

7.03 Exam 3

KEY

Name: _____

TA: Kunle Demuren

Marissa Fletcher

Levi Shmuel Teitz,

Aneesha Tewari

Exam starts at 11:05 and ends at 11:55

There are eleven (13) pages including this cover page

Please write your name on each page.

Please...

- Look over the entire exam so you don't spend too much time on hard questions leaving easy questions unanswered.
- Check your answers to make sure that they make sense.
- To help us give partial credit, show your work and state any assumptions that you make.

Question 1	36 points
Question 2	34 points
Question 3	30 points

Name: _____

1. An allele m that causes a recessive phenotype of microcephaly in wild hamsters has a frequency of $q=10^{-2}$ in this wild population. Microcephalic hamsters die at birth.

(a 4 points) What is the selective coefficient (S) for the m allele?

$$S = 1$$

+4 points

(b 8 points) While monitoring the number of live hamsters per birth in the wild, you notice that heterozygous (M/m) hamsters survive birth at higher rates than hamsters homozygous for the non-microcephaly allele (M/M). What is the heterozygous advantage of M/m hamsters in the wild?

At steady state, therefore $Sq^2=hq$

$$h=Sq$$

$$h= 1 \cdot 10^{-2}$$

+4 correct formula

+4 correct answer

-6 for correct formula without equating it to zero for steady state and no calculation

-2 minor formula error

Name: _____

(c 8 points) Excited by this finding, you decide to establish a very large colony of these hamsters in your animal facility. Your helpful and capable lab techs quickly realized that with specific care, they can deliver M/M hamsters at the same survival rate as M/m ones. By how much would the frequency of the **q** allele change in the hamster population in the animal facility after one generation? (assume for now that the mutation rate is negligible)

$h = 0$ since homozygotes and heterozygotes now have the same survival rate.

$$\Delta q_{\text{sel}} = -Sq^2$$

$$\begin{aligned}\Delta q_{\text{sel}} &= -(10^{-2})^2 \\ &= -10^{-4}\end{aligned}$$

+4 correct formula

+4 correct answer

-8 for saying q does not change

-8 for using $h = 10^{-2}$ with incorrect formula

(d 8 points) After many generations, the lab population of hamsters reached an allele frequency of $q=10^{-5}$ and did not change further. What is the mutation rate?

$$\mu = Sq^2$$

$$\mu = 1 \cdot (10^{-5})^2$$

$$\mu = 10^{-10}$$

+4 formula

+4 answer

-8 for including $h = 10^{-2}$ and using the wrong formula ($\mu = -Sq^2 + hq$)

-4 for including $h = 10^{-2}$ and using the right formula ($\mu = Sq^2 - hq$)

Name: _____

(e 4 points) You left on a trip, leaving a new tech in charge. When you returned, you realized that your tech **only** allowed matings between related hamsters. However, they were so forgetful, they could not remember whether those were uncle/niece (aunt/nephew) matings, or half first cousins. Luckily, your textbook tells you that the inbreeding coefficient for uncle/niece matings is $1/8$, but it does not provide information on half first cousins. What is the inbreeding coefficient (F) for half first cousins? For your convenience, pedigrees for the two relations are shown in the next page.

$$(1/2)^2(1/2)^2(1/4)*2 = (1/16)*(1/2) = 1/32 \text{ or } 0.03125$$

+2 for correct formula/fractions

+2 for correct F

(-2 for multiplying by 4 or 6 instead of 2)

(f 4 points) You find that the frequency of microcephalic hamsters in the litter is now 1.25×10^{-6} . Which matings did the technician do (assume that overall, **q** has not changed)? Show the calculations underlying your answer.

