EVOLUTIONARY GENETICS (Genome 453)

Practice problems for Final

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. 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?

Selection favors equal investment, which in this case would be making 1 female for every 10 males, or a sex ratio of 10:1 in favor of males. While the rare females will produce more grandchildren, they are more expensive. If females become rarer than 1 per 10 males, it will become advantageous to produce more females, and the female frequency will go up; if they become more common than 1 per 10 males, males will be more advantageous and the female frequency will go down. You can check this mathematically by counting grandchildren. At 10:1 an individual which specializes in male or female offspring will not gain more than the average number of grandchildren.

2. 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:

(a) Population bottleneck (tiny population size) Large karyotype changes are usually bad in the heterozygote (underdominant). A bottleneck could enable rapid fixation of the new form due to strong genetic drift.

(b) Inbreeding Karyotype changes are bad in the heterozygote. Inbred populations have fewer heterozygotes and thus suffer less from the bad consequences of such changes.

(c) Large numbers of transposons or other repeated sequences Repeated sequences can lead to mis-pairing in meiosis and rearrangement of the genome. If the Indian muntjac had an abnormally large number of repeated sequences scattered across its chromosomes—perhaps due to a recently activated transposon, or a virus— that might explain why its genome became so massively rearranged.

(d) Mutations in genes involved in DNA repair Failure of DNA repair might increase the frequency of chromosome rearrangements.

---

[Image of Chinese and Indian muntjac deer]
(e) Habitat fragmentation A fragmented habitat can encourage inbreeding and bottlenecks, both of which favor fixation of underdominant traits like karyotype changes.

(f) Harem keeping by male muntjacs Unequal reproductive success, such as one male monopolizing many females, reduces effective population size. This can promote fixation of underdominant traits. For example, if a successful male carries a new chromosome rearrangement, it may have a very high frequency in the next generation and get a big boost toward fixation.

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

<table>
<thead>
<tr>
<th>Species</th>
<th>GTC GCA ATA TGT</th>
<th>GTA GCA TTC TGC</th>
<th>GTA GCA TTC TGT</th>
<th>ATA CCG TAC TGT</th>
<th>ACA CTA TAG CCG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human (H)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chimpanzee (C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bonobo (B)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gorilla (G)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhesus (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Make a distance matrix (raw distances with no corrections) for these five species.

<table>
<thead>
<tr>
<th></th>
<th>H</th>
<th>C</th>
<th>B</th>
<th>G</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>–</td>
<td>3</td>
<td>7</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>4</td>
<td>–</td>
<td>1</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>1</td>
<td>–</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>G</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>R</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
<td>–</td>
</tr>
</tbody>
</table>

(b) Draw a UPGMA tree from your distances.

UPGMA will cluster C+B, then C+B+H, then C+B+H+G.

Reduced matrices for each step:

<table>
<thead>
<tr>
<th></th>
<th>H</th>
<th>CB</th>
<th>G</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>–</td>
<td>3.5</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>CB</td>
<td>3.5</td>
<td>–</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>G</td>
<td>7</td>
<td>4.5</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>R</td>
<td>11</td>
<td>9</td>
<td>7</td>
<td>–</td>
</tr>
</tbody>
</table>

HCBG | R
---|----
HCB | 5.75| 10 |
G  | 5.75| –  |
R  | 10  | 7  |

HCBG R
---
HCB | 8.5 |
G  | –    |
R  | 8.5  |

In Newick format the final tree is:

((((C:0.5,B:0.5):1.25,H:1.75):1.125,G:2.875):1.375,R:4.25):0.0;

We did not cover Newick format in class, but you can try to check your tree against this one.

