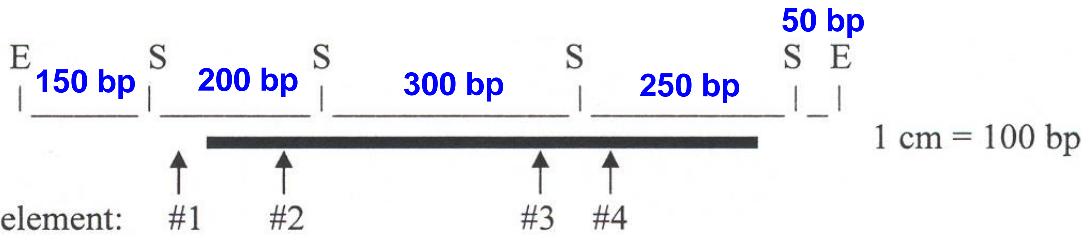
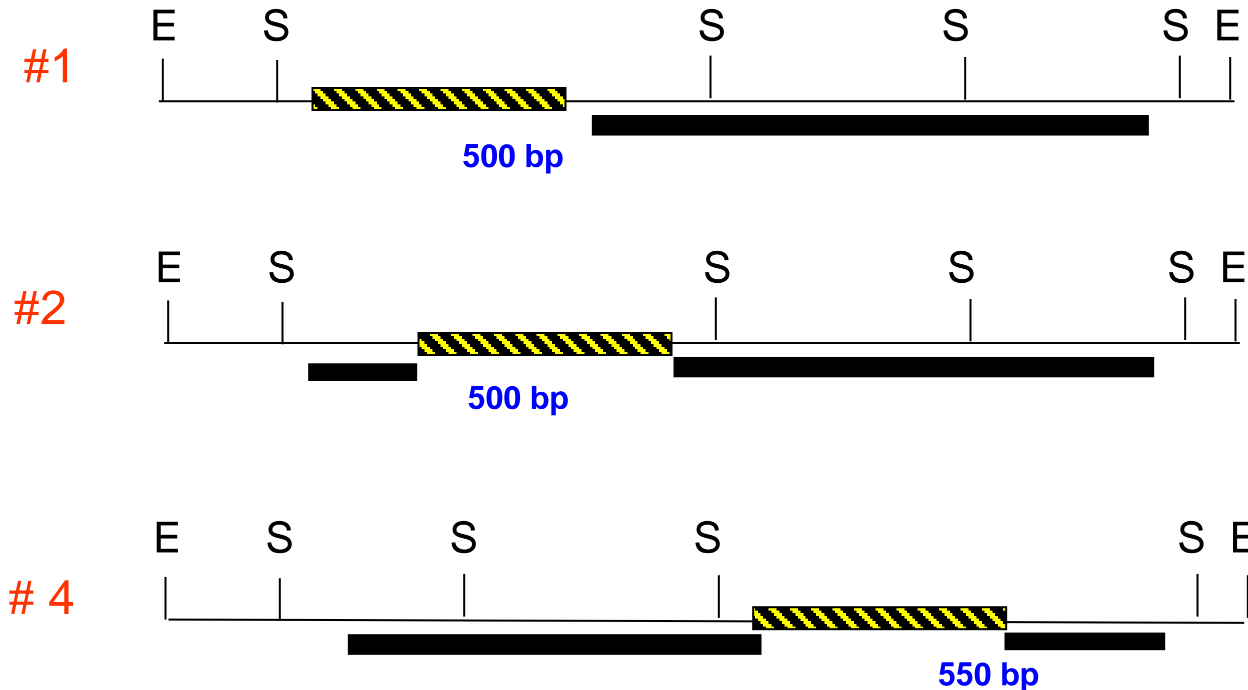
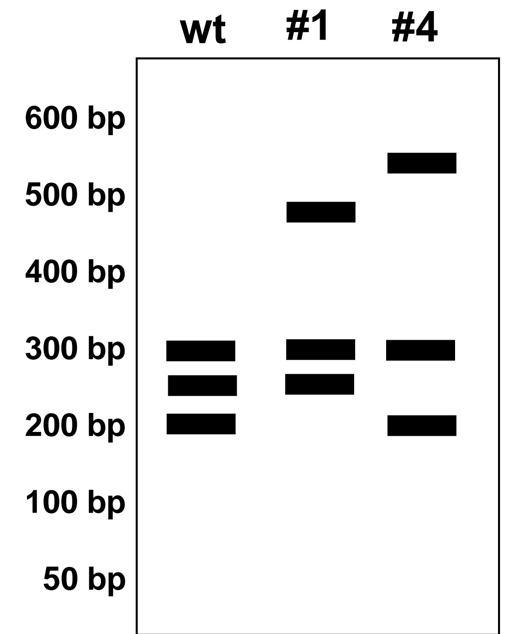
Problem set #3, Question 1



