

MIT OpenCourseWare
<http://ocw.mit.edu>

HST.161 Molecular Biology and Genetics in Modern Medicine
Fall 2007

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.

Problem Set 2: Answer Key

Problem #1

The incidence of cystic fibrosis is 1 in 1600 live births in Denmark, 1 in 3600 live births in Bulgaria and 1 in 10,000 live births in Hong Kong.

a. Calculate the allele frequency for normal and CF alleles in each population.

Denmark

CF disease incidence = $1/1600 = q^2 \rightarrow q = 1/40 = 0.025 = \text{CF allele frequency}$

Normal allele frequency = $p = (1-q) = 39/40 = 0.975$

Bulgaria

CF disease incidence = $1/3600 = q^2 \rightarrow q = 1/60 = \text{approx. } 0.017 = \text{CF allele frequency}$

Normal allele frequency = $p = (1-q) = 59/60 = \text{approx. } 0.983$

Hong Kong

CF disease incidence = $1/10,000 = q^2 \rightarrow q = 1/100 = 0.01 = \text{CF allele frequency}$

Normal allele frequency = $p = (1-q) = 99/100 = 0.99$

b. What is the risk of having a child with CF of a couple in which the wife is from Denmark and the husband is from Hong Kong, assuming no previous family history of CF for either parent?

Also reasonable to assume neither parent has CF, since as we learned they would know this by the age they were having kids.

Risk of wife (Denmark) being a carrier = chance of having 1 CF allele = $2pq = 2 * (1/40) * (39/40) = 0.049$

$(1/40) * (39/40) = 0.049$

Risk of husband (Hong Kong) being a carrier = chance of 1 CF allele = $2pq = 2 * (1/100) * (99/100) = 0.019$

$(1/100) * (99/100) = 0.019$

Risk of affected child (both parents must be carriers AND the child must inherit 2 CF alleles, which has a $1/4$ chance if both parents are carriers):

$.049 * .019 * .25 = 2.3 * 10^{-4}$

c. A woman from Bulgaria marries a man from Denmark who has a sister with cystic fibrosis. What is this couple's risk of having a child with CF?

Risk of wife (Bulgaria) being a carrier = chance of having 1 CF allele = $2pq = 2 * (1/60) * (59/60) = .033$

Risk of husband being a carrier, given that he has a sister with CF = $2/3$

His parents must both be carriers (genotype Aa where A is the normal allele and a the CF allele). If we didn't know anything about the husband, his possible genotypes and the probability of each would be $1/4$ aa, $1/2$ Aa, $1/4$ AA. But we know he is not affected (no chance of being aa) so instead his probabilities are $2/3$ Aa, $1/3$ AA. So his chance of being a carrier is $2/3$.

Risk of affected child = $.033 \times 2/3 \times 1/4 = .0055$

d. A cystic fibrosis test has been developed which identifies the 32 most frequent mutant CF alleles in the Danish population. The allele frequency of these alleles sum to 92% of the mutant CF alleles present in the Danish population. When the same test is applied to the Bulgarian population, the alleles it identifies sum to 78% of the mutant alleles in the Bulgarian population. Three couples (a Danish couple, a Bulgarian couple, and a couple in which the wife is from Bulgaria and the husband is from Denmark) are considering taking this test. How much will the risk of having a cystic fibrosis child change for each couple if each member of the couple has a negative test outcome?

Start with the Danish population. The normal CF allele frequency is 1/40. The risk of having two carriers is $2pq \times 2pq$, and then multiply by $1/4$ chance of transmission for having a child with CF in absence of the test: $.049 \times .049 \times .25 = .0006$

Now they will both take the test. We can assume the test is 100% specific and 100% sensitive for 92% of the diseased alleles since we are given no contrary information. Essentially, we have changed the allele frequency for diseased CF alleles in this population, with the new disease allele frequency being: $1/40 \times .08 = 1/500$

Thus our new carrier frequency, $2pq$, = $.004$

The couples' chance of now having a child with CF is $.004 \times .004 \times .25 = .000004$

