

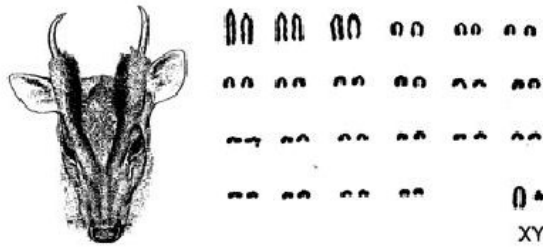
These are for your own information only; I won't be collecting or grading them. A solution key will be available on the Web. Some are a little longer and more open-ended than actual exam questions but otherwise they should be comparable.

1. Here are some useful dates in hominid history:

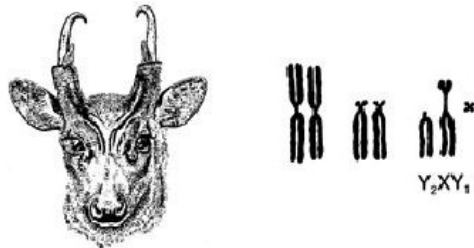
Human vs. chimp/bonobo	4.5 million years ago
Chimp vs. bonobo	1.0 million years ago
Modern human vs. Neanderthal	400,000 years ago
Human mtDNA "Eve"	200,000 years ago
Extinction of Neanderthals	40,000 years ago

- (a) There is considerable evidence that the nuclear genomes of Europeans contain Neanderthal alleles. However, no Neanderthal mtDNA has been found in any human population. Give two different possible reasons for the discrepancy between nuclear DNA and mtDNA.
- (b) Comparison of the bonobo and chimp genome sequences shows that almost all genes diverged about 1 million years ago; genetic diversity is similar in the two species. Chimps and bonobos have quite different appearance and behavior. In captivity, however, they freely mate and produce fertile offspring of both sexes. Modern-day chimps and bonobos live on opposite sides of the Congo River and are poor swimmers. Based on these facts, describe a likely scenario for the origin of these species, including presence or absence of post-speciation contact and gene flow.
- (c) The human and chimp genome sequences have very similar gene order overall. However, gene order on the Y chromosome is extremely different due to multiple paracentric (non-centromere-containing) inversions. Explain briefly why paracentric Y rearrangements might be better tolerated than paracentric autosomal rearrangements.
- (d) Would we expect the human and chimp X chromosomes to have similar gene order, like the autosomes, or very different gene order, like the Y? Explain briefly.
2. Two candidates for the closest living relative to the elephant are the hyrax (a small land herbivore) and the manatee (a large aquatic herbivore). All three are definitely more related to each other than they are to sheep.
- (a) Draw two **rooted** trees: one showing the hypothesis that hyraxes are the closest relative to elephants, and one showing the hypothesis that manatees are the closest relative to elephants. Use sheep as the outgroup in both trees.
- (b) For the following (fictional) protein sequence data, determine the parsimony score of each tree in part (a). Mark mutations on the trees, and choose the best tree under the parsimony criterion.
- |              |     |     |     |     |     |     |     |
|--------------|-----|-----|-----|-----|-----|-----|-----|
| hyrax (H)    | Ser | Arg | Glu | Phe | Ala | His | Leu |
| elephant (E) | Lys | Arg | Glu | Phe | Ala | Cys | Arg |
| manatee (M)  | Lys | Arg | Cys | Trp | Ala | Cys | Arg |
| sheep (S)    | Lys | Ser | Cys | Phe | Val | His | Leu |
- (c) These data give very poor bootstrap scores. We decide to sequence another gene, and identify five candidates. For each candidate, briefly explain why it would or would not be a good choice for clarifying the elephant/hyrax/manatee relationship.
- A highly conserved housekeeping gene
  - A moderately conserved gene involved in oxygen transport; variation in this gene affects how long an animal can hold its breath
  - One gene from the moderately conserved odor-receptor gene family, a large gene family involved in the sense of smell
  - A moderately conserved gene involved in digestion of plant material
  - A segment of "junk DNA" far from any genes
3. We examine three (fictional) varieties of flowering plants: Red snapdragon, Blue snapdragon, and White snapdragon. They have many differences besides flower color, and may or may not be separate species. As an outgroup we use Arabidopsis (the standard lab flowering plant).