	Response to drug M
Wild type (no IS in gene A)	Sensitive
Strain with IS at position #1	Resistant
Strain with IS at position #2	Resistant
Strain with IS at position #3	Resistant
Strain with IS at position #4	Sensitive



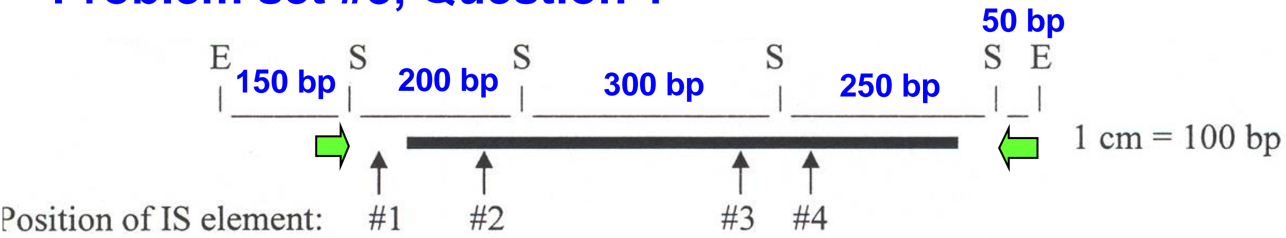


Model: *inactivation of gene A if there is an IS element in promoter region or within coding sequence*

except not if within C-terminal coding region

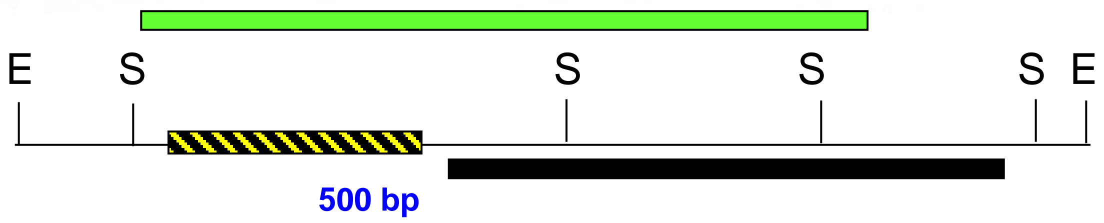
Transposon mutagenesis provides information about protein structure/function

Problem set #3, Question 1



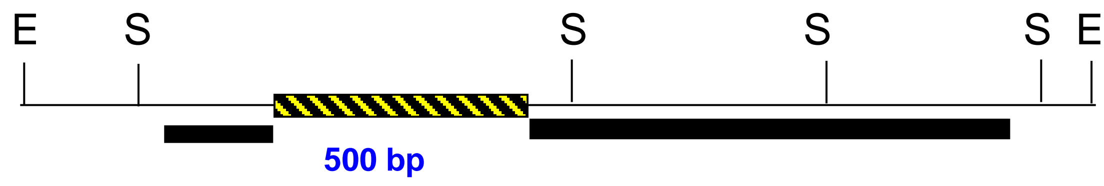
Response to drug M	
Wild type (no IS in gene A)	Sensitive
Strain with IS at position #1	Resistant
Strain with IS at position #2	Resistant
Strain with IS at position #3	Resistant
Strain with IS at position #4	Sensitive

