Fall 2005 Genomics Exam #3
Proteomics and Systems Biology

There is no time limit on this test, though I am hoping it will not take more than 6 hours. There are 4 pages for this exam, including this cover sheet. You are not allowed discuss the test with anyone until all exams are turned in at noon on Thursday December 15. **EXAMS ARE DUE AT NOON ON THURSDAY DECEMBER 15**. Submit your paper and electronic answers before the deadline. You may use a calculator, a computer, but only the web pages that appear in this exam. You are NOT allowed to explore the internet to take this exam. This is a new policy and is required if I am to shorten the length of the exams. You may take this exam in as many blocks of time as you need to. NOTE: I leave town on December 9 and return December 14 after a red-eye back from a meeting in SF,CA. If you have questions, I suggest you ask me before I leave. I will have *very limited* access to email so you can use gmail: <amalcolm.campbell@gmail.com>.

The **answers to the questions must be typed in a Word file and emailed to me as an attachment or handed to me on a thumb drive. I want paper and electronic submissions, please.** Be sure to backup your test answers just in case. You may need to capture screen images as a part of your answers which you may do without seeking permission since your test answers will not be in the public domain. Make sure the screen shots are big enough for me to read easily. Remember to explain your thoughts in your own words and use screen shots to support your answers. Screen shots without your words are worth very few points.

THIS IS A CLOSED BOOK EXAM TO HELP SHORTEN THE TEST.

-3 pts if you do not follow this direction.
Please do not write or type your name on any page other than this cover page.
Staple all your pages (INCLUDING THE TEST PAGES) together when finished with the exam.

Name (please print):

Write out the full pledge and sign:

How long did this exam take you to complete (excluding typing)?

20 Points

1) Open the file called "Exam1.pdf".

a. What was the ultimate goal of this research design?

The goal was to discover new protein interaction domains in many different proteins.

b. Explain what happens in each of the steps A – D.

A: find proteins that interact with a common protein (e.g. A, B, C, D, E, F interact with X).
B: Known domains for interacting are removed from the sequences of A – F.
C: Sequence analysis to find common motifs in proteins A – F after step B.
D: Proteins that contain the newly discovered motifs are collected, and the probability of these occurring by chance is calculated as a *p* value.

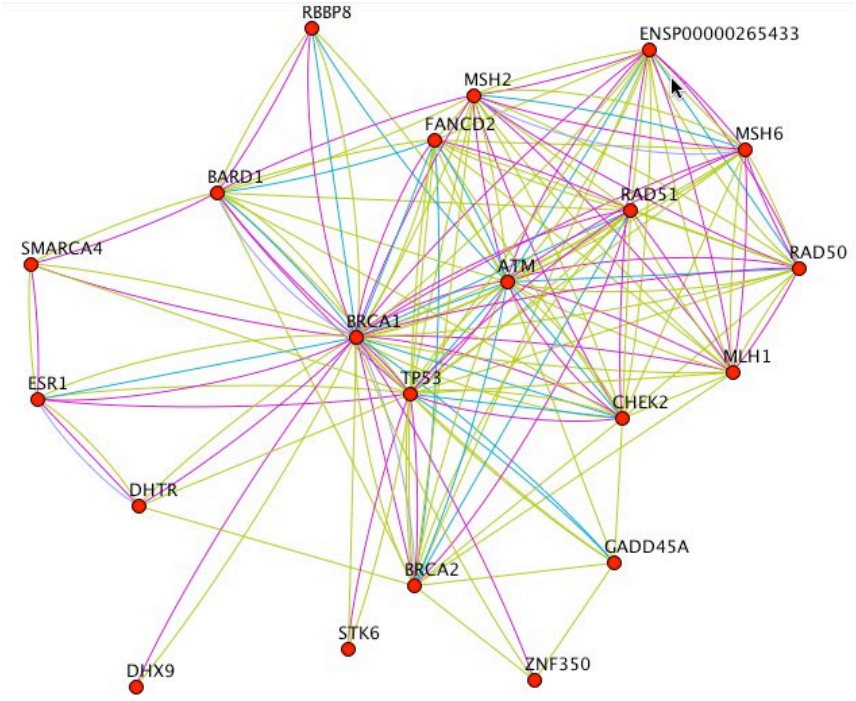
c. How could you experimentally test the results from step D?

