MCB 110
Spring 2017 Exam 1I KEY
FIVE PAGES

NAME:

SID Number:

Question	Maximum Points	Your Points
I	50	
II	50	
III	50	
IV	50	
	150	

Indicate which three questions you would like to be graded.

PLEASE WRITE your NAME or SID number on each page.
This exam must be written in PEN if you want the option of a regrade.
DO NOT USE WHITE-OUT: ask for a clean page of exam to start over if you need it.

Question I (50 points)

Some bacteria can undergo a "cellular differentiation" event called sporulation wherein cells switch from a normal "vegetative" growth state into a "dormant" spore upon nutrient starvation.

A. (5 points) During such a switch in induced cellular differentiation, what two classes of macromolecules would you monitor to determine global changes in gene expression as the bacteria exit the veg state and become spores?

Answer: all mRNAs (3 pts) and proteins (2 pts) found in the 2 cell-types

B. (15 points) Briefly describe an experiment (including an experimental flow chart) that would allow you to measure global changes in gene transcription during this process of going from vegetative growth to spore formation.

Answer: One good way to survey all changes in gene expression between veg and sporulating bacteria would be to use RNA-seq (8 pts) after isolating mRNA from cultures of veg and sporulating bacteria (3 pts). This procedure involves lysing the bacteria, extracting mRNA, fragmenting the RNA, carrying out reverse transcription to generate complementary double stranded DNA bearing linkers for sequencing and then performing high throughput sequencing to measure the frequency with which specific sequences representing mRNAs are obtained (ie, the more highly expressed a particular mRNA, the more frequently you will get a "sequence read" in your cDNA derived from mRNA) (4pts).
A possible alternative answer that I did not discuss in class: to extract all proteins from veg and sporulating cultures separately and subject them to mass spectroscopy to determine the identity and relative abundance of all proteins in these 2 cell-types. In this case, the experimental flow would include a protein isolation step followed by a protease fragmentation step and then mass spec and computational analysis of all the thousands of peptide fragments derived from gene products encoded in the genome of the bacteria that are expressed during veg versus sporulation growth.

C. (10 points) Given what you know about bacterial transcription systems in general, what changes in the bacteria's transcriptional machinery might you expect would allow its RNA Pol to switch from transcribing veg genes to sporulation genes?

Answer: The most parsimonious mechanism would be for veg and sporulating bacteria to use different sigma factors to program the same core RNA Pol to read and recognize different promoter elements that are responsible for turning on the genes for vegetative versus sporulation. Alternative answers may receive up to 6 points.

D. (20 points) Since this "switch" must respond to nutrient conditions, describe a "signaling" mechanism analogous to the PhoR/PhoB 2 component system in E.coli that could control the nutrient deprivation induction of spore formation. (HINT: think about how a 2 component signal response system might change the transcriptional apparatus to trigger sporulation.)

Answer: One mechanism would be for a membrane embedded receptor (NutS) that detects certain preferred essential nutrients (ie sugars, amino acids, phosphate etc) when these are abundant in the media and in the periplasmic space. When such nutrients become depleted, this receptor would no longer have a ligand bound and this causes a conformational change in the trans-membrane domain of NutS that activates a protein kinase transmitter domain capable of transferring the gamma phosphate of ATP from a His residue in NutS to a glutamic residue in a cytosolic response regulator protein SpoR. The key next step is that one of the targets of this activated response regulator is to turn on the transcription of

the sporulation specific sigma factor. Once this sporulation sigma factor is produced, it will then replace the veg sigma on many RNA pol core complexes and turn on a cascade of sporulation genes.
Up to 10 points for an accurate explanation of a PhoR/PhoB – like system
Up to 5 points for identifying how the bacterium would facilitate differential gene expression
5 points for identifying that this system relies on induction of the new sigma to alter its gene expression profile.

Question II (50 points)

A cancer biologist searching for changes in gene expression that promote certain cancers must first identify specific "aberrant gene expression signatures" in the tumor cells that correlates with cancer growth. If such changes in gene expression can be found, they must then demonstrate a direct mechanistic link to the tumor.

A. (5 points) What method we discussed in class would you use to compare the patterns of genes expressed as mRNA in the tumor versus genes expressed in normal cells at a genome-wide level?

Answer: Use the very sensitive RNA-seq method. After enriching for Poly-A-tailed RNA's using an oligo dT primer, the quantity of mRNA extracted from normal versus tumor cells and the number and identity of sequenced reads will tell you which if any, specific RNAs are highly over or under represented between the 2 types of cells. In many cases, one finds 1 or a few RNAs highly over-expressed in tumor cells.
5 points for RNA-Seq. 4 points for microarray. Microarrays can look at expression of many, many genes but cannot detect novel transcripts. 2 points for Northern Blot. Northern Blots cannot look at transcripts at a genome-wide level.

