PLEASE: Put your name on every page and SHOW YOUR WORK. Also, lots of space is provided, but you do not have to fill it all! Note that the details of these problems are fictional, for exam purposes only.

1. (18 pts) A large cat breeding business struggles with inherited kidney failure, which kills many kittens before they are old enough to reproduce. A database search suggests that the kidney failure is caused by the b allele of locus B. You measure genotype frequencies at this locus in newborn kittens and in the survivors at one year of age:

Genotype	BB	Bb	bb	Total
Kittens	321	159	20	500
Adult cats	272	154	8	434

(a) (5 pts) Compute a χ^2 to test whether the **adult cats** are in H-W for their own allele frequencies. A χ^2 table is on the last page of the exam.

Allele frequencies in adult cats:
pB=(272+0.5*154)/434=0.804; pb=1-pB=0.196

obs	exp	χ^2
272	280.54	0.260
154	136.78	2.168
8	16.67	4.509
		χ^2=6.937

Some students used the kitten counts or kitten allele frequencies as expectations, which is not what I asked for so I took off 2 points. If you did this there would be 2 df instead of 1 (because you didn't have to use the adult data to get the allele frequencies) and I graded (b) accordingly.

(b) (1pt) How many degrees of freedom does this test have?

One. Start with 3, lose one for allele frequencies and one for totals.

(c) (1pt) Do you **accept or reject** the null hypothesis of Hardy-Weinberg proportions at the 5% significance level?

Reject. The critical value is 3.84 and we are above it. With only one point to work with, I had to mark answers that said "accept" wrong even if accompanied by an explanation that it's not in H-W. If you said 2 df it's still reject.

(d) (2 pts) Calculate the fitness of the three genotypes by comparing the newborn kittens and the adult cats. (Since the kittens are in H-W proportions we can assume there are no deaths due to this locus before birth.) Normalize so that the highest fitness is 1.0.

Find survival of each genotype, then divide through by the highest.

Genotype	BB	Bb	bb
Survivorship	0.847	0.969	0.4
Normalized	0.874	1.0	0.413

(e) (5 pts) The manager proposes eliminating all bb animals, even if they live to adulthood, from the breeding program. Give a numeric prediction of the frequency of **the bb genotype in newborn kittens** if this strategy is continued for a very long time. *Note that this approach reduces the fitness of bb from 0.413 to 0.0 as they aren't allowed to breed: I took off 2 if this was missed.*

s = 1-0.874 = 0.126, t = 1-0.0 = 1.0
pB = t/(s+t) = 0.889 so pb = 0.112
pbb = 0.0125
I took off 1 point for answers that gave pb rather than pbb.
Some students tried 1 generation calculations, but that doesn't answer the question.

Problem 1 continued

(f) (4 pts) The manager then proposes using only BB animals for breeding. He argues that eliminating kidney failure will result in a cat population with better survival (higher fitness) than the current one. Do you agree that this plan will create a cat population with higher fitness? Explain your reasoning. You can disregard concerns about whether reducing the size of the cat population will lead to bad effects of genetic drift; assume the population of BB animals is large enough to avoid this.

No. This is an overdominant system, and the equilibrium is the maximum population fitness. (You could calculate this but it's not necessary.) A population of all BB is less fit than one at the equilibrium, even though some kittens die of kidney disease.

I gave 1 point for "No" without explanation or to explanations involving genetic drift or inbreeding, which I specifically ruled out in the question.

2. (19 pts) A wild salmon population on the Silver River contains 10,000 breeding individuals. The Black River, in a similar habitat, has been blocked off by a dam and has no wild salmon population. We start a fish hatchery on the Black River by catching 100 fish from the Silver River, breeding them in the hatchery and releasing their offspring into the Black River below the dam. In subsequent years we catch 100 fish from the Black River each year and breed them; all other Black River fish fail to breed due to the dam. (Remember that salmon normally come back to their home river to breed.) This situation persists for a long time.

(a) (3 pts) Which population (Silver or Black) will be more likely to fix a mildly harmful mutation? Why?

Black, because N_e is smaller so drift is stronger and can overcome selection.

(b) (5 pts) We measure selective pressures, and find that most of the selection on young Silver salmon involves avoiding predators, whereas most of the selection on young Black salmon involves resisting disease. On average, do we expect mating of "stray" Black fish with the Silver population will increase or decrease the fitness of the Silver population? Explain briefly.

Decrease, for two reasons. Black will have more bad alleles, as shown in (b), due to its small size; and the Black genome is adapted to a different environment and may contain alleles that are bad in the Silver environment. It is possible that Black could increase the genetic diversity of Silver–I gave 2 points for this answer–but given that Black is such a small population, it's not likely to have much to offer.

This is why hatcheries are controversial for salmon management.