This is a risk reduction of $.0006/.000004 = 150$ fold

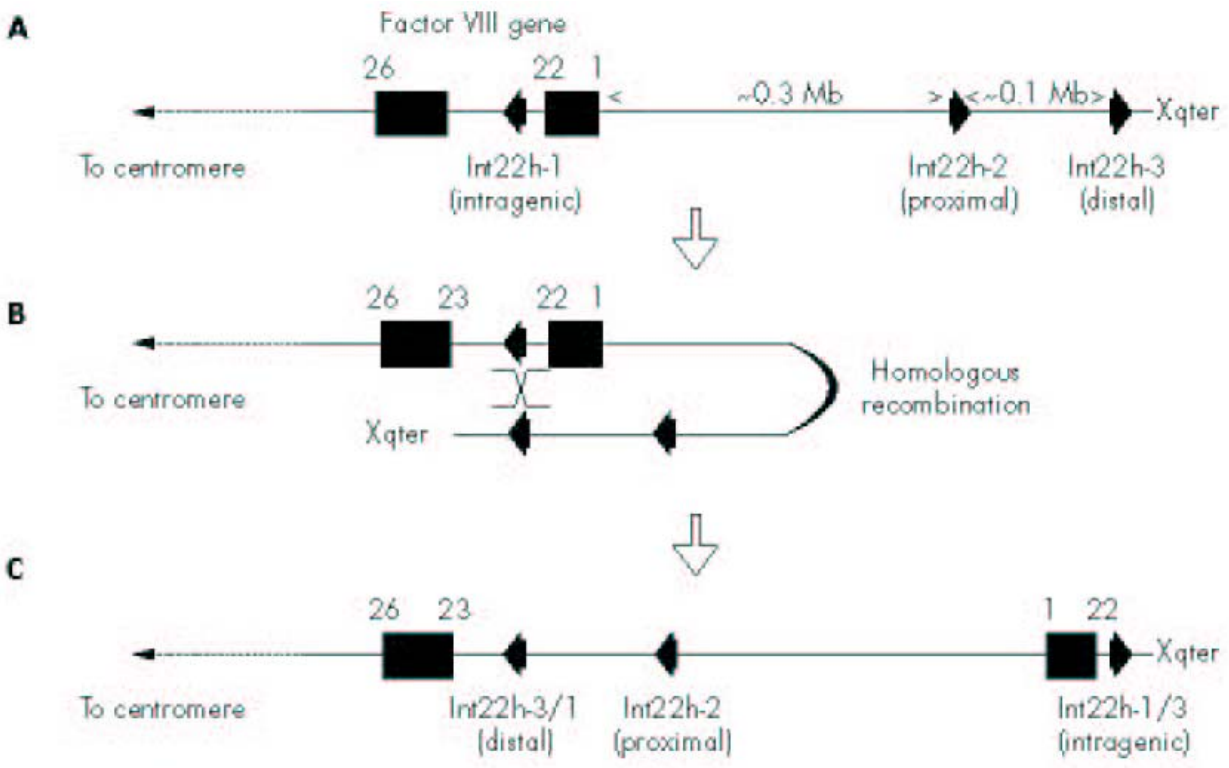
In the Bulgarian population, the original diseased allele frequency was 1/60. This leads to a carrier frequency of .033 (do the math!). This leads to an odds of having an affected child of $.033 \times .033 \times .25 = .00027$. We effectively eliminate 78% of these alleles, giving a new diseased allele frequency of $1/60 \times .22 = .00367$. Again our new carrier frequency will be $2pq = 2 \times .00367 \times .9963 = .0073$. Our new chance of having an affected child will be $.0073 \times .0073 \times .25 = .000013$ (Note, the Bulgarian population with the test is at a relatively higher risk than the Danish population because the test is not as good in this population1) Here the risk reduction is only $.00027/.000013 = 21$ fold, a significant difference! (Note, most doctors would actually describe these data to their patients in this way using this kind of analysis. As in, "if you take this test and it has a negative outcome for both of you, it will reduce your risk of having a child with CF by 150 fold!")

