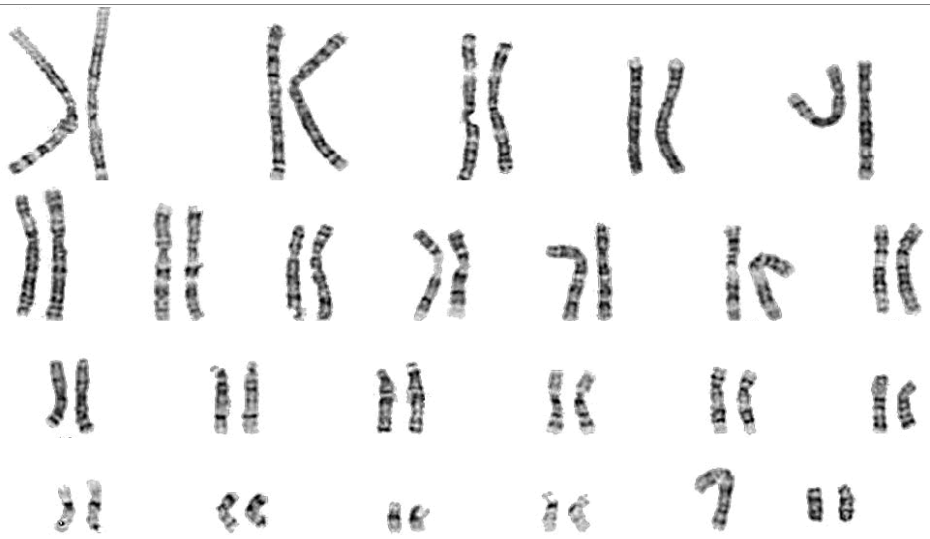
39.314 INTRODUCTORY CYTOGENETICS
FINAL EXAMINATION

December 4 , 2002
Time: 1:00 p.m. to 3 p.m.
Location: 134 Agriculture

This examination consists of questions totalling 100 points, and is worth 35% of the final grade.

---

1. (5 points) The human karyotype below is from an aneuploid individual. What is the aneuploidy?



2. (15 points)
Given the following definitions:

N: Genome size
f_{mr}: fraction of the genome that is middle repetitive
X : $KC_0t_{½\ (pure)}$, Complexity
l_{mr}: average length of a middle repetitive sequence

What do the following equations tell you?

a) $a = f_{mr}N$

b) $b = X/l_{mr}$

c) $c = a/X$

3. (15 points) Three different somatic cells of an individual were examined and were found to contain 59, 62 and 93 chromosomes. One of the cells is normal, one is polyploid and the other is aneuploid (not necessarily in that order).

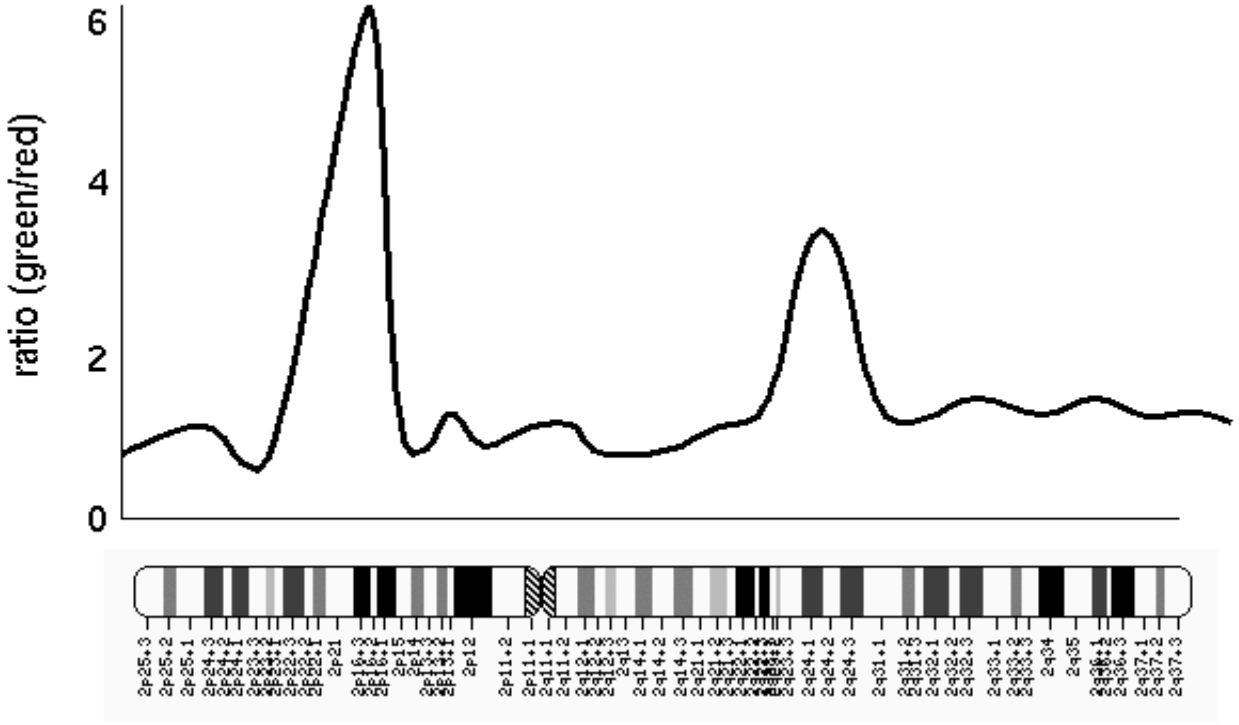
a) What is the haploid number of chromosomes in this species?