#1

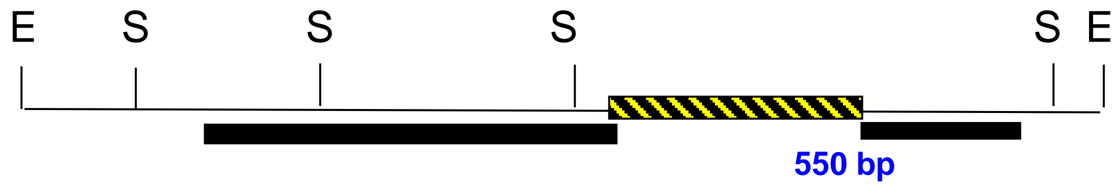


PCR method to screen new strains

#2



4



Are PCR products "wild-type" in length? *Gel electrophoresis (& size markers)*

For longer PCR products, analyze restriction profiles (using S enzyme) to localize IS element position *Gel electrophoresis (Southerns not needed)*

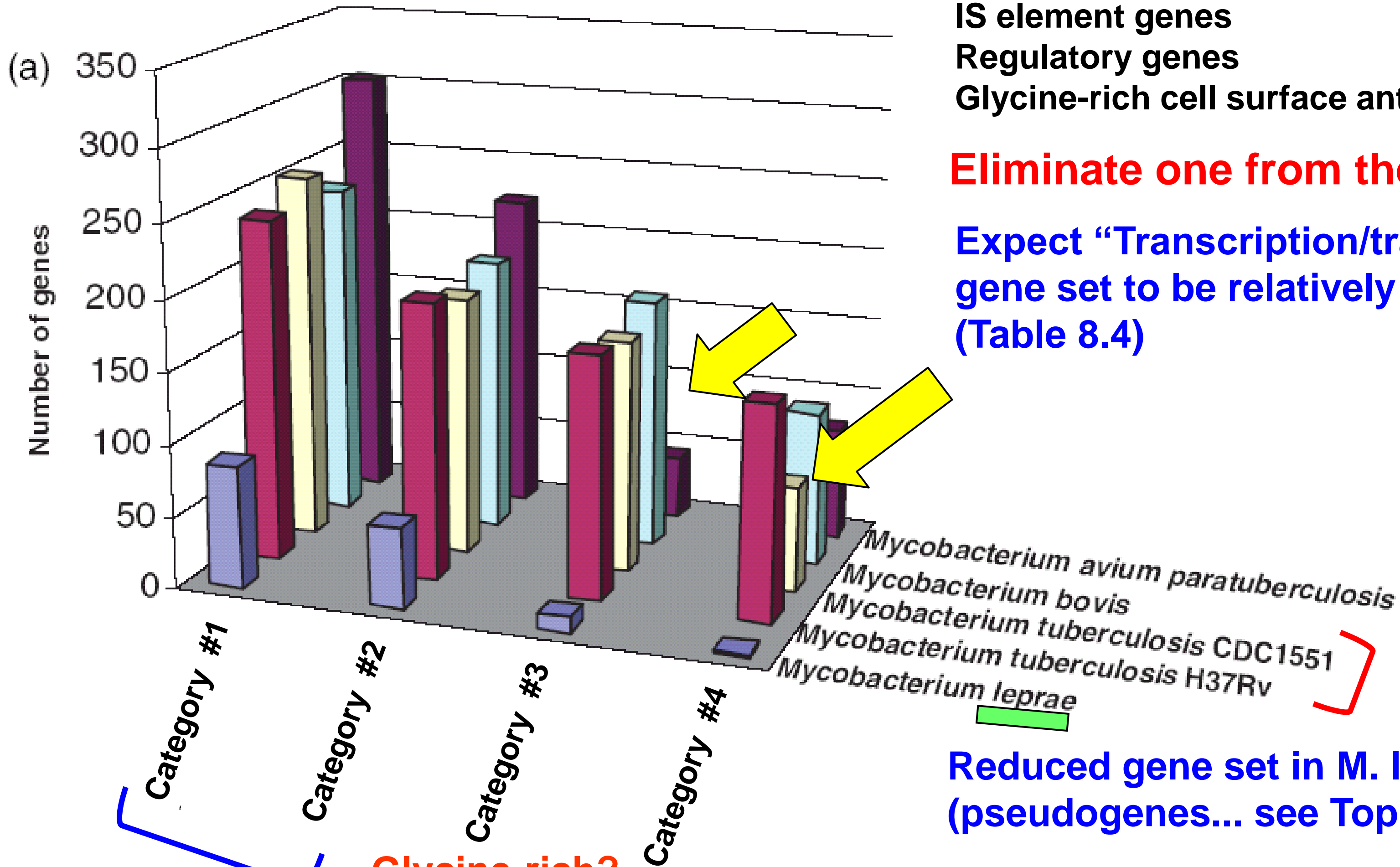
If IS element near 3' end of gene, maybe no negative effect on protein function

Problem set #3, Question 2

Lipid metabolism genes
Transcription/translation genes
IS element genes
Regulatory genes
Glycine-rich cell surface antigen genes

Eliminate one from the list?