Bulgarian/Denmark couple: Risk in absence of test is $.049 \times .033 \times .25 = .0004$

In the presence of the test, use the new/undetected allele carrier frequencies: $.004 \times .0073 \times .25 = .0000073$, For a risk reduction of $.0004/.0000073 = 55$ fold

Problem #2

Draw the possible outcome of crossover events between one of the repeat sequences located between 250 and 300 kbp upstream from the start of transcription of the factor VIII gene and the intron 22 copy of the repeat sequence. What transcription product would be produced by the gene produced as a product of this recombination event? Why would this outcome be expected to cause severe hemophilia A?



Due to this inversion event, an intact transcript will not be produced. Thus, factor VIII protein will not be present, and Hemophilia A will result.

What would be the outcome of PCR for each of the exons of the factor VIII gene from patients with severe hemophilia A as a consequence of a crossover event involving the repeat sequences?

PCR of each exon will yield normal products! All the DNA is still present, and will still be present in the same 5' to 3' orientation. The entire segment of exons 1-22 will be inverted and facing the other direction on the chromosome, but since all the sequence is still present in the same orientation PCR will proceed normally. (2 pts)

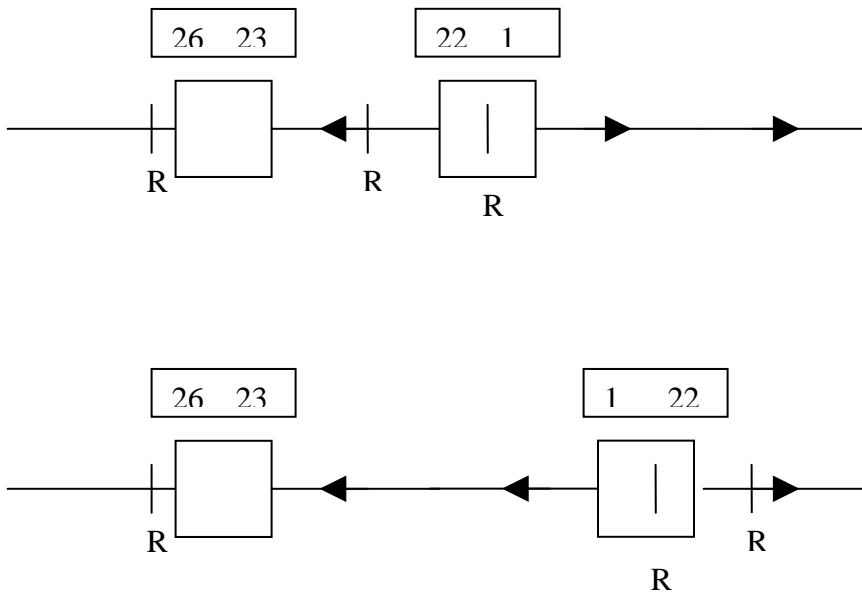
Describe how Southern blotting with a cDNA probe for factor VIII could reveal the presence of a crossing over event which causes the loss of function of the factor VIII gene.

The inversion is likely to alter the position of restriction sites relative to the coding sequence. For example, imagine the following pattern, where "R" represents a restriction

fragment cleavage site, and there are three sites for “R”, one 3’ to exon 26 (left), one between the repeat and exon 22, and one in the exon 1 to 22 region.

In this case, in the inversion, the restriction site would now be inverted relative to exons 23 to 26; thus, for example, probing with the cDNA fragment from exons 23 to 26 would detect a longer fragment using restriction enzyme “R” than previously.

Note that this makes the assumption that the fragment sizes as shown are appropriate for Southern blotting. Note also that there are many other possible solutions to this problem, which are conceptually identical. Note finally that whole genome analysis methods at present, like exon-specific PCR, are not suitable for detecting such rearrangements. Credit was given for any reasonable answer that demonstrated understanding of the concepts involved. *The inversion is unlikely to change hybridization!* Recall from the first TA problem session that a few bases of sequence complementarity can be enough to get probe binding. After inversion, all the sequence information is still present, it is the SIZE of fragments during a Southern blot that changes, not necessarily the signal intensity. (2 pts)



Explain how a PCR assay could be designed to detect the inactivation of the factor VIII gene as a consequence of such an event.