B. (15 points) Assume that one such highly expressed "cancer gene" is in a region of the human genome with no known gene expression "sign posts". After you have identified a gene that is abundant in tumors relative to normal cells, what techniques would you employ to map the promoter/enhancer elements responsible for regulating the expression of this putative "oncogene"? (HINT: You first have to generate a "gross" gene regulatory map of the control region surrounding the "cancer gene" – and then once that is obtained, you will need a different set of experiments to map more precisely the cis regulatory elements (both enhancers and promoters) that are responsible for activating the "cancer gene").

Answer: First, you will need to generate a DNAse hypersensitivity site map of the chromosome region surrounding the putative "cancer gene" potentially covering 10's-100 kb of DNA sequence. Once you have such a DNAse hypersensitive (DHS) region mapped, you would need to recombinantly clone these DHS regions into a vector linked to a reporter gene so that you can quickly assess if any of these regions contain cis-activating DNA sequences (ie enhancer/promoter elements). For example, if you identify three ~1000bp regions of DHS, you would clone each of these fragments into separate plasmids organized such that these test fragments are inserted "upstream" of a reporter gene such as luciferase. For mapping putative promoter elements, you can place the test fragment directly fused upstream of the reporter gene. For enhancers, you would need to also clone a basal or core promoter fragment directly upstream of the reporter and then insert your test fragment upstream of this promoter/reporter construct since enhancers cannot activate a gene without a proximal core promoter element. Once you have such plasmid reporters constructed, you will test their transcriptional activity by transfecting (introducing) them into an appropriate cell such as the tumor cell or a non-tumor cell line. To further carry out fine mapping of cis-elements, you would progressively delete or mutate these candidate cis-element fragments until you narrow down the active element to 10-20 bp. You should find that the cancer specific regulatory element would only activate genes in the cancer cells and not in normal cells.
3 points for use of DNAse hypersensitivity to identify potential promoters and enhancers.
3 points for cloning these potential promoters and enhancers into reporter plasmid (luciferase, GFP, etc.)
3 points for distinguishing between promoters and enhancers.
3 points for mentioning the need to transfect/introduce these plasmids into cells.
3 points for fine-mapping of cis-elements by progressively deleting or mutating these candidate cis-element fragments.
NOTE that, at this point, you have not purified your protein. In the absence of a purified protein, you would not be able to carry out DNA footprinting or EMSAs.

C. (15 points) Once you have "mapped" a specific 20bp region of the human genome responsible for regulating the expression of the cancer gene, what biochemical method would you use to purify the transcription factor responsible for recognizing and directly binding to this control element. Briefly describe both a purification procedure and how you would assay the activity of the TF during purification.

Answer: Since cis-regulatory DNA elements must be recognized and acted upon by TFs to turn genes on, you will need to identify and purify such factors and presumably, only the cancer cells will have such factors and not the normal non-tumor cell. Therefore, the best way to identify such activities is to make nuclear extracts from both the cancer cells and normal cells, and then test to see if one extract contains a protein that can bind specifically to your 20bp fragment of DNA that you mapped to contain a cis-regulatory element while the other cell extract does not display such a binding factor.
If you detect a cancer cell specific DNA binding factor using a technique such as DNAse footprinting or gel shift, the next step would be to attempt purifying this activity by sequence specific DNA affinity chromatography. The key to DNA affinity purification is to use multiple tandem copies of the putative 20bp cancer element attached to a column resin and then to apply extracts containing your binding activity in the presence of an excess of non-specific competitor DNA (ie random sheared DNA from salmon sperm or any other abundant cheap source of purified DNA). After each step of purification, you would assay for the presence of the correct DNA binding TF by using either footprinting or gel shift. Your control nuclear extract derived from normal cells should not show any such binding activity to your 20bp cancer element.
3 points for naming the biochemical method - DNA affinity purification.
6 points for describing how DNA affinity purification works
3 points for need to use multiple tandem repeats in order to purify the protein.
3 points for assaying activity of TF during purification with footprinting/EMSAs.

D. (15 points) After you have identified and purified a putative transcription factor that binds to your 20bp control element, how would you show that this protein actually recognizes and directly binds the 20bp element in the tumor cells?
Answer: So far, you have found a protein factor that can bind your 20bp target cancer element in test tube reactions. To confirm that this same TF actually binds to the "cancer" regulatory element in vivo, you could us the technique of ChIP- PCR or even better – ChIP-seq. To carry out either method, you first need to make specific antibodies against the "cancer activating" TF. An alternative is to make a "tagged- TF" but this is not as good as using Abs against the endogenous TF. Again, you would expect a ChIP- signal in tumor cells but not in the control normal cells.
10 points for ChIP-PCR or ChIP-Seq.
5 points for mentioning the need to make a specific antibody to carry out ChIP-PCR or ChIP-Seq.
3 to 6 for footprinting/EMSAs. Question specifically asked to show that this protein binds the 20 bp element *in vivo,* so any *in vitro* techniques (footprinting/EMSAs) would not address the question.
Up to 10 points for an orthogonal method of testing activity in vivo.

Question III (50 points)

The **C**lustered **R**epeats **I**nterspaced with **P**alindromic **R**epeats (CRISPR) are loci found in the chromosomes of many eubacteria. Within these repeats are short sequences (20-50 nucleotides long) that have homology to viruses that may infect the cell. The CRISPR loci are transcribed and the RNA is processed to form guide RNA's that assemble with proteins that become site-specific nucleases. These guide RNAs will target the nuclease to the invading viral DNA genome to effectively defend the cell from death caused by the invaders replication path.