Expect "Transcription/translation" gene set to be relatively constant (Table 8.4)



Reduced gene set in M. leprae (pseudogenes... see Topic 8)

Lipid metabolism or regulatory?

Glycine-rich? (defense)

IS elements?

Maybe "paraTB" pathogen has different defense mechanisms than others

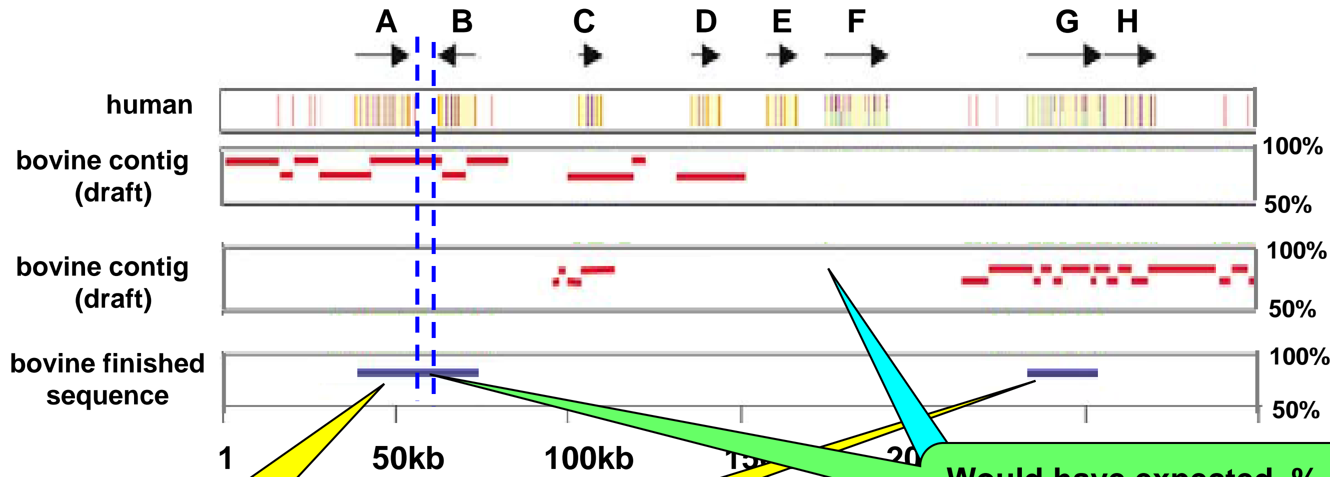
Aside: In paper, category #1: lipid metabolism, #2: regulatory genes, #3: glycine-rich, #4: IS elements

Adapted from Marri *FEMS Microbiol Rev* 30:906, 2006

Problem set #3, Question 3

MulitPipMaker plots

scoring sequence similarity between human & bovine DNA (300 kb region)



Homologues of human A, B & G genes in cow (~80% nt identity)

"Maybe these are housekeeping genes"

Would have expected % identities for spacer between genes A & B (as well as introns) to be lower than coding exons

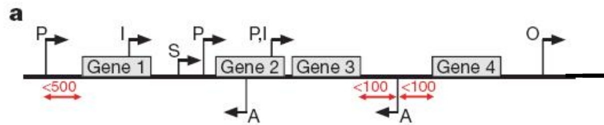
Maybe blue line is <u>average</u> value for whole region

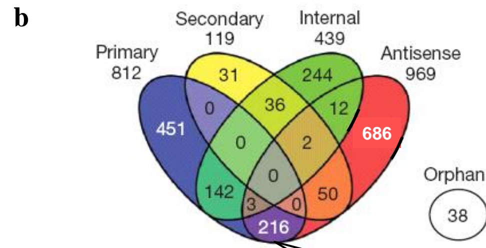
In "finished seq", gap is closed but genes E & F not there

Maybe they're somewhere else in cow genome…
…or absent

Rijnkels et al. *Genomics* 82:417, 2003

Problem set #4, Question 4





Promoters for non-coding RNA genes? or very short protein genes (so missed in annotation)? or pseudogenes?