You could put the motif in a yeast two hybrid as bait, or you could make the peptide and use affinity chromatography against the proteins listed in step D above.

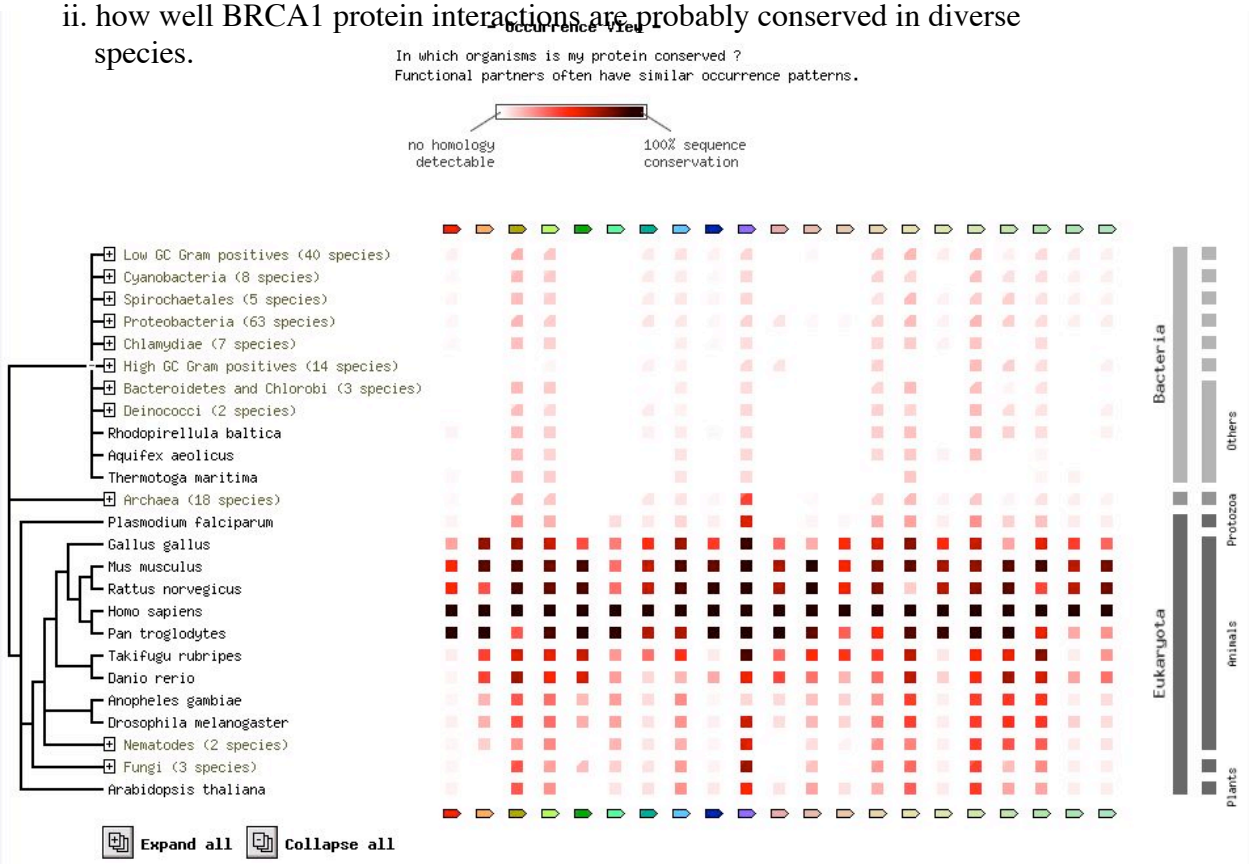
d. Go to this URL <http://string.embl.de/newstring_cgi/show_input_page.pl> and generate pictures that show:

i. which proteins interact with human BRCA1 similar to panel A in Exam1.pdf;

All of these nodes have direct interaction with BRCA1. In the static view, you can tell this because the nodes have color other than white. The different color edges indicate different sources of information. Need to set confidence at 0.7 or higher and expand view beyond 10.



ii. how well BRCA1 protein interactions are probably conserved in diverse species.



