1. A graduate student would like to clone a human DNA sequence from chromosome 21. She has a hamster tissue culture cell line containing all hamster chromosomes plus 1 human chromosome, chromosome 21.

a) By what simple procedure, using the light microscope, can she check whether the human chromosome is really chromosome 21 and whether roughly the entire chromosome is present? (Check text for ideas)

   Use a chromosome banding procedure – pattern of bands can tell if the chromosome is human chromosome 21 and if it is largely intact (within resolution of banding method)

b) She isolates total DNA from this cell line and uses it to prepare a genomic DNA library in phage. Most inserts are between 10-20 kb of DNA. To find clones corresponding to human DNA inserts, as opposed to hamster DNA, she screens thousands of colonies using DNA hybridization. As a probe she uses a particular human DNA repetitive sequence, named Alu.

   Why does this approach work to selectively pull out human DNA clones? Alu is primate specific, and dispersed throughout genome, present every ~5 kb on average.

   Can every potential sequence from human chromosome 21 be identified using this method? Explain. No- can’t find pieces that lack Alu. Statistically although Alu present over every 5 kb on average, one will find 10-20 kb pieces that lack Alu completely. Moreover Alu sequences are located nonrandomly and are enriched in gene rich chromosome bands.

c. Having made her library, the student now wants to screen a number of clones in order to select several that localize to a particular region of chromosome 21 which is thought to contain genes related to Alzheimer's disease. She does this using the method of in situ hybridization using her DNA probes and hybridizing to mitotic human chromosomes. Much to her surprise, she finds that each of her clones hybridizes over the entire lengths of all human chromosomes. WHY???

   Each 10-20 kb piece is likely to contain not only Alu but also other highly repetitive sequences found throughout the human genome.

   Can you think of a way of modifying her protocol so as to eliminate this problem? (The person who did this became well known and gave the modified procedure a glitzy name- Hint- what should the student add to the hybridization mixture?)
The student needs to add unlabeled, human repetitive DNA. A fraction of DNA (Cot-1) enriched in repetitive sequences can be obtained by melting human DNA and allowing the DNA to re-anneal, and isolating the ds DNA. At early times after re-annealing only the repetitive sequences will have had time to find homologous sequences to hybridize. The unique sequence will be single stranded and not purified.

2. Alu elements, as well as other transposable elements, are conspicuously rare over the four homeobox gene clusters- HoxA, HoxB, HoxC, and HoxD- as illustrated in the figure below for HoxD along with a similar region of chromosome 22 for comparison purposes. The lines going up are exons. The lines going down are transposable elements- these merge at this resolution into solid blocks for much of the ~ 1Mbp regions shown.

WHY DO YOU THINK TRANSPOSABLE ELEMENTS ARE SO RARE OVER THESE HOX CLUSTERS?

![Figure 4-31 MBoCS: The Problems Book © Garland Science 2008](image)

There must be elements between the genes which are important for function- ie proper transcriptional regulation- of these genes. By comparison with other genomic regions, these Hox regions appear to be unusually dense in such intergenic regulatory regions.

(There is an alternative, unlikely but not impossible, explanation, which is based on challenging the annotation. More specifically, one might propose that the regions between genes might in fact code for regions that are transcribed into non-coding RNA. In the last few years there has been increasing evidence of intergenic tx in the Hox regions. If this is the case one would have to conclude though that this noncoding RNA has sequence specific functions.)

3. In class, a cloning scheme was described which was used to identify centromere sequences from yeast.
Having found a centromere sequence from yeast which confers equal segregation of daughter chromatids, a student would now like to address the fidelity of chromosome segregation using this minimal sequence in circular plasmids.

He decides to start with the plasmid he has constructed which contains an ARS sequence, the HIS selectable marker, and the CEN centromere sequence. To this plasmid he adds a yeast gene, Ade 2, including appropriate regulatory sequences.

When Ade 2 gene is missing from yeast cells and they are grown on special types of plates, the yeast colonies become red. When the Ade 2 gene is present, the yeast colonies are white (wild type).

The student obtains yeast cells which lack a functioning Ade 2 gene from a friend. He now transforms his plasmid into yeast cells and plates them on these special plates which he has also made HIS minus. All colonies are white.

a. He now plates them on these special types of plates but in this second case he has used HIS + media to prepare these plates. Roughly 10 colonies out of 1 million is red instead of white. What is the accuracy of segregation of his plasmid? EXPLAIN.

The 10 colonies are red because they have lost the Ade2 gene by losing the plasmid with the CEN sequence. More specifically out of the $10^6$ cells plated (which each gave rise to a colony) 10 had lost the plasmid before plating presumably due to an error in segregation. Therefore the accuracy is $1/10^5$ on roughly $10^{-5}$ frequency of chromosome loss per cell division.

b. Occasionally he sees a "sectored" colony in which part of a white colony is red. Explain how these sectored colonies form.

During growth of the colony, one cell loses the plasmid. Therefore this cell and all the progeny from this cell become red producing a sector.

4. Below is the link for the UCSC genome browser, which allows easy comparison of different mammalian and vertebrate genomes:

http://genome.ucsc.edu/cgi-bin/hgGateway

The interface is relatively intuitive, and you should be able to learn to browse and look at sequences over a half hour or so of exploration.

a. Getting back to problem 2, look at the HoxD chromosome region. Can you get support for your answer to problem 2? What can you look at in this browser that might give you insights? Tell me what you find?
b. Look for gene deserts- regions that are devoid of genes. Without too much work, tell me what the largest desert is that you can find (give me the chromosome location so I can find the same region in the browser).

Are there highly conserved regions within this desert?

c. Look for gene rich regions, the opposite of gene deserts- tell me the most gene rich region that you can find.

d. Tell me something interesting- either an observation or an interesting question- that you discovered or came to mind while browsing!