants with increased levels of Spindlestop. Mutants 1-15 are mating type a (MAT a) and mutants 16-30 are mating type α (MAT α).

The analysis begins by pairwise mating of each mutant to a wild-type strain and to the mutants of the opposite mating type. The amounts of Spindlestop produced by the resulting diploids are shown in the table below ("wt" indicates wild-type quantities, "+" indicates about 10X wild-type levels).

MAT α Strains

	Wild-type	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Wild-type	wt	+	wt	wt	wt	wt	+	wt	+	+	wt	wt	wt	wt	wt	+
1	wt	+	wt	wt	wt	wt	+	wt	+	+	wt	wt	wt	+	wt	+
2	wt	+	wt	+	wt	wt	+	+	+	+	+	wt	wt	wt	+	+
3	wt	+	wt	+	wt	wt	+	+	+	+	+	wt	wt	+	+	+
4	wt	+	+	wt	+	+	+	wt	+	+	wt	+	+	wt	wt	+
5	wt	+	wt	+	wt	wt	+	+	+	+	+	wt	wt	wt	+	+
6	wt	+	wt	+	wt	wt	+	+	+	+	+	wt	wt	wt	+	+
7	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
9	wt	+	+	wt	+	+	+	wt	+	+	wt	+	+	wt	wt	+
10	wt	+	+	wt	+	+	+	wt	+	+	wt	+	+	wt	wt	+
11	wt	+	+	wt	+	+	+	wt	+	+	wt	+	+	wt	wt	+
12	wt	+	+	wt	+	+	+	wt	+	+	wt	+	+	wt	wt	+
13	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
14	wt	+	+	wt	+	+	+	wt	+	+	wt	+	+	wt	wt	+
15	wt	+	+	wt	+	+	+	wt	+	+	wt	+	+	wt	wt	+

⇧

MAT a Strains

a) Which of the mutants are dominant and which are recessive?

The phenotype of the diploid resulting from a mutant to wild type test cross indicates whether a particular mutation is recessive or dominant. A wild type diploid phenotype indicates that the mutation is recessive, while a mutant diploid phenotype suggests that the mutation is dominant.

Recessive: 1-6, 9-12, 14-15, 17-20, 22, 25-29
Dominant: 7-8, 13, 16, 21, 23-24, 30

b) Based on the observed properties what can you conclude about the function of the gene affected by Mutant 1 with respect to Spindlestop production? What can you conclude about the gene altered by Mutant 7?

Mutant 1 possesses a recessive mutation. Because recessive mutations tend to be loss of function mutations, we suspect that the gene affected by mutant 1 serves to suppress spindlestop biosynthesis. A recessive, loss-of-function mutation in this gene would increase spindlestop biosynthesis.

Conversely, mutant 7 possesses a dominant mutation. Because dominant mutations tend to be gain-of-function mutations, we suspect that the gene affected by mutant 7 serves to promote spindlestop biosynthesis.

c) What is anomalous about the behavior of Mutant 3? Provide a simple genetic explanation.

Mutant 3 resides in two different complementation groups. Given that a complementation group comprises a set of mutations in the same gene, mutant 3 must be a double mutant; it possesses mutations in two different genes in the spindlestop biosynthetic pathway.

d) Organize the 30 mutations into complementation groups (genes). Please indicate any remaining ambiguities.

A complementation group is a set of mutations in the same gene. Recessive mutants that fail to complement possess mutations in the same gene and thus belong to the same complementation group.

I. 1, 28, (3)
II. 2, 5, 6, 18, 22, 25, 29, (3)
III. 4, 9-12, 14-15, 17, 19-20, 26, 27
All of the dominant mutations (7-8, 13, 16, 21, 23-24, 30) are ambiguous. The complementation test can only be used to study recessive mutations.

e) Based on these limited complementation data, what is the absolute minimum number of genes that must comprise the Spindlestop biosynthetic pathway? What is the maximum number of genes?

Minimum: 3 Maximum: 11

The absolute minimum number of genes in the spindlestop biosynthetic pathway is equal to the number of identified complementation groups, each of which represents a particular gene in the pathway. The maximum number of genes is equal to the number of complementation groups plus the number of ambiguous mutations. In this example, there are 8 ambiguous mutations corresponding to the number of dominant mutations that we cannot identify with complementation analysis.

f) In an attempt to construct a yeast strain that produces even more Spindlestop than any of the existing mutants you decide to combine two of the <u>recessive</u> mutants to make a double mutant strain. Pick two of the mutants to combine and explain the reasoning for your choice.