Bidirectional promoters?

This is a discussion question about gene organization & expression in bacteria.

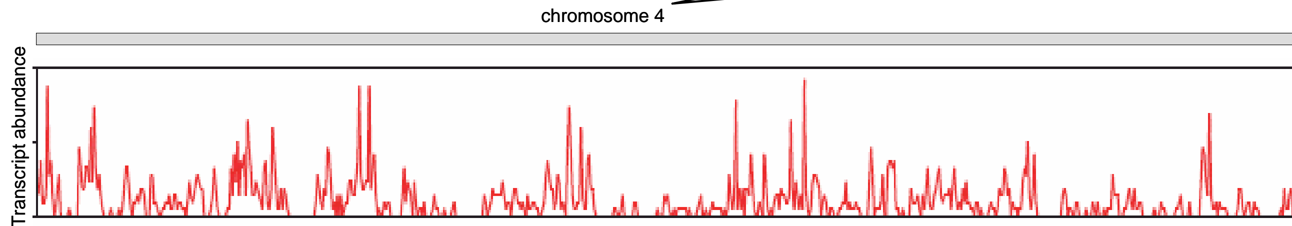
- transcription sites for operons can fall in strong (primary) & weak (secondary) categories (eg. depending on needs in different environments)

- internal sites can generate mRNAs for sub-sections of operons (maybe only certain proteins in that operon needed under some conditions, or different relative amounts needed)

- surprisingly high number of transcription initiation sites for antisense RNAs (maybe regulatory), maybe some reflect overlapping protein-coding genes (although not expect as many as in bacteriophage genomes)

Problem set #3, Question 5

~ 190 Mbp in length, so peaks reflect clusters of expressed genes (not exons vs. introns)



Uneven distribution of expressed genes along chr 4

Variation in abundance of transcripts among expressed genes

For housekeeping & liver-specific genes, can calculate approximate numbers (see Topic 6 slide)

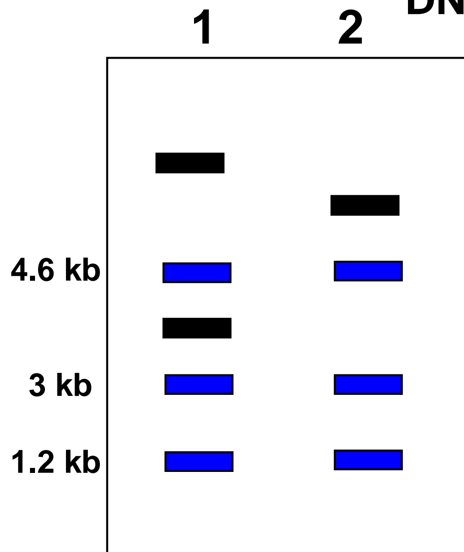
Not expect genes in heterochromatin, centromere regions etc.

For heat-stress conditions, expect most genes to be down-regulated, but small set of heat-shock responsive ones to be up-regulated

Problem set #3, Question 6

Alu-PCR fingerprinting (Fig. 4.15c in text)

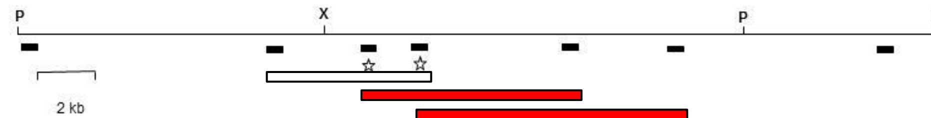
Use primers mapping to Alu repeat sequences, with template DNA : Pqr clone (lane 1) and Xyz clone (lane 2)