(c) (4 pts) A government agency argues that adding a few Silver fish to the Black population each year will improve its fitness. Give one reason why this might be true.

Adding a few new genotypes could let Black get rid of bad alleles that had fixed, and generally improve genetic diversity.

(d) (4 pts) ... and one reason why it might be false. *Silver fish adapted to a different environment might have alleles that are bad in the Black environment. They might also introduce genetic diseases that have died out in Black or never existed there, though this is less likely (2 pts) as the Silver genotype will have fewer bad alleles in general. I also gave full points to the epidemiological issue that mixing the fish might spread infectious diseases to Black.*

Problem 2 continued

(e) (3 pts) A colleague proposes to do a genetic survey of ocean salmon; she will lump Black and Silver salmon together as they are visually identical. Do you expect that she will generally find Hardy-Weinberg proportions in a random sample of ocean fish? If not, should she expect more homozygotes or more heterozygotes?

No; there will be more homozygotes for any loci where Black and Silver have different allele frequencies (the Wahlund Effect). I gave only 1 point for any answer that implied Black and Silver interbreed in the ocean.

3. (10 pts) The following analysis contains some logic errors or use of inappropriate formulas. (Assume that all arithmetic is done correctly.) Identify at least two errors and explain why they are wrong. Optionally, you can identify three errors and I will grade the best two.

In a rabbit population, adult weight is affected by many genes of individually small effect. A researcher does the following:

- He calculates total variance V_T as the average squared difference of each rabbit from the population mean and gets 100 g^2 (grams squared).
- He calculates environmental variance V_E by comparing identical twin rabbits, and gets 40 g^2.
- He calculates genetic variance V_G as $V_T - V_E = 60$ g^2.
- He then assesses a proposed breeding program to breed smaller rabbits. The current mean of the population is 1500 g, and the mean of the subset chosen for breeding will be 1400 g. He computes the expected response to selection in one generation as $R = S(V_G/V_T) = 100g * 0.6 = 60g$ and predicts that the next generation will weigh an average of 1440 g.
- His target is rabbits which weigh 1000 g. He plans to use a selection intensity of 100 g every generation, and predicts that it should take no more than 7 generations to reach his target based on losing 60 g per generation.

(a) (5 pts) First error

The researcher calculated broad-sense heritability V_G/V_T but needed narrow sense V_A/V_T. Some of his genetic variation will be unhelpful stuff like "males are larger".

(b) (5 pts) Second error

He assumed that if he gets a certain response in one generation, he'll continue to get it; but the response will likely go down as genetic variability is used up and/or natural selection interferes. If you followed his reasoning you'd think that in about 25 generations you would get rabbits of size 0, which is clearly false.

(c) (Optional) Third error

Comparing identical twins will underestimate V_E because they also share key environment factors: in particular, unless you use surrogate mothers, they share the pre-birth environment. (This is a nasty example of V_{GE} as twins have both high genetic similarity and high environmental similarity, since they develop in the same womb.)

I said there were no math errors, but some students noted that 60g/generation for 7 generations is only 420 of the needed 500 grams, and that's also a valid answer.

4. (28 pts) The Red Drum, a (non-fictional) fish of the Gulf of Mexico, has been estimated to have a headcount population size N (number of fish in the sea) of 3,000,000 but an effective population size N_e of 3000. Compare this species to a fictional fish, the Blue Drum, with both headcount and effective population size of 3,000,000. Assume that both species have had stable population sizes for a long time, and that there is no gene flow. Also assume that selection on individual loci is similar in both species.

   (a) (2 pts) If we sample a large number of genes across the genome, which species will have more heterozygotes?

   Blue. The effective population size is larger, so more variants can be maintained in the face of drift, and that increases heterozygosity.

   (b) (3 pts) At a particular locus, fitnesses in both species are as follows:

Genotype	AA	Aa	aa
Fitness	0.92	0.86	1.0

   Assuming that both populations start out fixed for A and a single copy of a arises by mutation in each species, which population is more likely to end up **fixing** a? (You do not need to calculate the probability.) Explain briefly.

   Red. This trait is underdominant and in a large population you'll lose the rare allele, but in a small population it could become frequent by chance. I gave only one point for answers with no coherent explanation.

   People who considered this a favorable allele reasonably concluded that Blue will fix it more often, but this ignores the fitnesses.

   (c) (8 pts) It has been suggested that in Red Drum most females never successfully reproduce, but a few lucky females have thousands of offspring each. It is not known whether Red Drum have cooperative breeding, in which one female gives up her own reproduction in order to help another rear her young. Assume that the common ancestor of Red and Blue Drum did not have cooperative breeding. Which species do you think would be more likely to evolve it? Explain.

