



Leveraging Next Gen Sequencing to Improve Brewery Quality Control

Dan Driscoll – Avery Brewing Company
Phillip Richmond – University of Colorado

ASBC Annual Conference, Chicago.
June 5th, 2014.

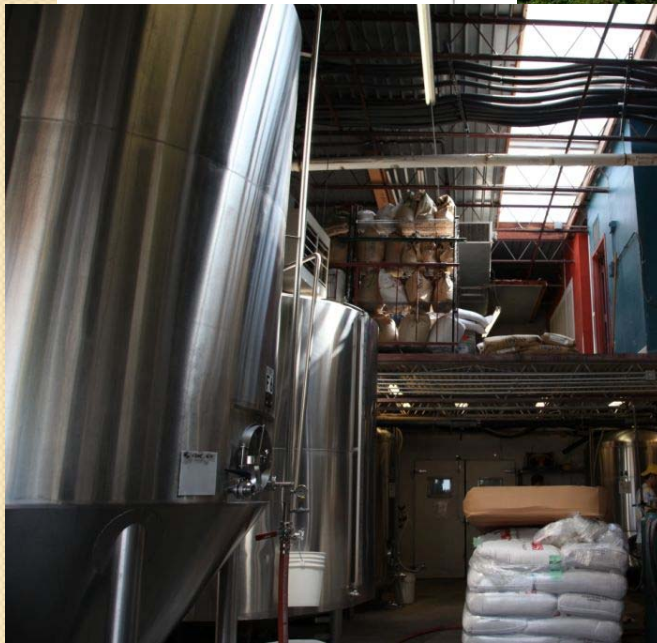
Presentation Outline

- Introduction to Avery
- Collaborative effort with CU
- Next Gen Sequencing
- Design of a brewery-specific diagnostic test
- Practical application
- Additional possibilities for NGS in brewing



Avery Brewing Company

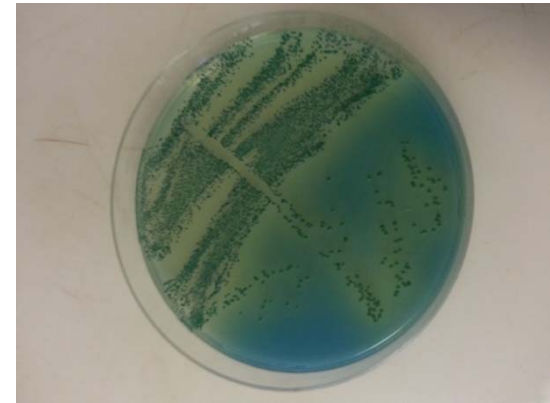
- History
- Production
- Setup
- Challenge



	2014	2013	2012	2011	2010	2009	2008	2007	2006	2005	2004	2003	2002	2001	2000						
6	556	472	598	859	1,13	1,54	2,00	2,76	3,40	5,12	5,16	6,92	10,6	12,4	14,8	4,15					
8	2,83	2,39	2,77	3,59	4,66	5,97	7,29	10,3	13,0	15,8	16,0	21,6	35,2	40,9	47,8	11,4					
	■ sales dollars per bbl produced																				
	\$202	\$202	\$202	\$203	\$200	\$196	\$197	\$216	\$239	\$243	\$259	\$275	\$268	\$260	\$323	\$321	\$320	\$302	\$304	\$309	\$364

QC Concerns at Avery

- Cross contamination of house yeasts
- Agar plating methods for detection now
 - 48 hrs, and subjective
- ASBC method for “fingerprinting” (Yeast -13)
 - No strain info
 - Are small amounts of “contaminant” DNA distinguishable?
- Generation of phenolic IPA
 - Destroying beer costs money and time
- What do we need?



Opportunity

- CU BioFrontiers Campus
- Next Generation Sequencing Lab
- Dowell Lab
 - Yeast lab that focuses on distinguishing individuals within a population, and correlating genotypes with phenotypes
- Collaboration and open exchange of ideas
- Publication potential?