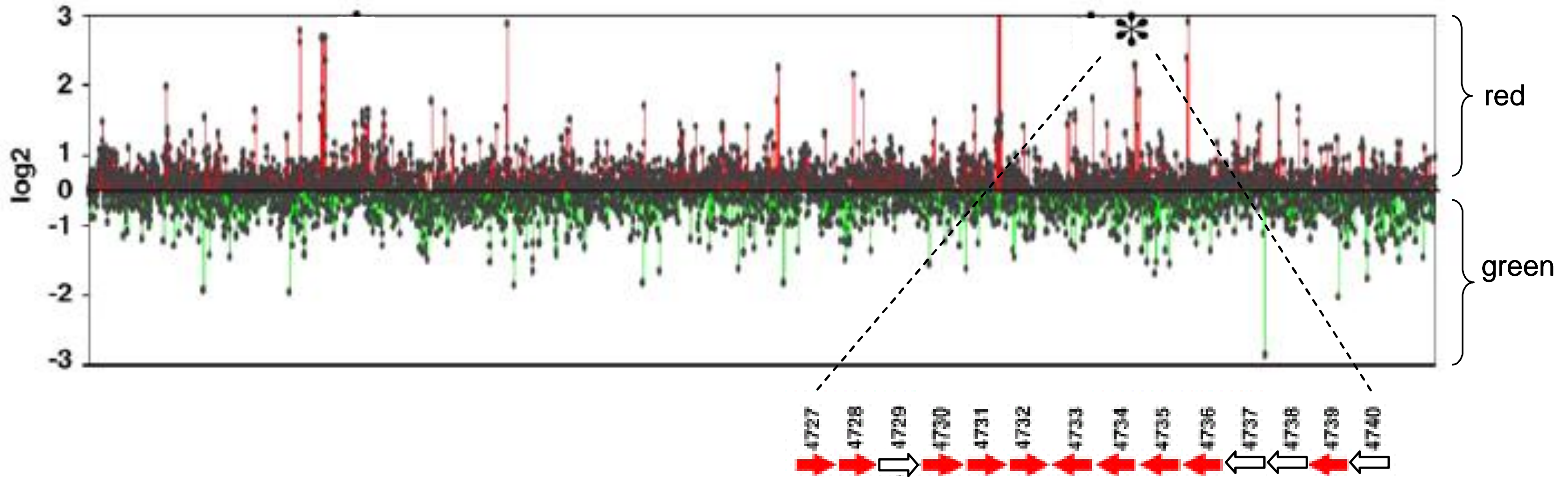
Lane 1 - clone from Pqr library
Lane 2 - clone from Xyz library

Ones in blue are shared PCR products , whereas black ones differ

Aside: You might also obtain longer PCR products (in which one Alu element has been "skipped over") assuming their distance apart is < 10 kb, and the two red ones would be informative as they are shared between the two clones.



Problem set #3, Question 7



X-axis represents whole M. smegmatis genome (ie. about 6000 genes)

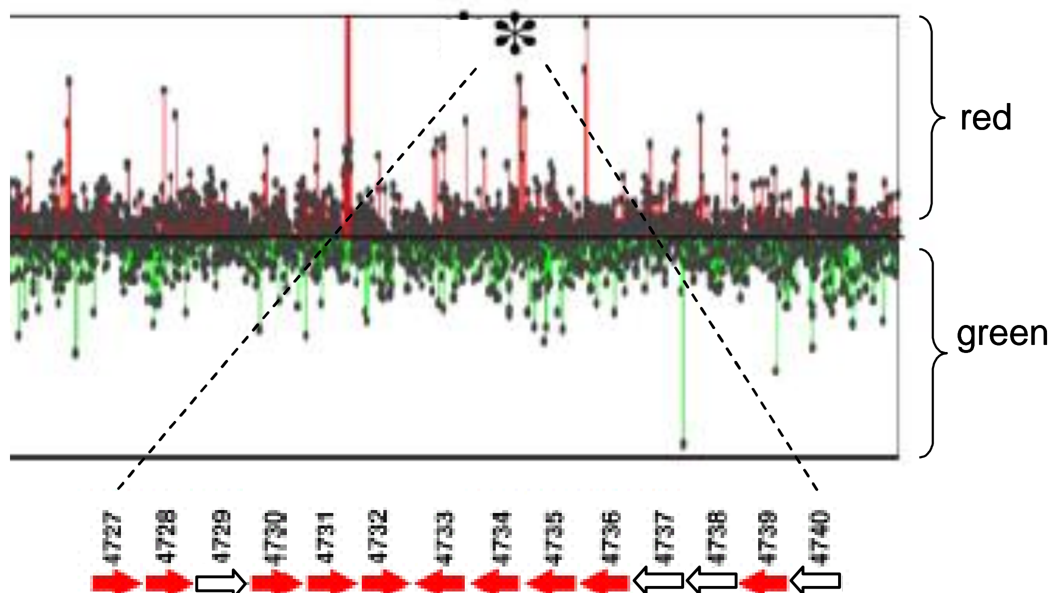
mRNA levels monitored in presence/absence drug A

