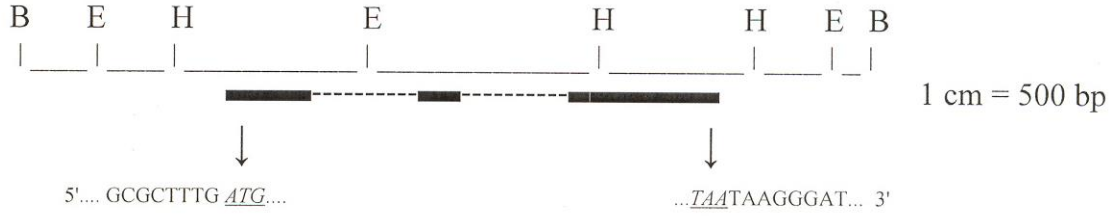
BPS3101 PROBLEM SET #2 FALL 2017

QUESTION 1

You have sequenced the region of an insect genome containing gene X, shown on the restriction map below (black bars = exons, dotted lines = introns). Arrows show the positions of short stretches of that sequence, with potential initiation and termination codons in underlined italics. Restriction sites are E: EcoRI, B: BamHI, H: HindIII.



a) Show what you would expect to see if **exon 3** were used as a hybridization probe in (i) northern analysis and (ii) Southern analysis - BamHI (lane 1), EcoRI (lane 2), HindIII (lane 3) EcoRI+HindIII (lane 4). Be sure to indicate size markers on your figures. Would your answer differ if **intron 2** were used as a probe? Give your rationale.

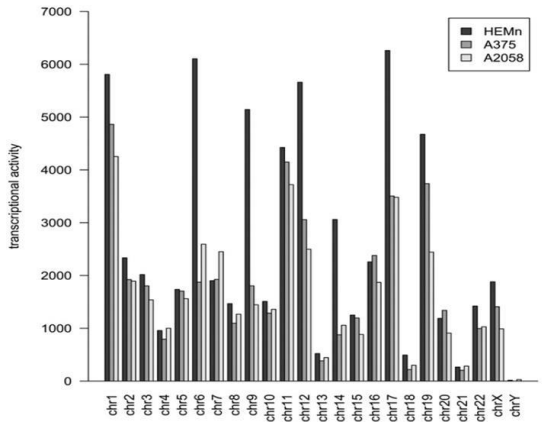
b) Design primers to generate an RT-PCR product for gene X and give its size(s). (Aside: see Practice set #1, question 9). If it were used as a probe in (i) a Southern hybridization experiment (as described above) or (ii) a northern experiment, what would you expect to see?

c) Suppose that the ATG which you thought was the initiation codon (italics in the figure above) is in fact an *internal* codon, and that gene X extends farther upstream (past what is shown in the restriction map).What strategy would you use to obtain a clone containing the rest of gene X? Use schematics in your answer.

d) Explain how you would use RACE to determine the position of the 3' end of mRNA X and show your strategy in schematics.

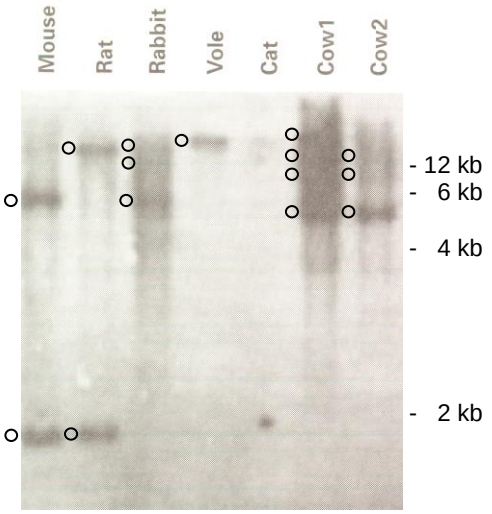
QUESTION 2

Give an interpretation of the data below (Zhao *Gene* 548:234, 2014). Transcriptional activity was determined from RNA-seq data for two cancer cell lines (A375 & A2058) and a normal cell line (HEMn).