b) What is the formula for the polyploid?

c) What is the formula for the aneuploid?

4. (5 points)

Genomic DNA from a normal cell line was labeled using a red fluorescent tag, and DNA from a cancer cell line was labeled using a green tag. Equal amounts of both probes were mixed, and hybridized to human chromosomes. The ratio of green to red fluorescence was measured as a function of position on chromosome 2, as shown above. What do the two main peaks probably indicate?



5. (10 points) Fill in the blanks:

a) Translocations and inversions drive _____ by reducing the viability of gametes in interspecific hybrids.

b) These effects are not seen when the individual is _____ for the translocation or inversion.

c) During meiosis, translocations don't cause duplications or deletions unless _____ occurs.

d) When hybrid progeny are produced from distantly-related species, the number of chromosomes pairing in chains due to translocations will be _____ than when closely-related species are hybridized.

6. (15 points) Suppose you wanted to create a mapping kit from a eukaryotic genome for which complete sequence information was available. You choose to create PCR-based markers based on known sequence tagged sites (STS).

a) STS sequences from the ends of BAC inserts are likely to be a better source of PCR markers for mapping, than would ESTs. Why?

b) Given that STS sequences are the better choice, what do we have to check, for each STS, before it can be used as a molecular marker? In other words, what type of sequence might not be able to uniquely identify a given genetic locus?

c) One might initially think that a good mapping kit could be created by choosing STS markers from clones equally spaced, for example, 10 Mbp apart. For the purposes of genetic mapping, what important consideration is being overlooked?

7. (20 points) Three of the four statements below contain an error. For each statement, indicate how the correct statement should read, and which part has been changed. (You don't need to rewrite the entire statement.) One of the statements has no errors. Indicate which statement that is.

Example: `DNA sequences that intrinsically lend themselves to amplification are called shellfish DNA.`

you could write the following:

`... selfish DNA`

a) Because only a small percentage of eukaryotic genomes codes for proteins, most mutations will occur within non-coding DNA, and will usually be deleterious.

b) After calculating gene order, a maximum likelihood method calculation of linkage distance begins by arbitrarily assigning different random distances between adjacent loci. Distances are systematically varied until the algorithm converges on the map that best explains the data.

c) Maximum likelihood methods for calculating linkage distance work because the closer two loci are, the more likely they are to cosegregate.

d) The Clark and Carbon formula, **N = ln(1-P)/ln(1-f)**, calculates **N**, the number of markers that must be screened to have a probability **P** of finding at lease 1 marker

linked within a distance of **d** megabases, given that each marker covers fraction **f** of the genome. **f** is defined as **2d/G**, where **G** is the size of the genome in centimorgans.

8. (15 points) The figure below shows a chromosomal region from two different species. The location and directions of transcription for 5 genes are shown. Describe possible evolutionary processes that may have occurred since the divergence of these two species. What are the implications for the evolution of genome size?