Next Gen Sequencing

- Speed (DNA to sequence in 26 hours)
- Accuracy (100x coverage per genome)
- Data quality (~1% error rate per read)
- Cost (~\$500 per strain)
- Bioinformatics and annotation
 - Raw data to assembled genomes

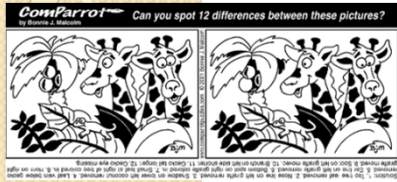


Strain coverage



- ~25 – 30 different brands per year, 6 yeast strains cover >98% of total 2013 production

Diagnostic Test Implementation

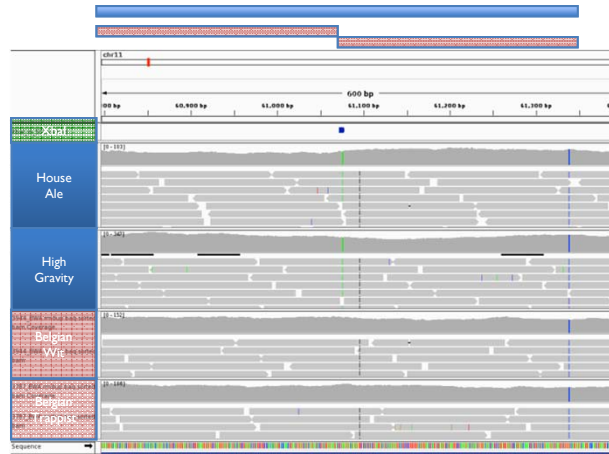


Find the strain-identifying SNPs

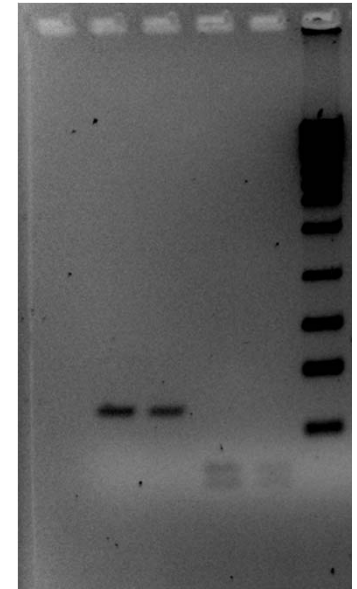
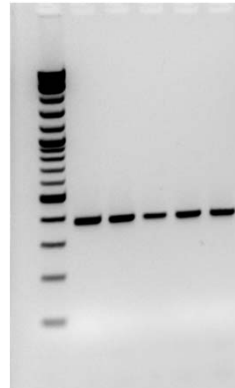


5'... TCTAGA... 3'
3'... AGATCT... 5'