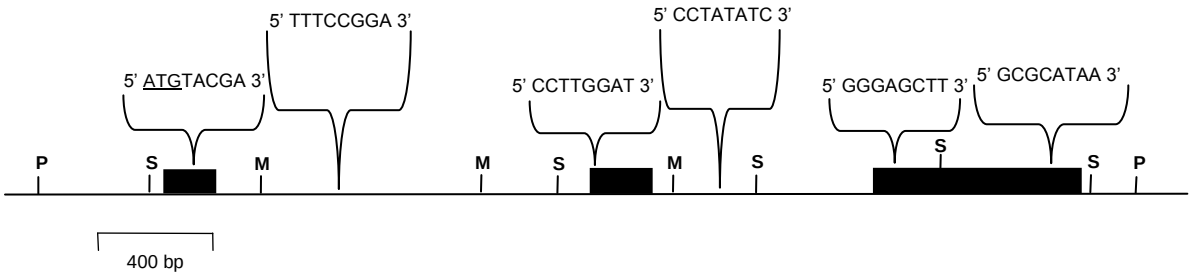
QUESTION 3

Suppose you have used gene G to probe a zoo blot of *Bam*HI-restricted DNAs from various animals and the data are shown below. [Tip: Open circles were put at the left of each band that was clearly visible in the original figure. Also, a vole is a rodent.] Give an interpretation of these results, along with your rationale, and propose further experiments to resolve any questions arising from these data.



QUESTION 4

Suppose that human gene Y shown below (exons = black bars; P, S, M = restriction sites) produces a protein that is 300 amino acids long in kidney cells, and one that is 230 amino acids long in the brain.



a) (3 marks) Propose a model for gene Y expression in kidney vs. brain, and give your rationale. [Tip: An important codon has been underlined].

b) Describe an RT-PCR based strategy to test your model. Show your specific oligomers (8'mers to save time) and templates in a schematic. Be sure to include 5' & 3'designations. Also show the expected results of your experiments and include a brief rationale.

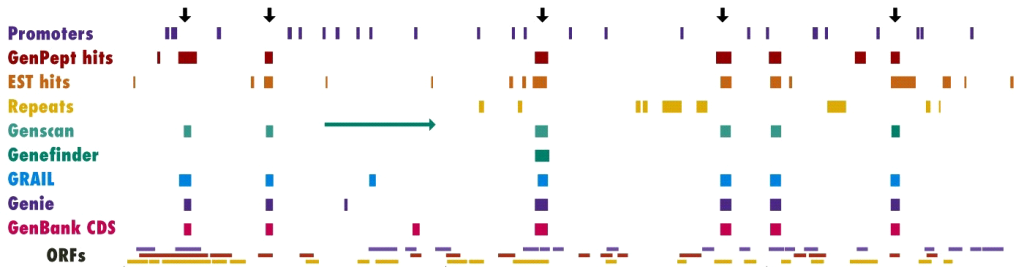
QUESTION 5
Suppose that the sequence below represents an internal section of fungal gene Y, and your bioinformatics analysis suggests that the region shown in lower case might be an intron.

5′..CCAGTGCGAAGTCGCC gtatcttactgcgggtatcgatgcgcgtactaacgtcttctttcgag CTAGCTCGTATCTGGC..3′

Describe **2** different **strategies** to test whether this lowercase sequence is indeed an intron, and show the **hypothetical results** of your experiments. For both strategies, include synthetic oligomers designed from the above sequence (6' mers to save time & show their positions). Be sure to include 5' and 3' designations in your schematics.

QUESTION 6
You suspect that a region of bacterial genome X which contains simple repetitive sequences is actually part of a protein-coding gene. Which **bioinformatics** and **experimental** strategies would you use to determine if this is indeed the case? Give your rationale.

QUESTION 7
What would you expect to see (and why) if 15 kb from the bacterial *Haemophilus influenza* genome had been used (instead of human DNA) in the analysis shown below (class notes & text Fig.5.10)?



QUESTION 8
a) How many genes (on average) would be expected in 400 kb of *C. elegans* DNA? Show your calculations and rationale. (I showed this question in class on October 5th).

b) Suppose that the genome of a newly-discovered bacterium (*Rideau canalium*) is five times larger than *E. coli*. How many genes would you expect it to contain? Show your calculations and rationale.