- (a) We choose one individual from each species and sequence 100 different loci. When we infer an evolutionary tree for each locus, 54 loci favor a tree which connects Red and White, and 46 favor a tree which connects Blue and White. Surprisingly, almost every locus has a bootstrap value over 90%. Assuming that we did not make a mistake in choosing or applying our phylogeny method, propose an explanation for these results and describe how your explanation fits the data.
- (b) We count chromosomes and find that Red has 4 pairs, Blue has 5 pairs, and White has 9 pairs. Given this additional information, what is the likely origin of the White snapdragon? Be as specific as possible.
- (c) A useful insect-resistance mutation occurs in White snapdragon. Would you expect it to spread to Blue or Red? Why or why not?
4. We collect some unusual silver salamanders. All individuals appear female. They have three copies of each chromosome (triploid). When we keep them isolated, no offspring result. In the same area we find some diploid blue salamanders of both sexes. When we put silver females in with blue males, they produce silver-colored daughters; despite several generations of adding more blue males, no blue daughters are ever produced, and no sons.
- (a) Are silver salamanders probably self-fertilizing, cloning, or reproducing sexually? Explain briefly. If more than one answer is possible, explain why.
- (b) Based on your answer to the previous question, describe two problems that silver salamanders may encounter over the evolutionary long term.
- (c) Given the problems described in (b), why do species like this arise in the short term?
5. *Note: This problem is pretty hard!* In the mosquito *Culex pipiens*, sex is determined by a single locus at which males have genotype  $M/m$  and females have genotype  $m/m$ . A closely linked locus distorts the sex ratio. It has two alleles,  $D$  and  $d$ . Homozygous  $d/d$  males produce 80% male offspring because their  $M$ -bearing chromosome destroys their  $m$ -bearing chromosome during spermatogenesis. The  $D$  locus has **no effect on females or heterozygous males**.
- (a) We examine a wild population containing only  $MD$  and  $mD$  chromosomes. Given what you know about sex ratio in stable populations, what are the likely frequencies of these two chromosomes? Explain briefly.
- (b) We make a cage population containing 100  $mD/mD$  females and 50  $MD/mD$  males from the wild population, and add 50  $Md/md$  males. In the first generation of offspring, what is the sex ratio? Assume that the  $M$  and  $D$  loci are so close together that they are never separated by recombination.
- (c) In the first generation of offspring, what is the frequency of the  $Md$  chromosome? Has it increased, decreased, or stayed the same?
- (d) We allow this population to evolve for a dozen generations or so, and examine a neutral genetic variant present on the  $md$  chromosome close to the  $M$  and  $D$  loci. Would we expect this variant to have increased, decreased, or stayed at about the same frequency as in the starting cage population? Briefly explain your reasoning.
- (e) The discoverers of this system were not sure whether  $M$  and  $D$  were two separate, closely linked loci or one locus (with four alleles  $MD$ ,  $Md$ ,  $mD$  and  $md$ ). Based on your answers to previous questions, do you think  $d$  would have an easier time becoming frequent in a population if it were a part of the  $M$  locus, or a separate locus that could occasionally be separated from  $M$  by recombination? Explain.
6. In a hypothetical marine invertebrate, male offspring can be produced cheaply—a mother can generate 10 male offspring for the same energy and resource expenditure as 1 female offspring, because males come from smaller eggs. Assume that this creature is diploid, reproduces only by male-female mating, and cannot self-fertilize or clone. Also assume that offspring sex is determined by genes in the mother. What is the equilibrium sex ratio in this species? Why?