Select any two mutants in different complementation groups; this will enable you to hit two different genes in the spindlestop pathway. Because the two mutations in mutant 3 (complementation groups I and II) show no additive effect on spindlestop biosynthesis, combining recessive mutants from complementation groups I and III or II and III is preferable.

2. Wild type mice are gray and in a large-scale breeding colony two white female mice arise from different parents. You would like to know whether the white phenotype is caused by two mutations in the same gene or in different genes. Using the concepts of dominance, recessivity, and complementation, describe a set of crosses and the interpretation of their outcomes that you would use to make this determination. Please be sure to indicate the circumstances that would prevent you from easily making this determination. (Assume you have available an unlimited number of true breeding wild-type gray mice.)

To determine whether two mutations occur in the same gene or in different genes, we must perform a complementation test. Before we can do this, however, we must first establish that the two mutations we're dealing with are recessive.

CROSS 1: Dominance/Recessive Test

Cross each white female with a true-breeding, wild-type gray male mouse. Denoting the wild-type allele as 'G', the white allele in Mutant 1 as 'g-1', and white allele in Mutant 2 as 'g-2', we have:

g-1/g-1 x G/G → G/g-1
g-2/g-2 x G/G → G/g-2

If 100% of the G/g-1 and G/g-2 heterozygotes from the F1 generation are phenotypically wild-type (i.e. gray), then we may conclude that the g-1 and g-2 alleles are both recessive. (Note: If we obtain additional white mice in the F1 generation from either cross, then the mutant allele from that particular cross is dominant, and we therefore cannot perform a complementation test with this mutant.)

CROSS 2: Generation of a g-1/g-1 OR g-2/g-2 white MALE mouse

Assuming that cross 1 showed the g-1 and g-2 alleles to be recessive, we would next perform a complementation test. However, we are told in the question that our two mutants are both female, and crossing these mice to one another is not possible. Therefore, we must generate a homozygous g-1/g-1 or g-2/g-2 mutant male mouse to perform a complementation test.

One could do this by crossing a G/g-1 heterozygote from the F1 generation of Cross 1 to the original g-1/g-1 mutant female. Half of the offspring from this cross will be g-1/g-1, some of which will be male. The same could be done with a G/g-2 F1 heterozygote and the original g-2/g-2 mutant female.

G/g-1 (F1 male from cross 1) x g-1/g-1 (orig. female mutant) → 50% G/g-1, 50% g-1/g-1

Also, male and female mice from the F1 generation of the g1 strain (or the g2 strain) could be mated to generate white male mice.

CROSS 3: Complementation Test

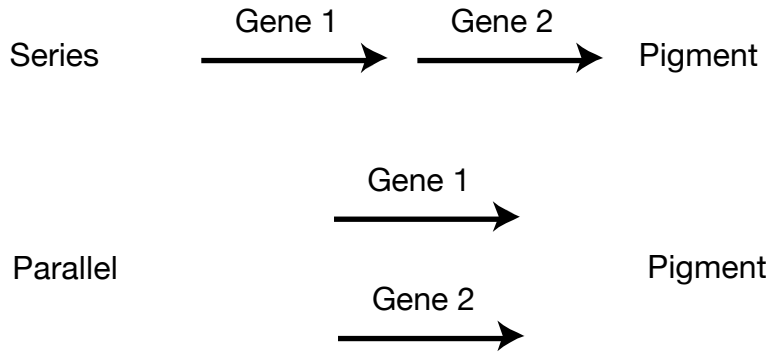
Now that we have a homozygous mutant male (g-1/g-1), we can cross this mouse to the original g-2/g-2 female mutant mouse. This is the complementation test. If the progeny from this cross are phenotypically wild-type (i.e. gray), then the mutations in g-1 and g-2 occur in different genes (complementation). If the progeny are phenotypically mutant (i.e. white), then mutations in g-1 and g-2 occur in the same gene (fail to complement).

g-1/g-1 (white male from Cross 2) x g-2/g-2 (original white female mutant)

or

g-2/g-2(white male from Cross 2) x g-1/g-1 (original white female mutant)

3. Genes that control coat color in mice can be thought of as steps in biochemical pathways whose products are pigmented compounds that give the fur its color. Albino mice have white fur because they lack the ability to make any pigment. Imagine mutations in two different genes that can, in certain combinations, block the production of pigment, yielding mutants with white fur. There are two different possible arrangements for two biochemical steps responsible for the formation of pigment: the two genes might act in *series* such that a loss of function of *either* gene would block the formation of pigment, or the two genes could act in *parallel* such that loss of function of *both* genes would be required to block the formation of pigment.



a) Say that you are given an albino mouse. When you cross this mouse to wild type all of the F1 progeny appear normal (i.e. like wild type). Consider the following three possibilities for the genetic basis of the albino trait: 1) recessive allele of a single gene, 2) recessive alleles of two genes acting in series 3) recessive alleles of two genes acting in parallel. For each of the three possibilities give the proportion of albino and normal looking mice among the F2 generation.