Some clusters of highly-affected genes, but majority are unaffected (as expected) eg. housekeeping genes

Peaks of different heights reflect <u>relative</u> degree of change in expression

Colangeli *PLoS Pathogen* 3:780, 2007

Problem set #3, Question 7



All 14 genes in "blow-up "region are affected by drug A - 10 up-regulated and 4 down-regulated

some genes encoded on opposite strand (so must be multiple operons)

… also different drug A response for some adjacent genes

For up-regulated genes – maybe transporter, detoxifying enzymes…

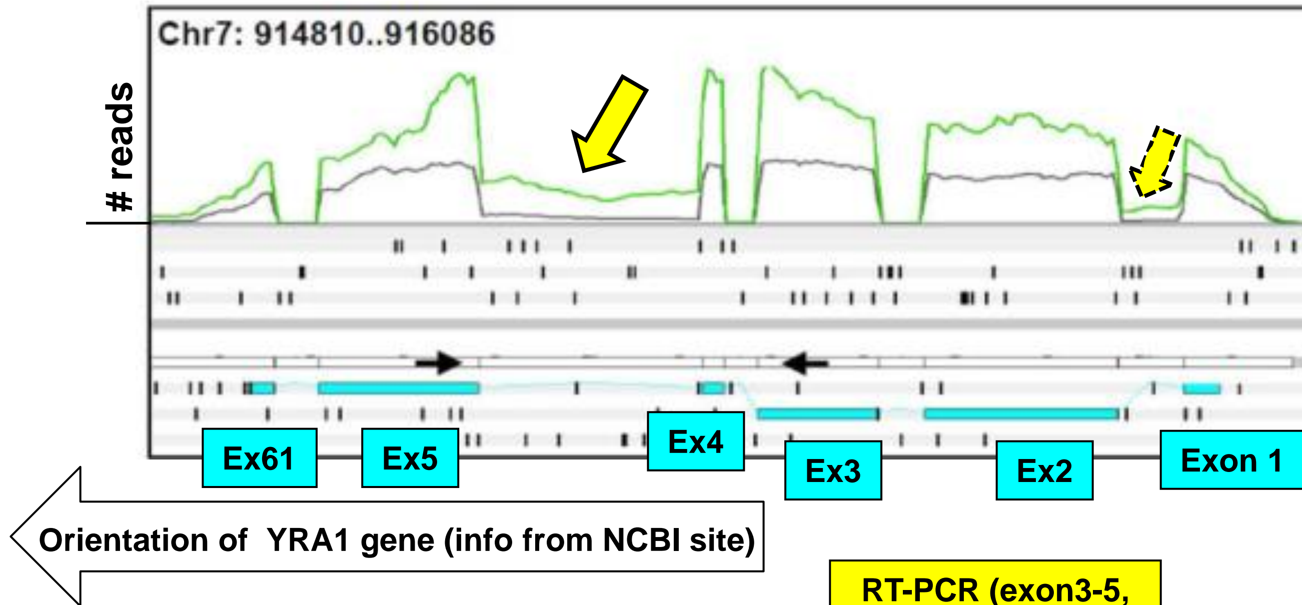
For down-regulated genes – maybe drug represses transcription (binds operator) or affects mRNA stability

Colangeli *PLoS Pathogen* 3:780, 2007

Problem set #3, Question 8

Higher mRNA levels in mutant (green)

C. neoformans YRA1 gene



RT-PCR (exon3-5, arrows)

Alternative splicing (intron retention)

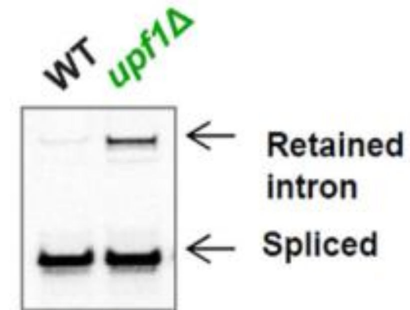


Figure from paper

Gonzalez-Hilarion *Sci Rep* 6:32252, 2016