Chinese muntjac deer  
*Muntiacus reevesi*



Indian muntjac deer  
*Muntiacus muntjak*

7. Chinese muntjac deer have a fairly normal mammalian chromosome set. Indian muntjacs have very few, large chromosomes and unusual sex chromosomes. For each of the following (hypothetical) factors, explain how it might have contributed to the transition between the Chinese and Indian karyotypes:

- Population bottleneck (tiny population size)
- Inbreeding
- Large numbers of transposons or other repeated sequences
- Mutations in genes involved in DNA repair
- Habitat fragmentation
- Harem keeping by male muntjacs

8. A short (fictional) DNA sequence is sampled from five primate species (only variable sites are shown):

Species				
Human (H)	GTC	GCA	ATA	TGT
Chimpanzee (C)	GTA	GCA	TTC	TGC
Bonobo (B)	GTA	GCA	TTC	TGT
Gorilla (G)	ATA	CCG	TAC	TGT
Rhesus (R)	ACA	CTA	TAG	CCG

- Make a distance matrix (raw distances with no corrections) for these five species.
  - Draw a UPGMA tree from your distances.
  - Modern genetic evidence strongly suggests that humans, chimpanzees and bonobos group together on the tree (the “third chimpanzee” hypothesis). Previously it appeared obvious that chimps, bonobos and gorillas must be most closely related because they all walk on their knuckles, while humans do not. How could this disagreement between genetics and morphology be explained? Note that knuckle-walking is not purely a behavioral trait—if you try it yourself you will discover that human arms and legs have the wrong proportions for successful knuckle-walking.
9. A morning glory plant is produced by fertilization of a normal haploid egg (1N) by an abnormal diploid sperm (2N), so that the resulting plant is triploid (3N).

- (a) Give two ways in which this triploid plant could start a new species of morning glory. Be sure to explain how it could manage to reproduce successfully.
- (b) Which of your two ways seems most promising in starting a species that will last for a long time? Why?
- (c) The plant hormone auxin is essential for normal plant development, but only one copy is needed. Would you expect the **allele frequency** of non-functional alleles of this gene to be higher or lower in a triploid morning glory species (assume that it has been triploid for many generations) than in the related diploid species? Why? (Please note that this question asks about allele frequency, not phenotype frequency!)

10. As a researcher, you have three genetic markers to choose from:

- A. A pseudogene with no function.
- B. A gene coding for a moderately conserved protein, such as alcohol dehydrogenase.
- C. A gene coding for an extremely conserved protein, such as cytochrome oxidase.

Which marker would you use for each of the following experiments? Briefly explain each answer. If you feel you need further information, explain what information you would need.

- (a) A study to determine how the different orders of mammals (primates, bats, rodents, whales, etc) are related to each other.
  - (b) A coalescent analysis of migration patterns in humans.
  - (c) An attempt to determine whether a newly-discovered hot-spring organism is a bacterium or an archaeobacterium.
11. A variant form of the mosquito  $Y$  chromosome, called  $Y^*$ , is discovered. A male with  $Y^*$  transmits it to 100% of his offspring (who are therefore all sons).

We find a wild population of 1000 ordinary  $XY$  males and 1000 ordinary  $XX$  females, and in an attempt to destroy them, dump in 100  $XY^*$  males. Assume that each male has an equal chance to reproduce and can produce an equal number of offspring.

- (a) Initially, what are the relative frequencies of  $X$ ,  $Y$  and  $Y^*$ ?
  - (b) After one generation, what will be the relative frequencies of  $X$ ,  $Y$  and  $Y^*$ ? (Remember that each mating must involve 1 male and 1 female.)
  - (c) Also after one generation, what will be the proportions of males and females?
  - (d) What kind of genetic event could save the mosquitoes from population extinction due to lack of females?
12. Two visibly different types of butterfly exist in the same region. Hybrids between them are never found in the wild. Their caterpillars eat different host plants. When adults of different types are put together in the lab, they do not attempt to mate. Artificial insemination can produce viable and fertile hybrid offspring, but these offspring do not appear very healthy – they develop slowly and have a high mortality rate as caterpillars.
- (a) Would you consider these butterflies separate species? Is there any additional evidence you need?
  - (b) Can you say anything about likely modes of speciation?