$$1.25 \times 10^{-6} = F * 1 \times 10^{-5}$$

$$F = 1.25 \times 10^{-6} / 1 \times 10^{-5} = 0.125$$

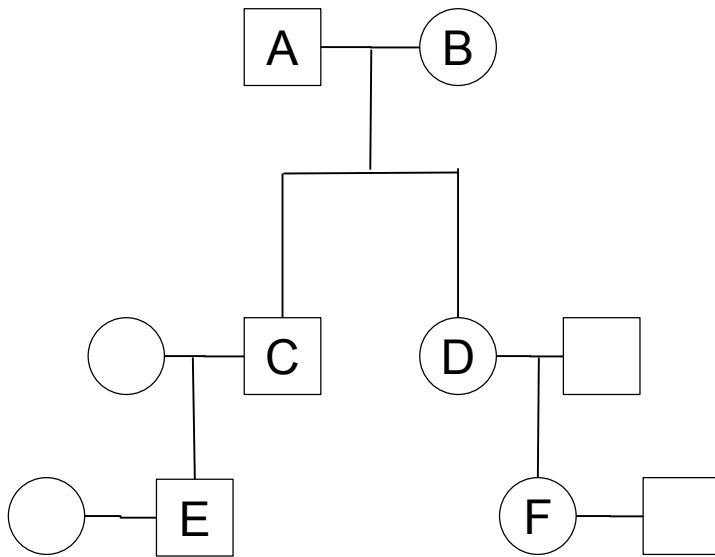
uncle/niece or aunt/nephew

+2 for solving correctly

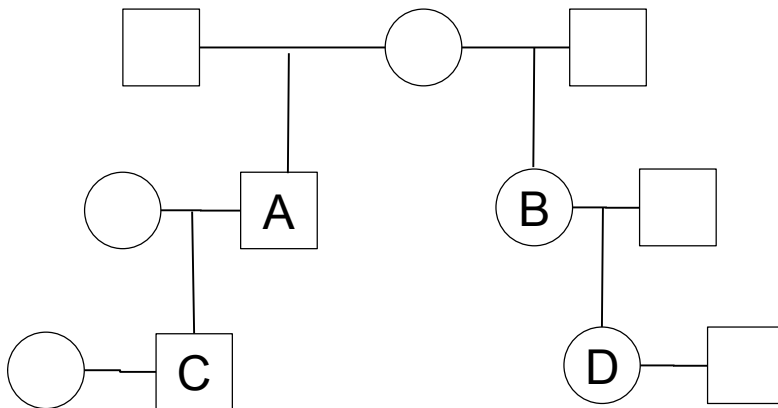
+2 for correct mating identified

(no penalty for replicating aunt/uncle typo)

Name: _____



C and F are uncle/niece



C and D are half first cousins

Name: _____

2. You are studying aphids, a major pest of plants. There are two aphid strains, green and red. The red strain infects ripe (red) tomatoes, helping it to evade predators while resting on the plant. Large quantities of special metabolites called keratins are needed to make the red color. The Aphid genome project has just concluded the sequencing of the red aphid genome and you set to analyze the aphid genes and chromosomes.

(a 6 points) First, you suspect one gene as the important one for making red coat color. You compare its sequence to that of all other sequences in the NIH's databases for protein coding sequences. The bet's result is shown in the DP matrix below. What is the score of the best **global** alignment? Mark its **path** in the DP matrix. **How many gaps** are in this alignment?

		P	A	A	L	P	G	T	S	D
	0	-4	-8	-12	-16	-20	-24	-28	-32	-36
Q	-4	-1	-3	-7	-11	-15	-19	-23	-27	-31
W	-8	-5	-2	-5	-9	-13	-17	-18	-22	-26
P	-12	-4	-6	-3	-3	-5	-9	-13	-17	-21
A	-16	-8	0	-4	-5	-5	-5	-8	-12	-16
L	-20	-12	-4	-1	0	-3	-7	-8	-11	-15
P	-24	-16	-8	-5	1	4	0	-4	-8	-12
T	-28	-20	-12	-9	-3	0	4	6	2	-2
S	-32	-24	-16	-13	-7	-4	0	4	11	7