Hypothesis 1) Recessive allele of a single gene

F1 A/a X A/a

F2 1 A/A : 2 A/a: 1 a/a

Because the allele is recessive, only a/a mice will show the albino phenotype. We would expect a **3 normal : 1 albino** phenotypic ratio among the F2 progeny.

Hypothesis 2) Recessive alleles of two genes acting in series

F1 A/a, B/b X A/a, B/b

F2 9 A/-, B/- : 3 A/-, b/b : 3 a/a, B/- : 1 a/a, b/b

Because the genes are acting in series, possessing recessive albino alleles for either of the genes will produce the albino phenotype. Mice with an a/a OR b/b genotype will be albino. As a result, we would expect a phenotypic ratio of **9 normal : 7 albino** among the F2 progeny.

Hypothesis 3) Recessive alleles of two genes acting in parallel

F1 A/a, B/b X A/a, B/b

F2 9 A/-, B/- : 3 A/-, b/b : 3 a/a, B/- : 1 a/a, b/b

Because the genes are acting in parallel, possessing recessive albino alleles for both of the genes is required to produce the albino phenotype. Only mice with an a/a, b/b genotype will be albino. As a result, we would expect a phenotypic ratio of **15 normal : 1 albino** among the F2 progeny.

b) You cross the normal looking F1 mice among themselves producing 40 F2 mice; 15 are albino and 25 appear normal. Determine whether these data are consistent with each of the three possibilities outlined in part (a) and draw whatever conclusions you can about the inheritance of albinism. The table below gives chi square values for 1, 2 and 3 degrees of freedom. Use the convention that for $p < 0.05$ there is a statistically significant difference between the observed results and the results expected for a given model and therefore we can reject the model on the basis of the experimental data.

p value:	.995	.975	0.9	0.5	0.1	0.05	0.025	0.01	0.005
do = 1	.000	.000	.016	.46	2.7	3.8	5.0	6.6	7.9
do = 2	.01	.05	.21	1.4	4.6	6.0	7.4	9.2	10.6
do = 3	.07	.22	.58	2.4	6.3	7.8	9.3	11.3	12.8

To set up the Chi Squared Test, we must first determine the expected number of normal and albino mice under each hypothetical model. Based on the expected phenotypic ratios for each model calculated in part a, we can easily determine the expected number of normal and albino mice.

	Gray Mice (wt)	Albino Mice	$\Delta = \lvert O\text{-}E \rvert$
Observed	25	15	
Exp(Hypothesis 1)	30	10	5
Exp(Hypothesis 2)	22.5	17.5	2.5
Exp(Hypothesis 3)	37.5	2.5	12.5

The degrees of freedom equals the number of phenotypic classes minus 1. In this example, there are 2 phenotypic classes, normal and albino.

Degrees of freedom = # of classes – 1
Degrees of freedom = 2 – 1 = 1

We can now calculate the χ^2 value for each hypothesis with $\chi^2 = \Sigma\ (O\text{-}E)^2/E$ by plugging in the values we obtained above for each hypothesis.

Hypothesis 1:
$\chi^2 = \Sigma\ (O\text{-}E)^2/E = 5^2/10 + 5^2/30 = 3.33$

For 1 degree of freedom and a χ^2 value of 3.33, $0.05 < p < 0.1$. Given the convention that we only reject hypotheses for $p < 0.05$, we cannot reject the hypothesis.

Hypothesis 2:
$\chi^2 = \Sigma\ (O-E)^2/E = 2.5^2/22.5 + 2.5^2/17.5 = 0.63$

For 1 degree of freedom and a χ^2 value of 0.63, $0.1 < p < 0.5$. Given the convention that we only reject hypotheses for $p < 0.05$, we cannot reject the hypothesis.

Hypothesis 3:
$\chi^2 = \Sigma\ (O-E)^2/E = 12.5^2/2.5 + 12.5^2/37.5 = 66.6$

For 1 degree of freedom and a χ^2 value of 66.6, $p < 0.05$. Given the convention that we reject hypotheses for $p < 0.05$, we can reject hypothesis 3. The statistically significant difference between the observed data and the expected results of the model warrants rejection of the hypothesis.