A. (8 points) The CRISPR/Cas9 enzyme has two nuclease active sites. One, called "HNH", that cuts the target DNA strand paired with the guide RNA and the other, called "RuvC", that cuts the non-target DNA strand. Both leave 3'-OH and 5'-phosphate ends. What kind of damage results from these nuclease cuts?
Answer: A Double strand break (8 pts)

B. (8 points) Name two DNA repair mechanism that could address the damage caused by CRISPR/Cas9 cleavage.
Answer: Homology-directed repair / homology-dependent repair (HDR), microhomology-mediated end joining (MMEJ), non homologous end-joining (NHEJ), single-strand annealing (SSA), homologous recombination, recombinational repair (4 pts each)

C. (20 points) Is the Cell's own DNA a target for the nuclease? Explain your answer.
Answer: No the chromosome of the cell will not be a substrate. This is because during the adaptive phase – when the cell or it's ancestor first encountered the viral genome and survived – enzymes cut out target sequences in the virus next to but not including so called PAM sequences (protospacer associated modules) also found in the viral genome and in the specific snip proximal to the cut. Then these snipped out short duplexes from the virus get integrated into the Cell's CRISPR loci. The PAM sequences (no need to have the nomenclature for full credit) are the first sites for binding by the mature nucleases and these protein:DNA interactions are distinct from the guide interactions.
No credit for "yes" answers.
Up to 8 points – target DNA comes from the virus etc.
Up to 10 points – explanation of PAM/guide RNA
2 points – Cas9 interacts with PAM (protein-DNA interaction)

D. (14 points) Breifly describe why the CRISPR/Cas9 technology has so quickly swept the molecular biology field. In particular, give an example of how this technology has been applied in either biomedicine or agriculture.
Answer: CRISPR/Cas9 genome editing is fast, easy to execute and universally applicable to just about cell-type prokaryotic or eukaryotic. It greatly simplifies targeted double strand breaks for genome manipulations and has revolutionized gene editing for biomedical applications such as correcting genetic defects such sickle cell anemia in patients blood or genetically editing plant genomes to alter and improve metabolic pathways for more nutritious food production or pest resistance without resorting to GMO since there is no "foreign DNA" being introduced into the natural organisms.
Up to 10 points – making the guide RNA is cheap/easy/fast etc. relative to proteins like ZFNs and TALENs.
Up to 4 points – how it is used in either biomedical/agriculture. "KO a disease causing gene" was not a sufficient answer for full credit.

Question IV (50 points)

Below is a diagram of a primary transcript produced in a eukaryotic cell. The following questions relate to the processing of this transcript.

(Drawing here)

A. (10 points) Name three modifications that happen to the initial transcript in humans.
Answer: 7-methyl-G cap, 2'O-methylation, and Poly A additions or Poly A tails
1 out of 3 modifications: +2 to 3 points
2 out of 3 modifications: +7-8 points
3 out of 3: all 10 points

B. (10 points) In the **first chemical** step of splicing, what is the nucleophile?
Answer: The 2' OH of the branch point adenosine. 10/10 if mentioned branchpoint A and 2' OH. If student said 3'OH of branchpoint A, -5 points.

C. (5 points) Name the specific protein/RNA complex that recognizes this first step of splicing?
Answer: This question can be interpreted 2 different ways since it could be asking about the formation of the spliceosome onto the transcript (for which U1/U2 binding would be accepted) or the actual catalytic step (for which U2/U6 would be correct). If student said only one U2 or U6 alone, +2 to 3points. If student said spliceosome, +1 point.

D. (5 points) What is the electrophile in the second transesterification reaction?
Answer: Phosphate at 3' splice site. Student needed to understand that the electrophile is the phosphate of the 3' splice site. – 3 points if student guessed phosphate, but did not specify 3' splice site.

E. (20 points) In liver cells where the transcript depicted above is expressed, an RNA editing activity exists that is programmed to target a different mRNA. However, the editing activity erroneously finds the second exon shown above and creates, by conversion, a translational stop codon. This occurs in 10% of the mRNA's being produced. Assuming that the correct mRNA and the edited version attract pioneer ribosomes with equal efficiencies and ribosome levels are limiting. What would be the ratio of correct to short protein you predict under these conditions? Why?
Answer: Close to 100% of the protein product would be wild-type. Because of Non-sense mediated decay, the mRNA with the aberrant premature stop codon would be rapidly eliminated from the mRNA pool so that the steady state levels of full length protein synthesis would not drop. The EJC is deposited on the sliced RNA and the pioneer ribosome confronting the EJC downstream in the stop conformation will attract the RNAse machinery to destroy the defective mRNA. If student did not understand that there would be <<10% of short peptide with PTC, -9 points. If student did not understand this would be because of NMD, -11 points. If student did not explain pathway PTC, EJC, UPF1 were worth 3 points each.