(c) Modern genetic evidence suggests that humans, chimpanzees and bonobos group together on the tree (the “third chimpanzee” hypothesis). Previously it was thought that chimps, bonobos and gorillas must be 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. Either humans lost knuckle-walking, perhaps as an adaptation to their plains lifestyle, or chimps and gorillas evolved it independently. A less likely possibility is species hybridization which allowed knuckle-walking to spread from, say, gorillas to chimps. Alternatively, the genetic evidence could be wrong for some reason, such as genetic convergent evolution, but across the whole genome this is highly unlikely.

4. 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. I know of at least four:

i. Cloning, so no sexual reproduction required.

ii. Double all of its chromosomes so that it becomes hexaploid (6N) and can now go through meiosis normally.

iii. Double its chromosomes in oogenesis and halve them in spermatogenesis, or vice versa, so that sexual 3N offspring can be produced. I have never heard of a plant doing this—only amphibians—but theoretically it could happen.

iv. Lucky break in meiosis that produces a gamete with exactly 1N or 2N. (1N does not seem likely to produce anything new, but if two 2N gametes are made, you could get a tetraploid species.)

(b) Which of your two ways seems most promising in starting a species that will last for a long time? Why? Cloning can lead to long-term problems (unless you are a rotifer) due to Muller’s Ratchet. Hexaploidy and tetraploidy should be fine once they are established. The sexual-triploid strategy seems hard to get started but should be okay once it’s established. If I had to bet, I’d bet on hexaploidy: this has happened in cultivated tobacco and wheat.

(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!) Higher. The alleles can hide in a heterozygote a lot more easily in a triploid. If it is asexual they would never risk coming together at all, and if it is sexual they would need to come together in 3’s in order to be eliminated.

Please be careful of questions like this. If you are asked for allele frequency, don’t mistakenly give phenotype or genotype frequency instead.

5. 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. Probably the moderately conserved protein is best. The pseudogene is unlikely to be recognizable among such distant species, and cytochrome oxidase will probably not have enough changes to be useful.

(b) A coalescent analysis of migration patterns in humans. I would prefer the pseudogene as it will have the most variability and no selection to confuse the issue. Even a moderately conserved protein won’t vary among humans enough to help, unless it is selectively important (Duffy, hemoglobin, etc) and then I’d worry about convergent evolution.

(c) An attempt to determine whether a newly-discovered hot-spring organism is a bacterium or an archaeabacterium. This must be a highly conserved protein (ribosomal RNA would also work). Nothing else will be reliably recognizable among such distantly related species.

6. 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*? Just by counting chromosomes, we find \( p(X) = \frac{3100}{4200} = 0.738 \), \( p(Y) = \frac{1000}{4200} = 0.238 \), \( p(Y^*) = \frac{100}{4200} = 0.024 \). Check that they add to 1.
(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.) All females give \( X \) to their offspring so we can ignore them. All we care about is the relative frequency of \( XY \) males and \( XY^* \) males. We have 1100 males, 91% \( XY \) and 9% \( XY^* \). So 91% of our offspring will come from normal males; half will be \( XX \) and half \( XY \). 9% will come from supernales; all will be \( XY^* \). This means \( p(X) = 3/4 \ast 0.91 + 1/2 \ast 0.09 = 0.7275 \), \( p(Y) = 1/4 \ast 0.91 = 0.2275 \), \( p(Y^*) = 1/2 \ast 0.09 = 0.045 \). The frequency of \( Y^* \) has approximately doubled due to its meiotic drive. Check that these add to 1.

(c) Also after one generation, what will be the proportions of males and females? We’ll have 45.5% females and 54.5% males.

(d) What kind of genetic event could save the mosquitoes from population extinction due to lack of females? Females could develop the ability to recognize and reject \( XY^* \) males; a female who could do so would have daughters, which would spread her genes rapidly among a mostly-male population.

An \( X^* \) chromosome mutation which overcame the meiotic drive of the \( Y^* \) so that \( X^*Y^* \) produced some daughters would also do the trick (though if all offspring inherited \( X^* \) you would just replace one problem with another!)