Your Input:

		Neighborhood	Gene Fusion	Cooccurrence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
BRCA1	Breast cancer type 1 susceptibility protein (1863 aa)									

Predicted Functional Associations:

ESR1	Estrogen receptor (ER) (Estradiol receptor) (ER-alpha) (595 aa)	X		X	X	●	●	●		0.999
MSH2	DNA mismatch repair protein Msh2 (934 aa)	X		X	X	●		●		0.999
MSH6	DNA mismatch repair protein MSH6 (MutS-alpha 160 kDa subunit) (G/T mis [...]	X		X	X	●		●		0.999
GADD45A	Growth arrest and DNA-damage-inducible protein GADD45 alpha (DNA- dama [...]	X		X	X		●	●		0.999
ZNF350	zinc finger protein 350; zinc-finger protein ZBRK1 (561 aa)	X		X	X	●		●		0.999
BARD1	BRCA1-associated RING domain protein 1 (BARD-1) (777 aa)	X		X	X	●	●	●	●	0.999
RAD50	RAD50 homolog isoform 1 (1315 aa)	X		X	X	●		●		0.999
BRCA2	Breast cancer type 2 susceptibility protein (3419 aa)	X		X	X	●		●	●	0.999
RAD51	DNA repair protein RAD51 homolog 1 (hRAD51) (HsRAD51) (339 aa)	X		X	X	●	●	●		0.999
TP53	Cellular tumor antigen p53 (Tumor suppressor p53) (Phosphoprotein p53) [...]	X		X	X	●		●		0.999
ATM	Serine-protein kinase ATM (EC 2.7.1.37) (Ataxia telangiectasia mutated [...]	X		X	X	●	●	●		0.999
FANCD2	Fanconi anemia, complementation group D2 (1471 aa)	X		X	X	●	●	●		0.999
CHEK2	Serine/threonine-protein kinase Chk2 (EC 2.7.1.37) (Cds1) (543 aa)	X		X	X	●	●	●		0.999
MLH1	DNA mismatch repair protein Mlh1 (MutL protein homolog 1) (756 aa)	X		X	X	●		●		0.998
RBBP8	Retinoblastoma-binding protein 8 (RBBP-8) (CtBP interacting protein) ([...]	X		X	X	●	●	●		0.998
STK6	Serine/threonine kinase 6 (EC 2.7.1.37) (Serine/threonine kinase 15) ([...]	X		X	X			●		0.998
DHX9	ATP-dependent RNA helicase A (Nuclear DNA helicase II) (NDH II) (DEAH- [...]	X		X	X	●		●		0.998
SMARCA4	Possible global transcription activator SNF2L4 (SNF2-beta) (BRG-1 prot [...]	X		X	X	●		●		0.997
ENSP00000265433	nibrin; p95 protein of the MRE11/RAD50 complex (754 aa)	X		X	X	●		●		0.997
DHTR	Androgen receptor (Dihydrotestosterone receptor) (920 aa)	X		X	X	●		●		0.997

All of the proteins are conserved in mammals and birds, though there is one exception (MLH1) in rats.
BRCA1 is poorly conserved in the other taxa, so it is not clear if any other interactions are conserved.

e. Which proteins from your figures above would you use for the process described in Exam1.pdf?
I would use all these proteins except BRCA1 to find the motifs.

20 Points
2) Open the file called "Exam2.pdf".
a. Summarize the behavior of circuits simulated in panels C – G.
C: One slow loop is slow to produce output but is resistant to noise. Output decays slowly.
D: Two slow loops are faster than one, is equally resistant to noise, decay is similar to C.
E: One fast loop is fast to produce output but is very sensitive to noise. Output decays rapidly.
F: Two fasts loops are faster to produce output, more sensitive to noise, output decays similar to E.
G: Two mixed loops are fast to produce output, (like E and F) resistant to noise (like C and D), output decays slowly, like C and D. It appears to be the best of both worlds.

b. Explain why C and D give different outputs.
The difference is in the rate of output. Two loops provides twice the ability to produce output and therefore its delay is reduced.

c. Explain why circuits E and F exhibit the outputs they do and contrast your answer to the one you provided for C and D.
E and F are almost the same, except F has a little faster output similar to D having faster output than C. F is a little more sensitive to noise because its induction is faster than E. This is most notable at the beginning and end.