Overlap with restriction digest sites

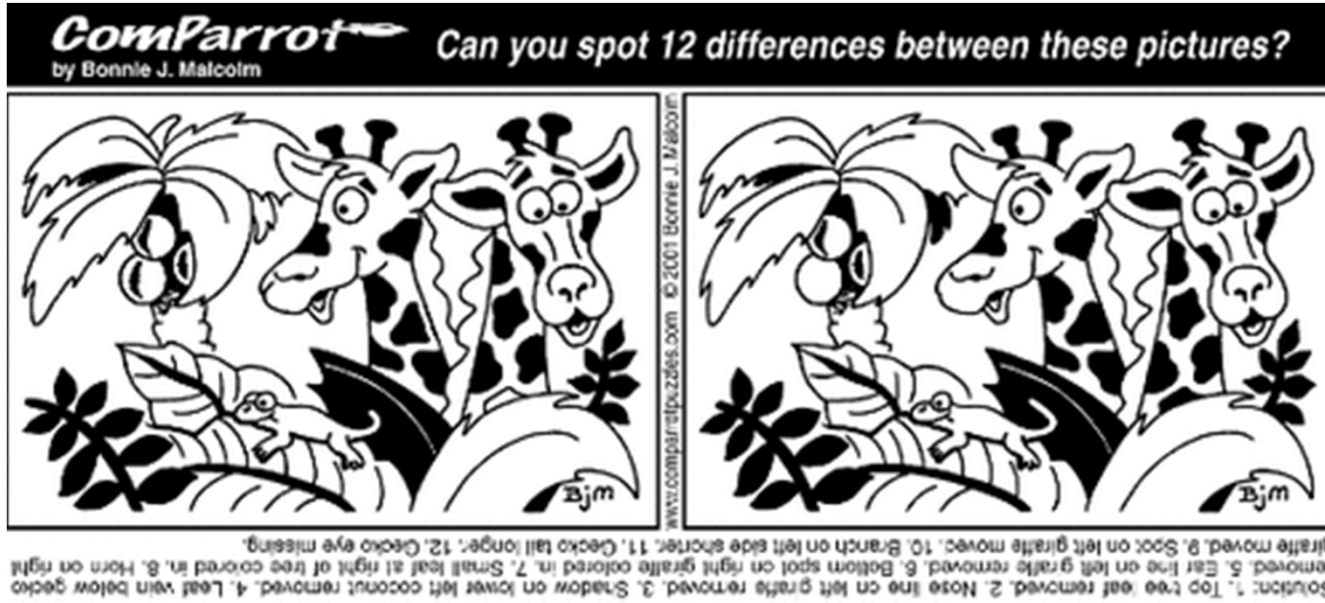


Design primers and PCR amplify a region encapsulating the SNP

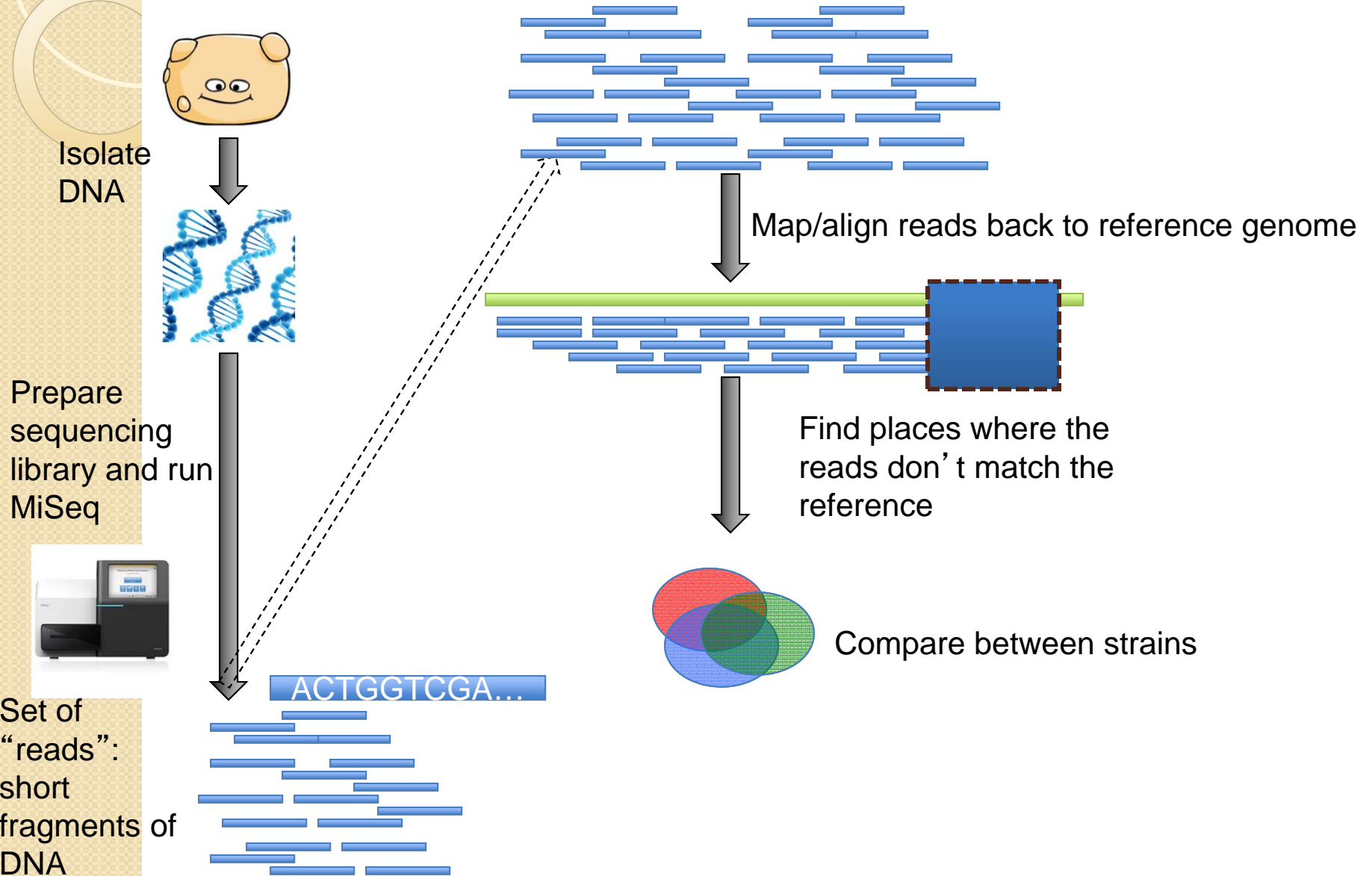