score = 7

gaps = 5

+2 for score

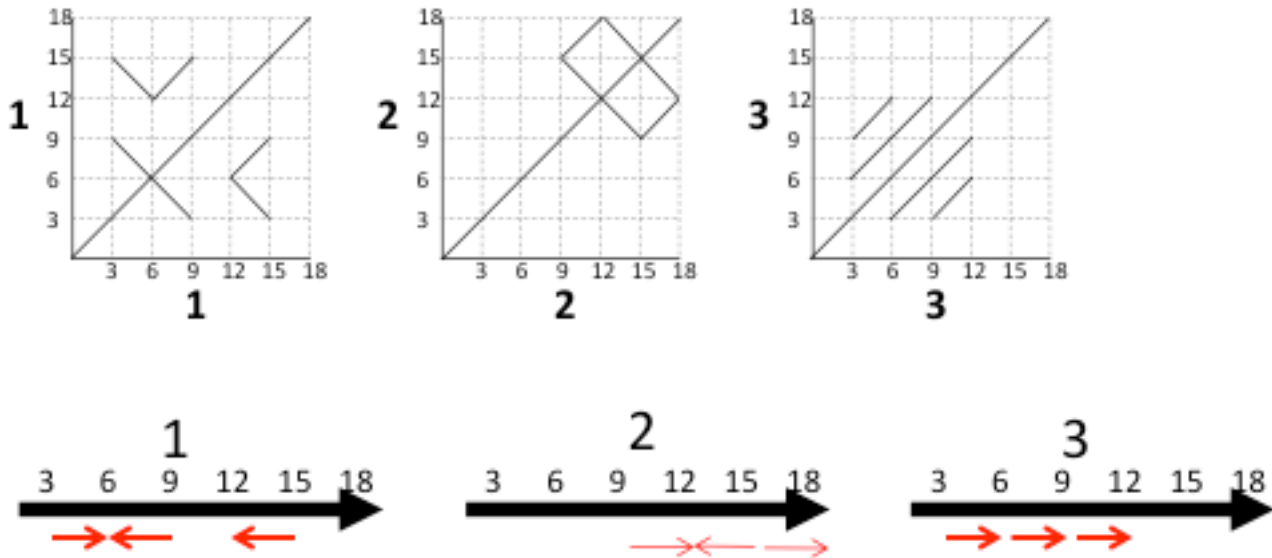
+2 for gaps

+2 for correct path

(no double penalty for wrong gap answer consistent with incorrect path)

Name: _____

(b 12 points) Since your first line of investigation did not yield very promising results, you now study the organization of each of the three red aphid chromosomes (1,2,3) using a dot plot of each chromosome's sequence against itself. Draw the organization of each chromosome, marking repetitive sequences by arrows, demarcating their relative organization and start and end positions. Your helpful friend already solved one of the cases (chromosome 2), as illustration.



[2 points per arrow]

(b 5 points) What would be candidate regions to search for enzymes important for infecting red plants? Write their coordinates (e.g. Chromosomes 7, 15-18). Explain your answer in one sentence.