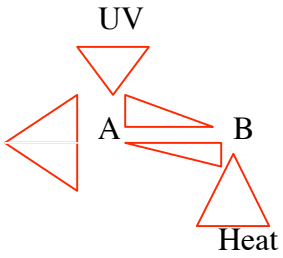
d. Which of these 5 circuits would you want controlling calcium signaling inside a cell, such as a cardiac cell? Explain your answer.
You would want a fast response time and one that is robust. Therefore, circuit G is the best cardiac system.

20 Points
3) a. In only one paragraph, explain how this multi-colored device works.
When UV is on, A is off while B and the reporter are on.
When heat is on, B and the reporter are off while A is on.
Without heat or UV, either A rules and reporter is off, or B rules and the reporter is one. This would be a stochastic choice.

b. Write a truth table for this device.

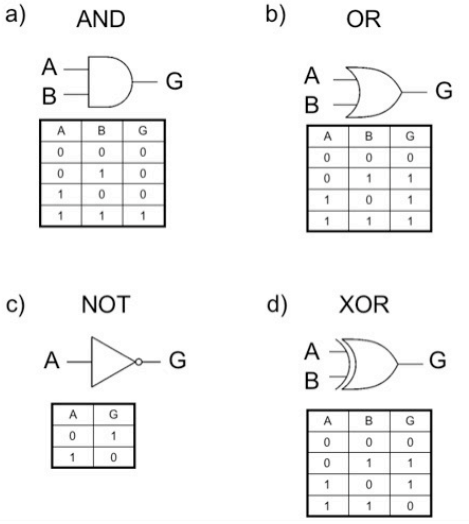
UV	A	B	Heat	Reporter
1	0	1	0	1
0	1	0	1	0
0	1	0	0	0
0	0	1	0	1
1	0	0	1	1

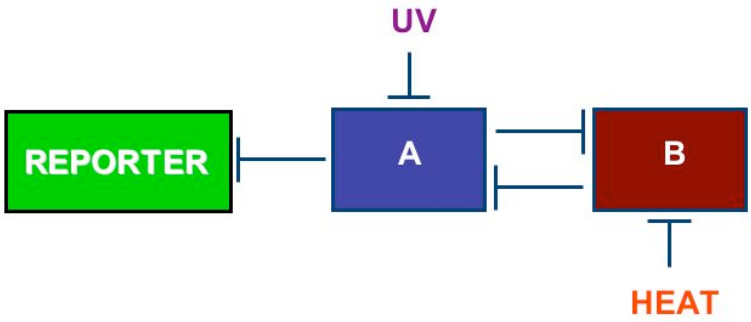
c. Redraw this device using electrical circuit symbols from the subset provided here.



The exact shape of the triangles does not matter. They are all NOT statements, or inverters.

d. Is this a toggle switch, a bistable toggle switch, or neither? Explain your answer.
This is a bistable toggle switch because it flips to one outcome and stays there until external stimulus flips it the other way.





20 Points

4) Open "Exam4.pdf".

a. Summarize briefly in words how a cell can produce different levels of noise for each of the panels A – D.

A: If the promoter is weak but translation is efficient, then the amount of protein produced will vary a lot and thus is noisy.
B: If the promoter toggles between active and inactive infrequently due to transcription factor complex stability and transcription is efficient, then protein production will be noisy.
C: Single locus copy of a gene is noisier than polyploidy or gene duplication.
D: Genes that have negative feedback on their own activity produce less noise than one which has no mechanism to detect its own level of production.

b. Make a generalization about what one process a cell should regulate if it wants to minimize noise. Support your generalization with examples in this figure.

It seems that controlling mRNA production is a critical factor in the amount of noise (i.e. protein production) in the system. With high rates of mRNA production (panels A and C), the noise is

reduced with slower step 2 (i.e. translation). Panels B and D indicate that a gene's ability to quickly alter its on status is the other trend to reduce noise. Indirectly, this also relates to the amount of mRNA (the first of two steps), though it appears to be more related to reducing stochastic bursts of excess mRNA rather than producing precise levels of mRNA.

c. Give one general rule or common example that we studied in class that would be subject to regulation as shown in panel C.