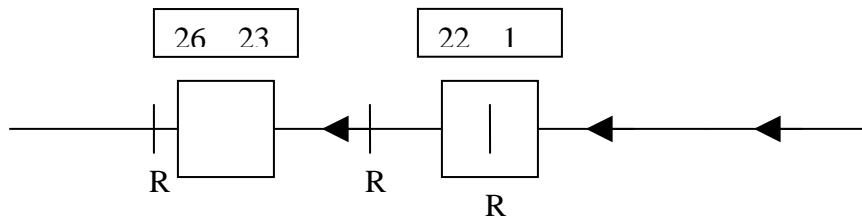
Specific PCR primers could be designed for genomic DNA sequencing flanking the inverted repeat elements. You should realize that any priming event with primers on both sides of the point of inversion would yield a simple +/- test in a PCR reaction for whether or not inversion occurred. The key point is that the PCR is flanking the inversion site, and not located entirely within the inverted sequence (which would yield no difference b/w WT and inverted).

Many of you came up with more elaborate methods. If they made sense, you got credit. (2 pts)

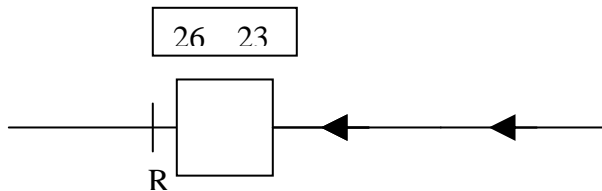
What would be the possible genetic consequences if the repeat sequences were present in the same position upstream of the factor VIII gene but in the opposite orientation, while the copy in intron 22 of the repeat sequence was not changed in orientation.

In this case, two different homologous recombination events could lead to severe hemophilia A by completely deleting exons 1 through 22 of factor VIII, as shown below.

Inverted repeats upstream of factor VIII



Recombination between inverted repeats could cause this product.



Problem #3

First, you should note the blood groups that result from each of the following genotypes.

ii – blood group O

$I^A I^A$ and $i I^A$ and $I^A i$ – blood group A

$I^B I^B$ and $i I^B$ and $I^B i$ – blood group B

$I^A I^B$ – blood group AB

From the O group frequency we can determine the allele frequency of i . Then, by solving a quadratic equation in I^A we can determine the I^A frequency. Lastly, since $I^A + I^B + i = 1$, we can solve for I^B .

Thus, for City 1:

$$O = 0.42 \rightarrow i^2 = 0.42 \rightarrow i = \mathbf{0.648}.$$

$$I^A I^A + I^A i + i I^A = 0.30 \rightarrow (I^A)^2 + 2(I^A i) = 0.30 \rightarrow I^A = \mathbf{0.200}$$

$$\text{Therefore, } I^B = 1 - 0.201 - 0.648 = \mathbf{0.152}$$

To check your work, you can multiply $2 * I^A * I^B = 0.06$. This is the AB frequency reported for city 1. **What happens when you do this for City 2? Why?**

The same calculations can be done for the other cities and all are summarized below.

City 1:

$$i = \mathbf{0.648}$$

$$I^A = \mathbf{0.201}$$

$$I^B = \mathbf{0.151}$$

City 2:

$$i = \mathbf{0.62}$$

$$I^A = \mathbf{0.249}$$

$$I^B = \mathbf{0.131}$$

City 3:

$$i = \mathbf{0.70}$$

$$I^A = \mathbf{0.10}$$

$$I^B = \mathbf{0.20}$$

City 4:

$$i = 0.624$$

$$I^A = 0.376$$

$$I^B = 0$$

City 5:

$$i = 0.60$$

$$I^A = 0.30$$

$$I^B = 0.10$$

Based on the calculated allele frequencies, we can easily match the cities to their names.