13. The Greek islands of Rhodos and Crete were connected to the mainland until about 5.3 million years ago, when the Mediterranean flooded. Since then they have been separated by salt water. Suppose that originally the area had one species of water frogs and one species of small songbirds. Water frogs cannot cross saltwater at all and were completely isolated after the flooding, whereas songbirds can occasionally fly from one island to another.

Today we observe that each island has its own species of frogs and of songbirds. In answering the following questions, you are not expected to use specific biological knowledge about birds and frogs; the essential difference is that birds can migrate and frogs, in this situation, cannot.

- (a) Would we expect more pre-mating reproductive isolation in birds or in frogs?
- (b) Which mode or modes of speciation likely explain the bird species?

- (c) A researcher proposes that the Rhodos water frog originated by peripatric (tiny isolated population) speciation. What would this predict about the genetic diversity of Rhodos frogs compared to mainland frogs?
- (d) We sample twenty protein-coding genes from the songbirds and draw a phylogenetic tree of each gene. To our surprise, the trees for different genes do not agree. About half the genes group Rhodos songbirds with Crete songbirds; the other half group Rhodos with the mainland. Give two different explanations for the discrepancies in these trees. Assume that we used an appropriate phylogeny method and that all of our trees have high bootstrap scores.
14. The common ancestor of all modern human mtDNA appears to be about 200,000 years ago (“Mitochondrial Eve”).
- (a) If we assume that this is about the expected value, roughly how long ago would we expect the common ancestor of a random nuclear locus to be? Don’t forget that nuclear loci are diploid and are contributed by both parents, while mtDNA is haploid and contributed by the mother only.
- (b) Roughly how long ago would we expect the common ancestor of a Y chromosome gene to be?
- (c) Give three (or more) reasons why the common ancestors of mitochondria and Y chromosomes might be at different time depths.
- (d) The common ancestor of modern alleles at the HLA locus HLA-DR is earlier than the human/chimpanzee split (6 to 8 million years ago). Why might this locus have a much older common ancestor than the average locus? (HLA genes are involved in immune system recognition of pathogens and cancer cells.)
15. In the bacterium *E. coli* and its relatives, housekeeping genes (genes which code for proteins involved in DNA replication, repair, transcription, translation, and basic metabolism) are seldom successfully transferred among species. Other genes (genes for exploiting a particular food source, resisting pathogens and toxins, antibiotic resistance, etc.) are frequently transferred among species, even distantly related species.
- (a) If we draw a tree of a housekeeping gene, what kind of information will we be able to gain from it?
- (b) If we draw a tree of an antibiotic resistance gene, what kind of information will we be able to gain from it?
- (c) What would you expect from a tree made by mixing housekeeping and non-housekeeping genes?
16. We sample a specific odor-receptor gene from humans, mice, rats, and dogs. The length of the branch leading to humans is much greater than expected based on current theories of the relationship among primates, rodents and carnivores.
- (a) Why might this be?
- (b) What traits would you look for in the DNA sequence in order to confirm or disconfirm your theory?