Given that the p value for hypothesis 2 is larger than that for hypothesis 1, the model of two genes acting in series is more statistically consistent with the observed results. While this doesn't provide complete certainty that hypothesis 2 is accurate, we can reasonably say that this model represents the most likely mode of inheritance of the three models tested.

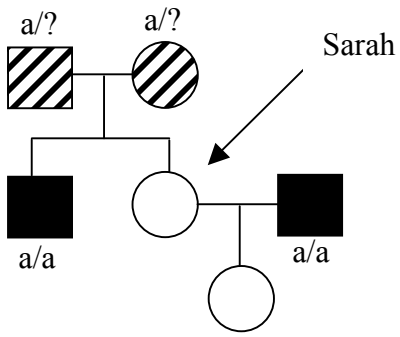
4. PKU is an autosomal recessive genetic disorder resulting from a loss of the enzyme phenylalanine hydroxylase, which converts phenylalanine to tyrosine. Without this enzyme, phenylalanine and its breakdown products accumulate to toxic levels resulting in mental retardation. Fortunately, individuals homozygous for the disease allele can be spared by a phenylalanine-free diet.

Say that Sarah has a brother with phenylketonuria (PKU), but she doesn't have the disease herself.

a) What is the probability that Sarah is a carrier for PKU?

Sarah's husband was detected to have PKU at birth but he has been treated with a phenylalanine-free diet. Sarah and her husband recently had their first child and are relieved that their child shows normal phenylalanine levels and thus is not homozygous for the disease allele.

b) Based on this new information use Bayes Theorem to compute an updated probability that Sarah is a carrier for PKU.



a) Despite our incomplete knowledge of the genotypes of Sarah's parents, we may calculate the probability that Sarah is a carrier for PKU as follows:

p(Sarah is carrier) = p(Parents are both Aa)*p(Sarah gets 1 A and 1 a) +
2*p(1 parent is aa and one parent is Aa)*p(Aa parent gives A to Sarah)

Notes:
(i) The combination of parental genotypes must allow Sarah to be 'Aa'.
(ii) Because Sarah's brother is affected, we know that each parent must have at least one 'a' allele.
(iii) The remaining possible combinations for parental genotypes are (1) Aa and Aa, (2) Aa and aa, and (3) aa and Aa. Combinations (2) and (3) are equally likely, and we may therefore condense the algebraic expression above by simply doubling the second term.
(iv) With three equally likely combinations of parental genotypes, the probability of any one combination is 1/3.
(v) Because Sarah is unaffected, her genotype cannot be 'aa'. If her parents are both Aa, this leaves 3 possibilities – AA, Aa, and aA. Her probability of being a carrier is then 2/3.
Taking these notes into account, we can insert values into the above equation to calculate the probability that Sarah is a carrier:

Taking these notes into account, we can insert values into the above equation to calculate the probability that Sarah is a carrier:

p(Sarah is carrier) = (1/3)*(2/3) + 2*(1/3)*(1) = 8/9

From the information given, there is an 88.9% chance that Sarah is a carrier for PKU.

****ASSUMING THAT BOTH OF SARAH'S PARENTS ARE CARRIERS****

a) p(Sarah is carrier) = 2/3 (See note (v)

****ASSUMING UNKNOWN GENOTYPES FOR SARAH'S PARENTS****

b) Because Sarah's child is unaffected, this new information should lead us to predict that the new probability that Sarah is a carrier (as opposed to AA) is less than 0.89. We will apply Bayes' Theorem to confirm:

X = Sarah is a carrier for PKU
Y = Sarah's child is unaffected by PKU

p(X) = 8/9
p(not X) = 1 – 8/9 = 1/9
p(Y | X) = ½
p(Y | not X) = 1

Bayes' Theorem:
p(X | Y) = p(Y | X)*p(X) / [p(Y | X)*p(X) + p(Y | not X)*p(not X)]
p(X | Y) = (1/2)*(8/9) / [(1/2)*(8/9) + (1)*(1/9)] = 0.80

Accounting for the new information that Sarah's child is unaffected by PKU, we find that, as expected, the probability that Sarah is a carrier decreases from 0.89 to 0.80.

****ASSUMING THAT BOTH OF SARAH'S PARENTS ARE CARRIERS****

X = Sarah is a carrier for PKU
Y = Sarah's child is unaffected by PKU

p(X) = 2/3
p(not X) = 1 – 2/3 = 1/3
p(Y | X) = ½
p(Y | not X) = 1

Bayes' Theorem:
p(X | Y) = p(Y | X)*p(X) / [p(Y | X)*p(X) + p(Y | not X)*p(not X)]
p(X | Y) = (1/2)*(2/3) / [(1/2)*(2/3) + (1)*(1/3)] = 0.50