There could be an “arms race” between \( X \) and \( Y \); some researchers feel they see evidence of this in genomes.

An autosomal mutation that blocked the drive of \( Y^* \) would get into daughters and spread as a result.

Catching Wolbachia, which tends to transform males into females, could help—again, it might eventually replace one problem with another.

If the too-frequent sex were females, parthenogenesis (female cloning) could help, but as far as I know there are no species in which males can reproduce themselves without females.

7. 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

(a) Would you consider these butterflies separate species? Is there any additional evidence you need? I would consider them species by just about any definition; they have separate gene pools and probably cannot share adaptations.

(b) Can you say anything about likely modes of speciation? There is strong pre-mating isolation and weaker post-mating isolation. This suggests species that arose in the presence of gene flow—parapatric, peripatric, or sympatric speciation. The different host plants make sympatric speciation, which is usually difficult, seem fairly likely here. By the way, if words like “peripatric” cause trouble you can paraphrase.

8. 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? In birds. Frogs don’t have any selection for pre-mating isolation as they never meet each other.

(b) Which mode or modes of speciation likely explain the bird species? If you don’t remember the Latin names, paraphrase.

(c) Parapatric probably fits the best (speciation with some gene flow).

(d) 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? If the Rhodos frogs came from a tiny isolated population, they should have low genetic diversity compared to the mainland because of lower population size.

(e) 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
9. 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 not. About four times longer, or 800,000 years. One factor of two comes from diploidy and the other from inheritance through both males and females.

(b) Roughly how long ago would we expect the common ancestor of a Y chromosome gene to be? The same as mtDNA, or 200,000 years.

(c) Give three (or more) reasons why the common ancestors of mitochondria and Y chromosomes might be at different time depths. Statistical fluctuation (the coalescent has high random variability). Different effective population sizes in males and females due to either different survival or different reproductive success (Genghis Khan effect). Different migration rates in males and females. Natural selection on mitochondria or Y chromosomes.

(d) The common ancestor of 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.) Balancing selection. This gene probably has overdominant alleles (or alleles which are favored when rare—we can’t be sure which) and therefore multiple alleles have been preserved throughout the history of both species, all the way back into the common ancestral population.

10. 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 nearly unrelated species.

(a) If we draw a tree of a housekeeping gene, what kind of information will we be able to gain from it? This probably carries information about the relationships between bacterial species, or at least the housekeeping parts of their genomes.

(b) If we draw a tree of an antibiotic resistance gene, what kind of information will we be able to gain from it? This is more likely to give a history of gene transfer events, and may tend to clump together species which share a common set of environmental challenges, rather than ones whose genomes are overall closely related. For example, all hospital bacteria might cluster together.

(c) What would you expect from a tree made by mixing housekeeping and non-housekeeping genes? Confusion! If you mix genes that come from different trees of descent, the resulting tree will be mixed up and impossible to interpret.

11. We sample a specific smell-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? The gene may be unused or even dysfunctional in humans, while it is still functional in the other species. Alternatively, it might be under rapid selection for a new function in humans, while it retains a conservative function in the other species. The first theory is more likely given that humans are known to have a poor sense of smell, but the second can’t be ruled out.

(b) What traits would you look for in the DNA sequence in order to confirm or disconfirm your theory? If the gene is a pseudogene in humans, it may have \( \omega \) close to 1, and may also show damage such as stop codons, frameshifts, and very nonconservative changes. If it is being selected for a new function, it may have \( \omega \) greater than 1, but will be free from blatant damage (unless the new function only requires the 5’ end, which sometimes happens). In the other species where the gene is presumably functional, it will have \( \omega \) less than 1. We could also do an HKA test using a non-selected sequence such as an intron as the control. If the polymorphism/divergence ratio in the human gene is similar to that of an intron, the gene is probably not important to humans.
A chimp or gorilla sequence would also be useful here, to see if this receptor is changing unexpectedly fast among other primates.