Due to assortative mating in Monoangeles, this city is clearly not in Hardy Weinberg equilibrium. You will note that when you check $2 * I^A * I^B$ for City 2, you do not get the reported AB frequency. This further confirms that HW equilibrium does not apply for City 2. Therefore, **City 2 is Monoangeles.**

Note that City 2 has allele frequencies between Ariga and Svenska, but is closer to Svenska. Thus, **City 3 is Ariga and City 5 is Svenska.** Note that for each allele, City 1 has frequencies exactly midway between Ariga and Svenska. Thus, **City 1 is Polyangeles** (in which Ariganians and Svenskans marry each other and themselves with equal probability).

City 4 is Cousinia. Because this city was founded by two couples, presumably neither of which had I^B , the I^B frequency in this city is 0. This is an example of the founder effect.

Problem #4

This question is designed to enforce the concept of linkage disequilibrium. The first notable piece of information is that the SNPs mentioned are located in intron 2 of the genes of interest. Therefore, the SNPs in question are tightly linked to (in fact, located *within*) the genes of interest (beta globin, cftr, fgfr2, factor VIII). This means that the chances that the SNP and mutant gene can be dissociated by recombination are vanishingly small.

Basically, if mutations in the gene of interest are ancient, then (almost) all individuals carrying the mutant allele will also have the SNP marker (SNP A in the case of beta globin, C in the case of cftr, etc.). In contrast, if mutations in the gene of interest are new, then presence of a mutant allele will not be strongly associated with a particular SNP in intron 2. With this simple logic, we can now predict what would be expected for each of the diseases mentioned:

Cystic fibrosis – Most individuals with CF carry a mutant allele that of ancient origin. Thus, most of these individuals would be expected to be homozygous for the C allele at the SNP site. Since we know that the allele frequency of SNP C is 0.5, statistically 0.5^2 or 25% of the population would be expected to be homozygous for the C SNP. However, we would expect that among CF patients, almost 100% should be homozygous for the C SNP.

Sickle Cell – As with CF, beta-globin mutations causing sickle cell arose in an ancient ancestor. Thus, we would expect the same situation as above.

Achondroplasia – Most cases of achondroplasia arise by new mutations. Thus, we would not expect the SNP in intron 2 of fgfr2 to be in linkage disequilibrium with the mutant fgfr2 allele. Therefore, among patients with achondroplasia, we'd expect only ~25% of the patients to be homozygous for the C allele.

Hemophilia A - As with achondroplasia, most cases of hemophilia A arise due to de novo mutations. Therefore, we would expect only about 25% of patients with hemophilia A to be homozygous for the G allele in the factor VIII gene.

Problem #5

- a. A father has three daughters who each give birth to a son with Duchenne Muscular Dystrophy. The father does not have DMD himself. Propose an explanation for the occurrence of DMD in this family.

As you learned in class, DMD is inherited in an X-linked recessive fashion. Therefore, males who inherit a diseased X-chromosome are affected. Females are unaffected unless they happen to inherit two defective X-chromosomes. In the scenario in this question, the father's wife is presumably a carrier of DMD. His daughters all happen to be carriers as well – presumably they all inherited the defective X from their mother (and a normal X from their father). These daughters then pass on the defective X to their sons, who have DMD.

- b. Identical twin girls are born. One is diagnosed with DMD, the other is normal. The father is normal, while the mother is shown to be a heterozygous carrier of a deletion of exon 44 of the dystrophin gene. Propose an explanation for these findings.

Males have only one X chromosome while females have two. Dosage compensation refers to the phenomenon whereby one of the two X chromosomes in female cells is randomly inactivated. Although the initial choice of X chromosome for inactivation is random, this decision is heritable once made (i.e. all progeny cells inactivate the same X). In the twin with DMD, the normal X chromosome was presumably inactivated in a progenitor cell that gives rise to most of the skeletal muscle. Therefore, this twin expresses mutant dystrophin in her muscle cells and thus displays DMD. The other twin presumably inactivated the mutant X in her muscle progenitor cells, and therefore has a normal phenotype.