Chromosome 1: 3-9, 12-15

Chromosome 2: 9-18

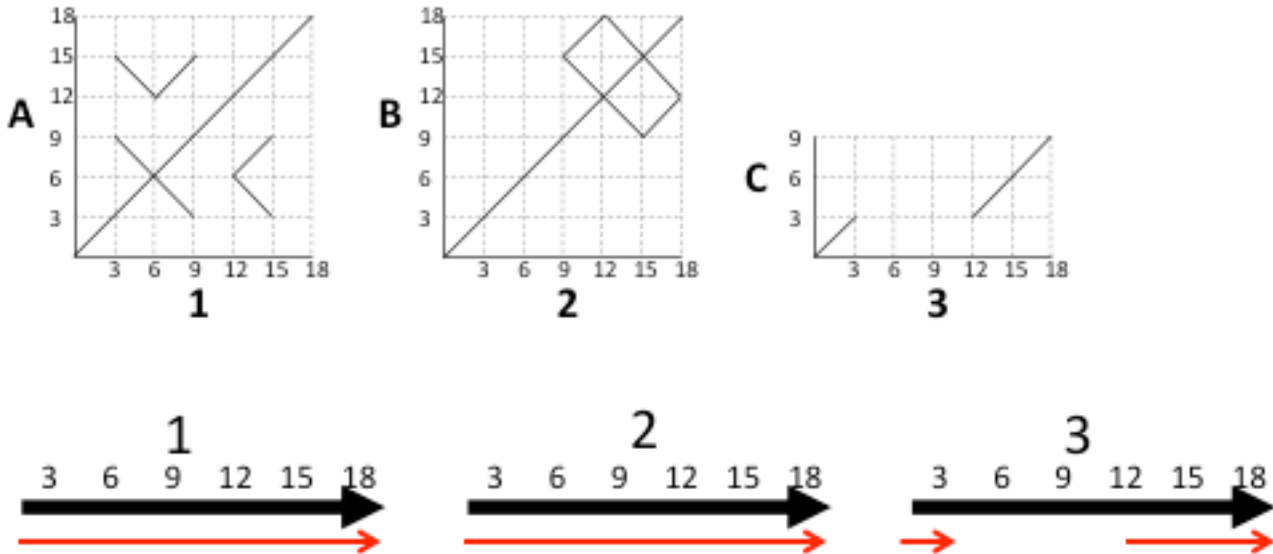
Chromosome 3: 3-12

Many copies of keratins are needed for red color, so we'd expect multiple copies of the keratin-producing gene; therefore, we would look at repeat regions.

[1 point per region; 1 point for explanation]

Name: _____

(c 6 points) To gain better understanding, you now compare the red aphid chromosomes (1,2,3) to those of the green aphid (A,B,C). Draw the organization of the green chromosomes (A,B,C) relative to the red ones (shown as arrows with coordinates below). (Hint: if a red aphid chromosome is identical to a green aphid chromosome, just draw one straight arrow for it)



[2 points per chromosome]

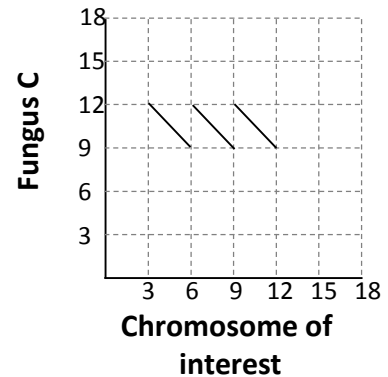
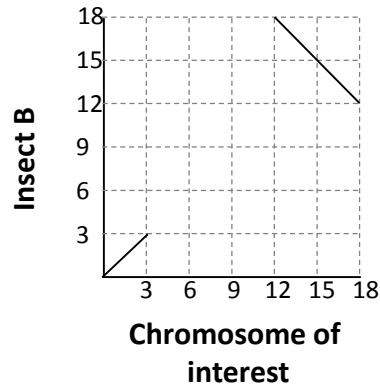
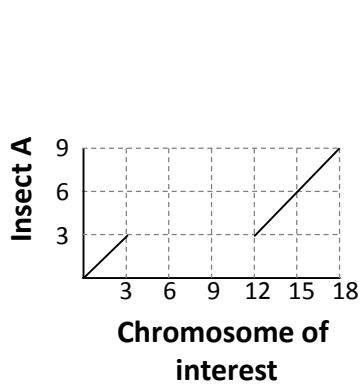
(d 3 points) Given your answer in (c) which region is the best candidate for the enzymes important to make red color? Explain your answer in one sentence.