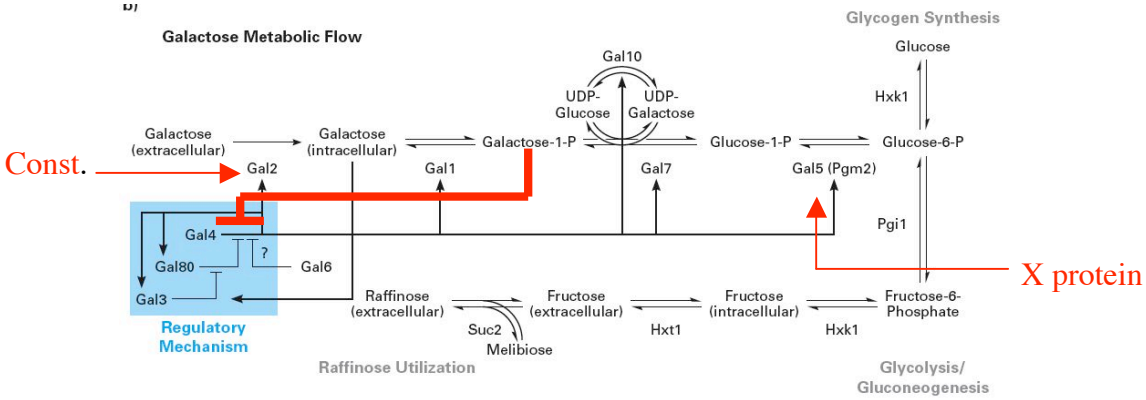
We talked a lot about redundancy due to isozymes/paralogs. Many genomes contain duplicate genes and often the selection pressure behind this duplication is unappreciated.

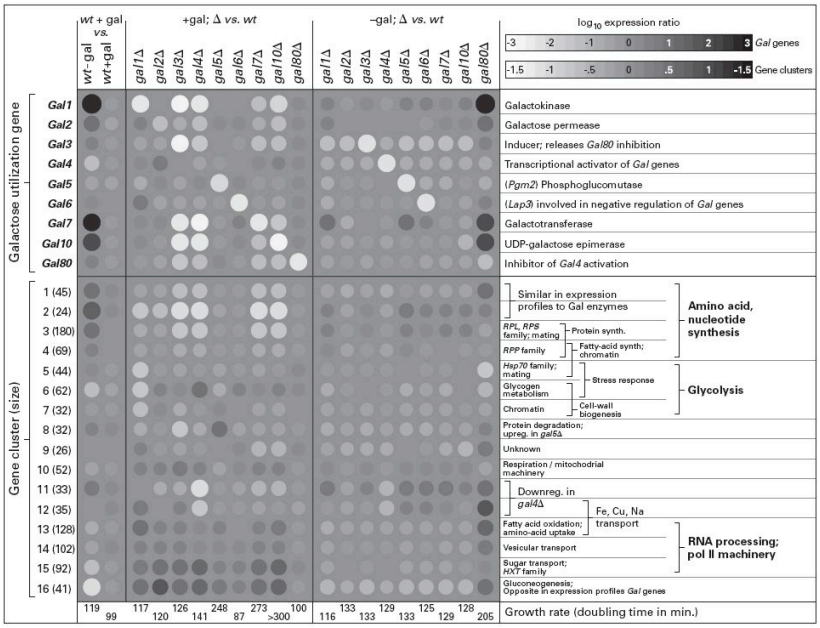
d. What evolutionary implications are inherent in the example you have chosen for question c above?

Processes that require precise levels of protein production should favor redundant coding capacity. If low-noise translation is advantageous for a particular process, then we might expect to see isozymes/paralogs for this particular gene.

20 Points

5) These figures should be very familiar.