   It's more likely in Red. A female who is unlikely to succeed in reproducing has less to lose if she helps another. If females have any way to predict their likely success, those with poor chances could optimize their fitness by helping female kin who have a better chance. (For example, small fish might help while large ones breed, if small fish can seldom breed successfully.) A second valid answer is that the low N_e increases relatedness and thus the benefit of helping.

   Some students answered the question "Do you think Red actually breeds cooperatively now?" and I gave up to 6 points for that.

   (d) (8 pts) Would cooperative breeding be more likely to evolve if Red Drum return to the river mouth where they were born in order to spawn, or if they choose a river mouth at random? (The Gulf of Mexico has a large number of river mouths.) Explain briefly.

   Definitely if they return to the same river mouth. Altruism via kin selection needs to involve helping kin, and you are much more likely to be around kin if all fish go back to where they were born in order to mate.

   I gave 4 points to the alternative explanation that going to the same river mouth reduces environmental variability in outcomes and thus enhances selection. This might be true but it doesn't seem like a sure thing.

Problem 4 continued

(e) (7 pts) An alternative explanation for the observed N_e values is that our assumption of stable population size over time is wrong. Briefly explain how the unusual Red Drum N_e could be caused by variation in population size.

If Red Drum had had a bottleneck in the past, their N_e might still be low, reflecting slow recovery from the bottleneck (as in cheetahs). If Red Drum oscillate in size like snowshoe hares, their N_e might reflect the bad years more than the good ones.

I also gave credit for answers involving very unequal sex ratio in reproductive fish, though that's not quite what I asked.

I gave only 1 point for "We estimated N or N_e wrong." This is a trivial explanation for ANY observation but doesn't really answer the question.

5. (17 pts) The human genome contains a (fictional) HLA gene, HLA-P. Compared to other HLA genes, only the first half of the protein coding sequence is translated due to an early STOP codon. This removes the part of the protein which inserts into the cell membrane, so instead of sitting on the surface of the cell, HLA-P protein falls off into the liquid surrounding the cell.

A researcher proposes that HLA-P is a pseudogene in humans with no useful function. To test this, she measures silent (synonymous) and coding (nonsynonymous) substitutions among many human copies of HLA-P. She looks only at the part which is translated.

	silent	coding
Observed mutations	14	15
Potential sites for mutation	24	117

(a) (5 pts) Compute the ω statistic from her data.

dN is coding mutations per coding site or 15/117=0.128. dS is silent mutations per silent site or 14/24=0.583. Their ratio dN/dS is 0.220.

A lot of people forgot to use the potential-sites data, but it is essential to make this test work. Silent sites are naturally rarer than coding sites due to the structure of the genetic code, and you must correct for this.

(b) (2 pts) Does this value of ω support her hypothesis? Explain why or why not. (You do not need to do a significance test.)

I graded this according to the number given in part (a). Her hypothesis predicts ω around 1.0. If it is much larger or smaller this is likely not a pseudogene. In fact she got a low ω which strongly suggests purifying selection.

(c) (5 pts) In other HLA genes, ω is around 3.0 for the first half of the gene and about 0.2 for the second half (the part that's missing in HLA-P). What type of selection could cause that finding?

In other HLA loci the first half of the gene is under directional or diversifying selection and the second half is under purifying selection. The first half has either a rapidly changing mission or one that rewards diversity. (In fact it's the latter: there is heterozygote advantage and probably also frequency-dependent selection.) The second half has an important, stable mission so mutations there are usually lost to selection.

I gave 3 points for a correct answer for either half of the gene. I think some students were surprised to see two kinds of selection on the same gene.

Problem 5 continued

(d) (5 pts) What do these ω values suggest about differences between HLA-P and other HLA genes?

They are under different selection: other HLA loci value change or diversity in the first half and consistency in the second half, whereas HLA-P values consistency in the first half and clearly doesn't need the second half at all. Probably these genes have very different functions.

This problem was inspired by old data on the HLA-G gene but does not accurately represent what is now known about this gene.

6. (8 pts) Vertebrate cells have several genes whose sole function is to cause the cell to self-destruct ("apoptosis") if it receives a particular signal. It has been argued that the existence of apoptosis alleles is a problem for evolutionary theory because if the allele kills the cell carrying it, it obviously is not transmitted to the next generation—and if it does not kill the cell, it has no phenotype and should drift randomly, which means it will usually be lost.

Briefly critique this argument: if possible, suggest how such a trait can evolve.

A propensity to self-destruct under certain circumstances like viral infection or cancer can be selected for because the germ cells, with the same genome, will have an increased chance to reproduce and pass on the allele due to this self-destruction. I did not give full credit for answers that just said "It's good for the organism or population" without mentioning the close genetic relationship between the self-destructing cell and the reproductive cells. Be very careful about "good for the population" arguments–it is very hard to select for things that are good for the population but not good for self or kin. This one is clearly good for kin.