Digest PCR product and run on gel to find different banding patterns

Find all differences between the strains

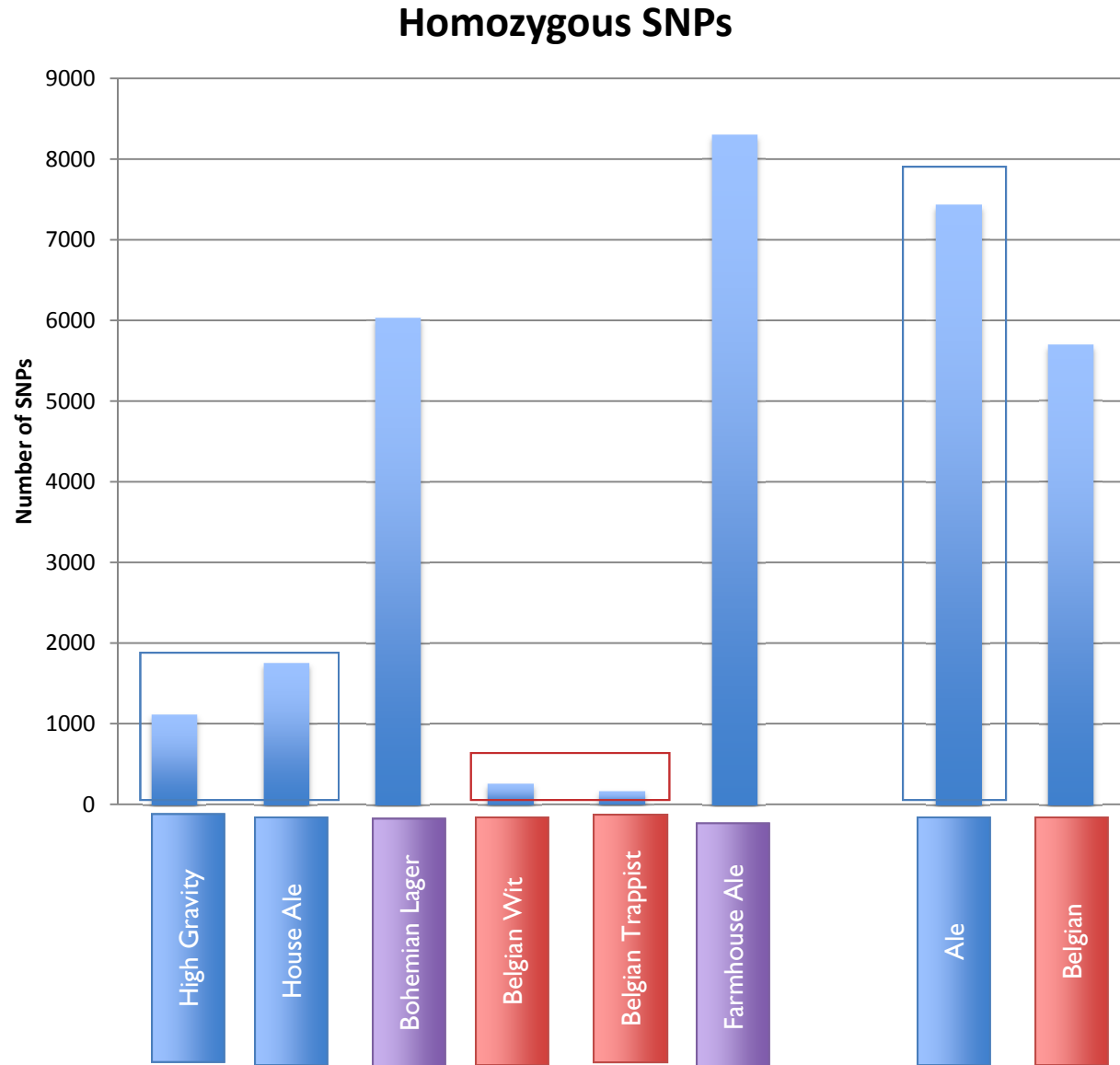


Finding Single Nucleotide Polymorphisms (SNPs)