Chromosome 3, 3-12. This region is present in red aphids but not in green aphids, so it is a likely candidate for the difference in phenotype between the two species.

[2 points for region, 1 point for explanation]

Name: _____

(e 2 points) To understand the source of these sequences, you now compare your chromosome of interest to those of all other fully sequenced species. You obtain the following results. Where did the genes for your enzymes of interest likely come from? Indicate the species and coordinates and explain your reasoning in one sentence.



Fungus C. The candidate region for the genes that code for keratins is not present in Insect A or Insect B, but it is present in Fungus C.

[1 point for correct species, 1 point for explanation]

Name: _____

3. You study a population of unusual rodents. You follow two genes. One of the genes affects hair, and has a Hairy and a Bald allele. Another affects eye color and has a Pink and a Green allele.

Please use three decimal places in your calculations, to avoid rounding errors.

(a 7 points) You measured 674 gametes from mainland rodents, and obtained the following sample:

P _{HairyGreen}	P _{HairyPink}	P _{BaldGreen}	P _{BaldPink}	N
20	4	600	50	674

What are the frequencies of each of the alleles in this population?

$$f(\text{Hairy}) = 24/674 = 0.036 \quad +2$$

$$f(\text{Bald}) = 650/674 = 0.964 \quad +2$$

$$f(\text{Green}) = 620/674 = 0.920 \quad +2$$

$$f(\text{Pink}) = 54/674 = 0.080 \quad +1$$

Name: _____

(b 7 points) What are D and r^2 (allele correlation)?

$$D = P_{HG} * P_{BP} - P_{HP} * P_{BG}$$

$$D = (20/674) * (50/674) - (4/674) * (600/674)$$

$$D = 0.030 * 0.740 - 0.006 * 0.890$$

$$D = 0.002 - 0.005 = \mathbf{-0.003}$$

+2 for equation, +2 for correct D value (points awarded w/in rounding error)

$$r^2 = D^2 / f(H) * f(B) * f(G) * f(P)$$

$$r^2 = 9 \times 10^{-6} / 0.003$$

$$r^2 = \mathbf{0.003}$$

+2 for equation, +1 for correct r^2 value (points awarded w/in rounding error)

(c 8 points) Is the degree of allele correlation (r^2) you found in the sample statistically significant? Show your statistical test (P values are in the table below).

Is the population in LD or LE?

(Consider 0.05 as your P-value threshold to reject the null hypothesis.)

	Probability of exceeding the critical value				
df	0.1	0.05	0.025	0.01	0.001
1	2.706	3.841	5.024	6.635	10.828

H_0 = in LE

H_A = in LD

$$\chi^2 \sim r^2 * \text{Number of gametes}$$

Name: _____

$$= 0.003 \times 674 = 2.022$$

DF=1 so according to the χ^2 chart, $p > 0.1$

Therefore we accept the null hypothesis, the population is in LE

-1 if did χ^2 test for D and not r^2

-2 if correct interpretation, but incorrect math

-2 if squared r^2

-2 if doubled the number of gametes

Full credit awarded if carried through incorrect r^2 value from part b but otherwise calculations and logic was consistent.

(d 8 points) You found that a small population of **50** rodents has just been shipped off to inhabit a new island, which was not inhabited by the rodents before. This founder population had the following haplotypes.

P _{HairyGreen}	P _{HairyPink}	P _{BaldGreen}	P _{BaldPink}	N
18	0	65	17	100

Is the population in LD or LE? Explain, without any calculation or statistical test.

LD, only 3 of 4 possible haplotypes are represented in this new population, so we will have maximal LD

+4 for LD identification

+4 for correct explanation

No credit given if did not state that missing a haplotype is why we are in LD. Small populations do NOT guarantee LD- you would need to do a statistical test.

Name: _____

Name: _____

Grading section

Question 1 36 points:_____

Question 2 34 points:_____

Question 3 30 points:_____

Total :_____