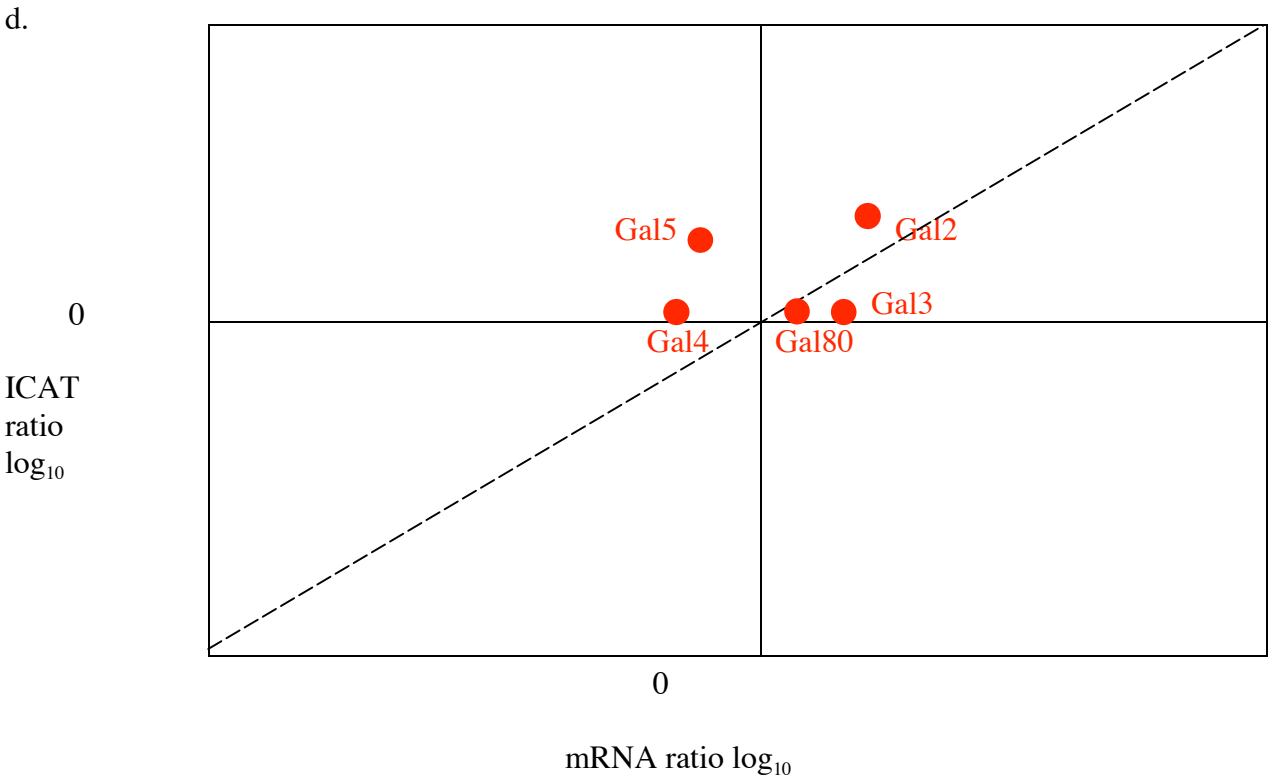


a. Name two emergent properties became evident during this research?
1) Gal-1-P has a negative feedback capacity similar to Gal80p.
2) Gal5 transcription is regulated by more than just Gal4p.
3) Gal2 transcription must be constitutive since some Gal2p must be present at all times.

b. Improve the blue regulatory area and redraw it in any format you prefer.
see above

c. Do you think the galactose pathway is a system with high or low intrinsic noise?
Support your claim with data.
Low intrinsic noise because there is negative feedback with Gal-1-P similar to panel D. Furthermore, it seems that Gal4p has a fast on and off rate which is similar to panel B with low intrinsic noise.

d.



Use the drawing tool in Word to plot where you think *Gal2, Gal4, Gal3, Gal80,* and *Gal5* would be in *wt* cells + gal/ *wt* cells - gal based on the data reproduced in question #5. Be sure to label each gene. If you have trouble doing it in Word, you may do it by hand on your paper copy.

All mRNA levels were taken from the first column in the figure above. For protein levels:
Gal2 and 5 have alternative sources of transcription and thus protein production.
You don't need much Gal4p, 80p, or 3p since there are limited binding sites and thus the protein levels would not change much +/- galactose.

Extra Credit: 3 points
Go to the URL below and take the post-semester survey. You are part of a national effort to measure how well students are learning DNA microarrays. You will be asked for other information as well so we can correlate your scores with those of other students around the country. The URL is: <http://www2.davidson.edu/survey/survey.asp?s=01181106106124014>.

Done – names and completion of survey verified by database search.