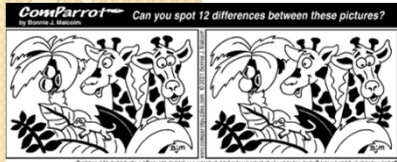


Homozygous SNP counts

- Need to use homozygous locations for the test to work
- House Ales and the two Belgian Strains are similar they have fewer unique SNPs, but they share a high number of SNPs between them that can be used to separate them from the American Ales



Diagnostic Test Implementation

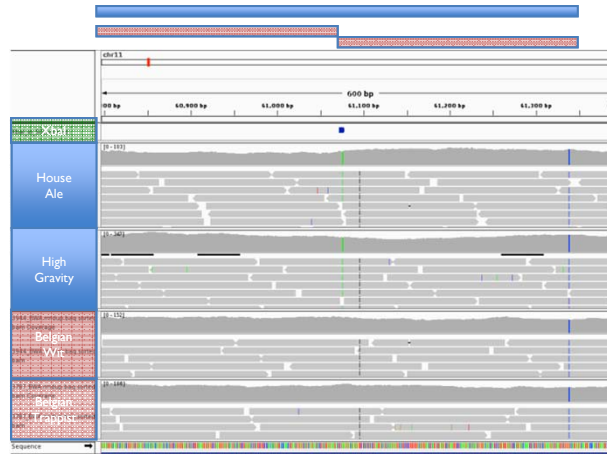


Find the strain-identifying SNPs

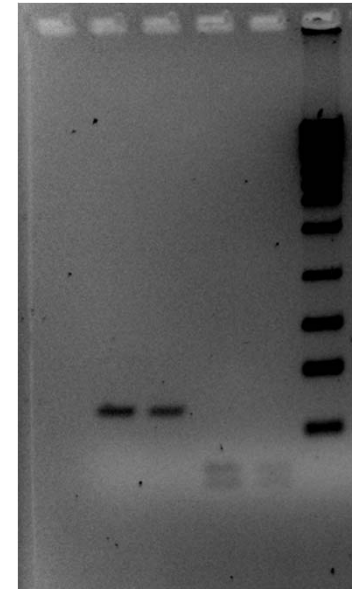
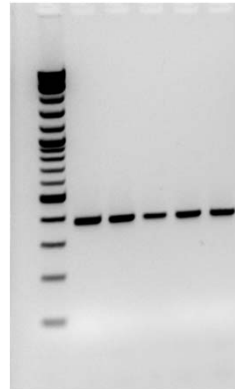


5'... TCTAGA... 3'
3'... AGATCT... 5'