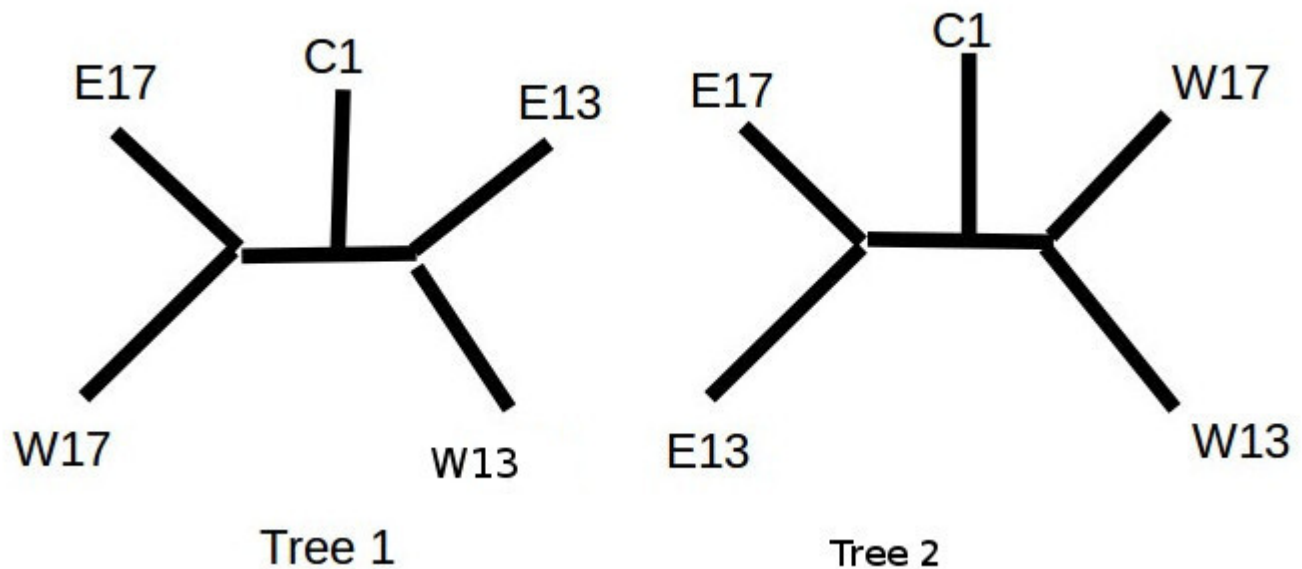
\*\*\*\*\*

The remainder of these practice problems are the entire midterm from 2016. This should give an idea of the likely scope of a midterm.

(The details of these problems are fictional, for exam purposes only; they don’t represent real data. –Except for *P. esculentus*, which behaves just as described in the exam; I could never invent anything so weird. )

17. (23 pts) We search a country and find four species of periodic cicadas: eastern 17-year (E17), eastern 13-year (E13), western 17-year (W17), and western 13-year (W13). (The “year” designations tell how long the insect remains underground as a larva before emerging to mate and die.) We also find a species of one-year cicada (C1) which is clearly an outgroup to the periodic cicadas.

The following unrooted trees show two hypotheses about relationships between periodic cicadas. (There are other possibilities, but we will focus on these two.)



- (a) (4 pts) Draw the two trees as rooted trees, using C1 as an outgroup. (Please draw them large!)
- (b) (7 pts) We collect protein-coding sequences from the five species and find the following data (only variable sites are shown):

Position	1	2	3	4	5	6	7	8
E17	Val	Ser	His	Val	Arg	Glu	Gln	Trp
E13	Val	Lys	His	Val	His	Glu	Gln	Trp
W17	Val	Lys	Trp	Val	Arg	Ser	Gln	Trp
W13	Val	Lys	Trp	Ser	Arg	Ser	Val	Trp
C1	Ala	Lys	His	Val	Arg	Ser	Gln	Phe

Next to each of your trees on the previous page, write the parsimony score for these data on that tree. Circle the score of the tree that is preferred by parsimony. Marking changes on the branches may be helpful for partial credit but is not required. NOTE: for your convenience, these data are reprinted on the tear-off sheet (last page of exam).