Overlap with restriction digest sites



Design primers and PCR amplify a region encapsulating the SNP

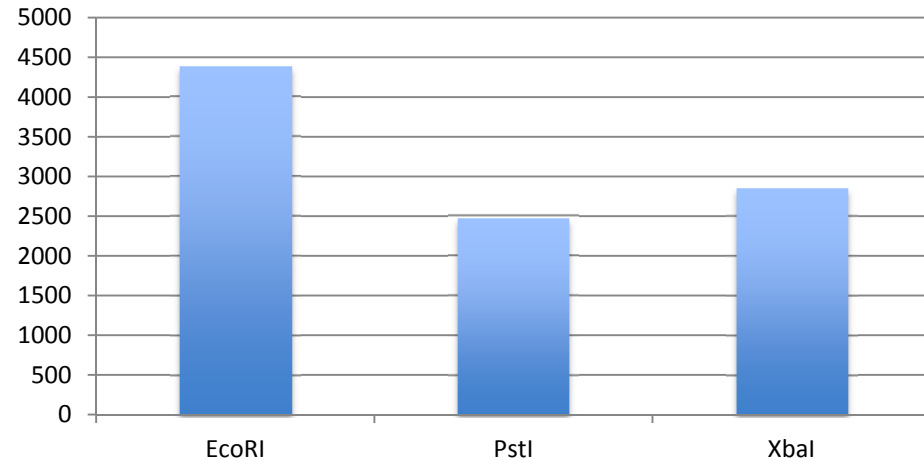


Digest PCR product and run on gel to find different banding patterns

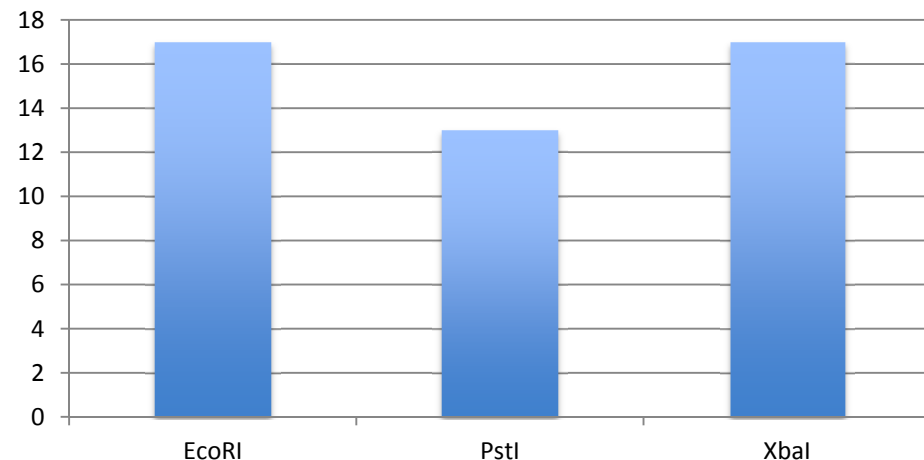
Restriction Digest Sites

- Took the recognition motif for EcoRI, PstI, and XbaI and annotated the genome
- Overlapped with set of SNPs occurring in Ale strains, but absent in Belgian Strains