For the following two questions, make use of this additional information:

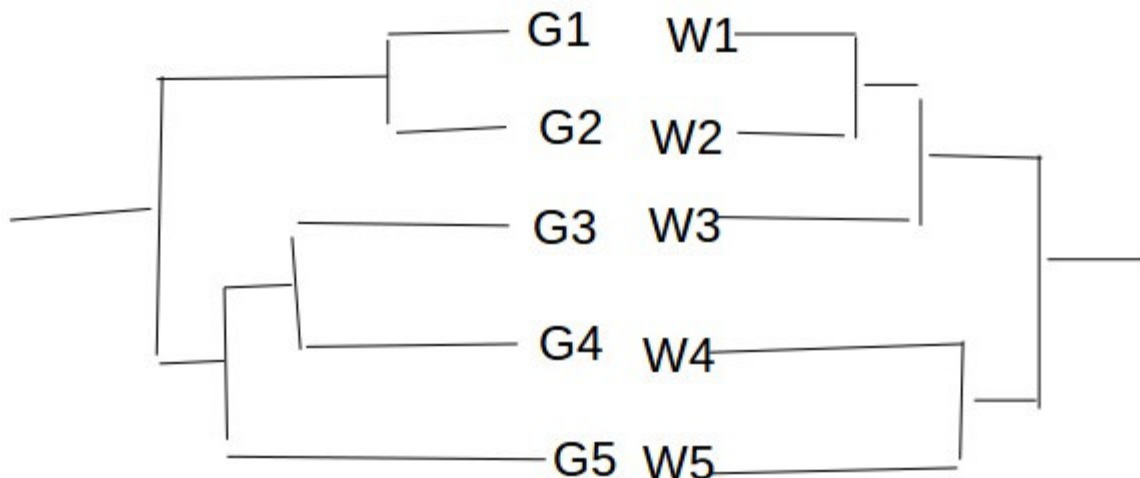
There is no geographic overlap between eastern and western cicadas due to a mountain range. In the east, however, E13 and E17 live in the same area, and in the west W13 and W17 live in the same area. C1 lives throughout the country.

The two 17-year species emerge in the same year; the two 13-year species also emerge in the same year (almost always a different year than the 17's).

When a cicada emerges in the wrong year it generally makes a mistake of 4 years. A 17-year cicada will thus occasionally emerge in 13 or 21 years. Genetic factors can influence how often a cicada makes a mistake, and whether it is early or late.

- (c) (6 pts) If Tree 1 is correct, give a hypothesis for the origin of the four periodic cicada species. What speciation events probably happened, and in what order?
- (d) (6 pts) If Tree 2 is correct, give a hypothesis for the origin of the four periodic cicada species. What speciation events probably happened, and in what order?
18. (30 pts) The bacterium *Wolbachia* lives within insect cells, and is generally transmitted only from mother to offspring. In a certain gnat species, *Wolbachia* manipulates the sex ratio, causing 3/4 of offspring of an infected female to be female. This does not cause a reduction in total offspring produced.
- (a) (6 pts) Suppose that gnat males are fully capable of fertilizing all females even when the males are only 1/4 of the population. Would there still be evolutionary pressure to return the sex ratio to 50/50? Explain briefly.

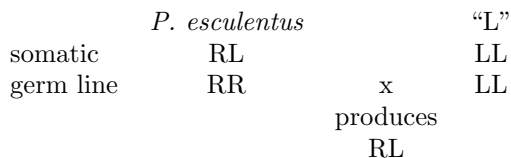
- (b) (6 pts) Suppose that mutations in *Wolbachia* increase its ability to create female offspring, so that fewer and fewer males are produced. In response, infected female gnats might begin to produce fertile eggs by mitosis, creating offspring that are genetic copies of their mother. What long-term problem will this strategy cause for the gnats?
- (c) (6 pts) Alternatively, infected gnats might become hermaphrodites that produce both eggs and sperm and self-fertilize. What long-term problem will this strategy cause for the gnats?
- (d) (6 pts) No one has ever found a strain of *Wolbachia* that transforms insects into males. Why not?
- (e) (6 pts) A researcher wishes to know if gnat *Wolbachia* is ever able to infect a new host (as opposed to being inherited from the mother). She draws phylogenetic trees of gnat species and the *Wolbachia* that infect them, with the following results:



In these trees, G1 is gnat species 1, and W1 is the *Wolbachia* that infects it; similarly for G2, G3, etc. Do these data support the theory that gnat *Wolbachia* occasionally infects new hosts? Explain briefly. You may want to circle aspects of the trees that support your answer.

- 19. (32 pts) The frog *Pelophylax esculentus* arose from an ancient hybridization of two species, which for simplicity I will call L and R. As far as we know, the original hybridization only happened once.

In preparation for meiosis, *P. esculentus* destroys its L genome and duplicates its R genome (in germ line cells). It then undergoes meiosis and mates with an L individual. The offspring therefore have one L genome and one R genome, and are *P. esculentus* again. This has been going on for thousands of years and has produced a widespread population of *P. esculentus*. An illustration:



- (a) (4 pts) We catch several *P. esculentus* frogs and sequence their L and R genomes. Would we expect more linkage disequilibrium (non-random association of alleles along the chromosome) among the L genomes or the R genomes? Explain briefly.  
For the following three questions, please predict if the new mutation will **increase** due to natural selection, **decrease** due to natural selection, or **drift randomly**, and explain your answer briefly.
- (b) (5 pts) A mutation arises in the L species which causes it to reject mating with *P. esculentus*. Assume that L x *P. esculentus* matings and L x L matings produce the same number of surviving offspring on average. What will happen to this mutation in the L population?

- (c) (5 pts) A **harmful recessive** mutation arises in the **R genome** of a *P. esculentus* individual. Assume that this mutation does not act in the germ line, and that it does not exist in the L species. What will happen to this mutation in the *P. esculentus* population?
- (d) (4 pts) A paracentric (non-centromere-containing) inversion arises in the **R genome** of a *P. esculentus* individual. Assume that the inversion itself does not change any genes. What will happen to this mutation in the *P. esculentus* population?
- (e) (4 pts) A paracentric (non-centromere-containing) inversion arises in the **L genome** of a *P. esculentus* individual. Assume that the inversion itself does not change any genes. What will happen to this mutation in the *P. esculentus* population?
- (f) (5 pts) Given your answers to the previous questions, what do you predict the R genome of *P. esculentus* will look like if this species lasts for a long time?
- (g) (5 pts) Logically, crossing two *P. esculentus* should produce an R individual:

	<i>P. esculentus</i>		<i>P. esculentus</i>
somatic	RL		RL
germ line	RR	x	RR
		produces	
		RR	

*P. esculentus* prefer not to mate with each other, but when they do, the tadpoles (baby frogs) always die early in development. Based on your answers so far, what is a likely reason for the inviable tadpoles?

20. (15 pts) Three children with the same father, but different mothers, all have abnormal brain development. When their genomes are sequenced, each child is found to have one normal copy of chromosome 3 and one abnormal copy which is missing several genes and has extra copies of several other genes.

The mothers are developmentally normal, and each has two normal copies of chromosome 3.

The father is developmentally normal, but unfortunately his genome is not available. The clinician who collected these data is very surprised that the three children appear to have inherited their problem from their **unaffected** father.

- (a) (5 pts) If we could sequence the father's copies of chromosome 3, what do you predict we would find? Be as specific as possible, and be sure to account for both copies. You may draw a picture, but it is not required.
- (b) (5 pts) Does your explanation predict that future children of his will all be developmentally abnormal, or might some healthy children be produced? If so, how would they be produced? (Assume that any child with a deletion/duplication genotype will be abnormal.)
- (c) (5 pts) Is the father's genetic abnormality likely to be due to a recent event (for example, in him or one of his parents) or is it plausibly an ancient variant in the human population? Explain briefly.