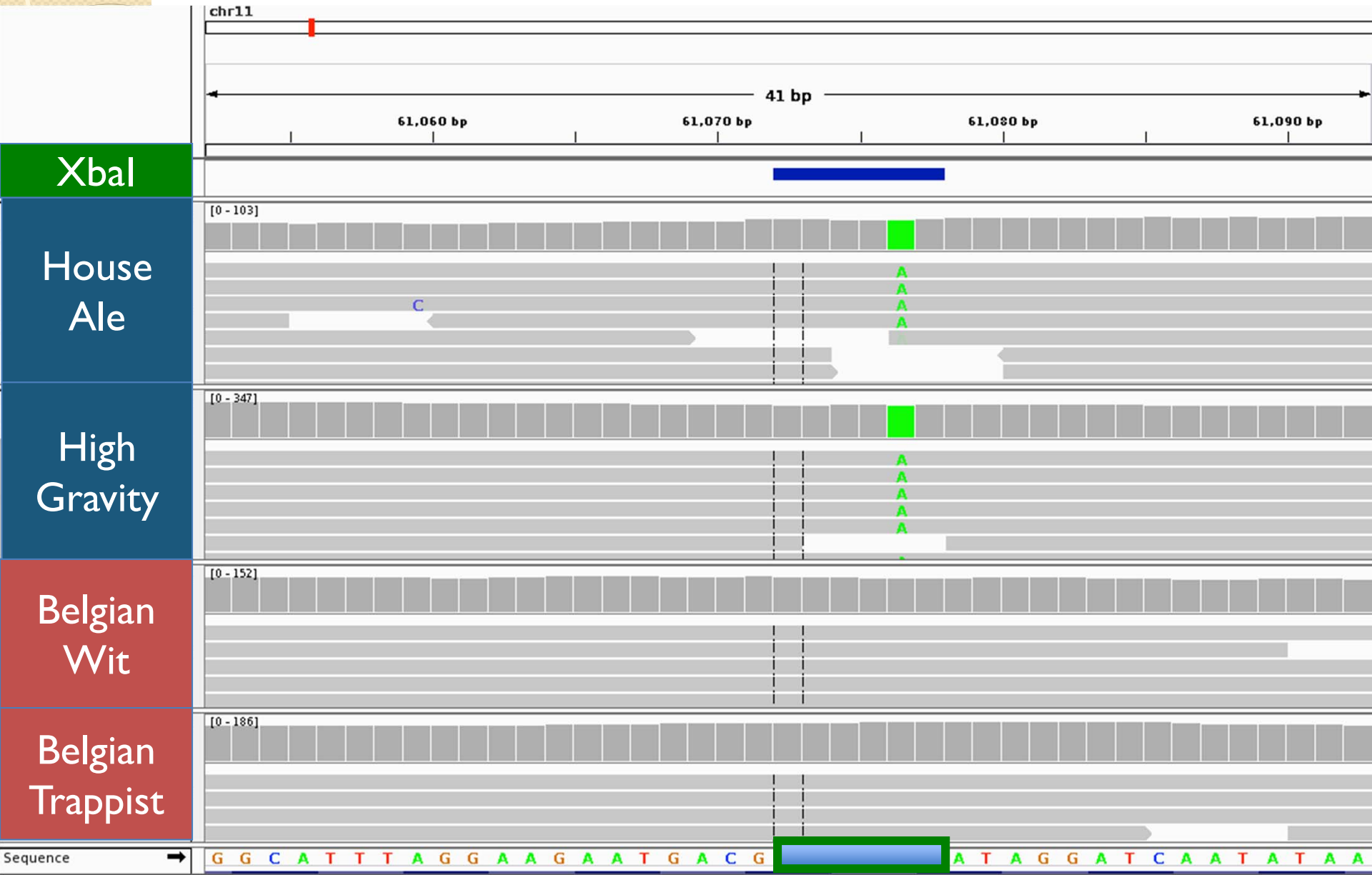
Restriction Digest Sites



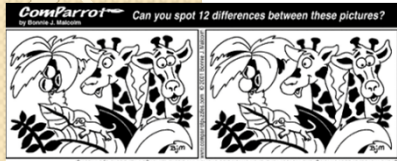
Ablated Digest Sites in Ale Strains



Ale Ablated Digest Site



Diagnostic Test Implementation

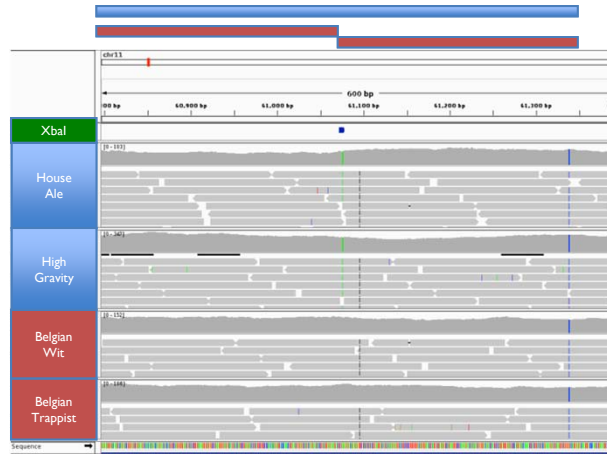


Find the strain-identifying SNPs

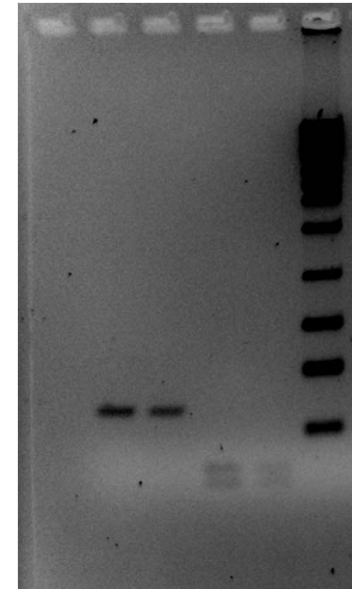
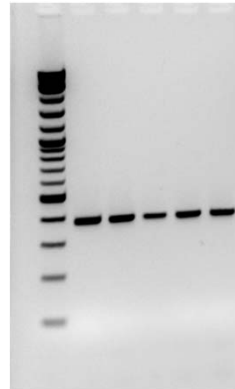


5'... TCTAGA... 3'
3'... AGATCT... 5'

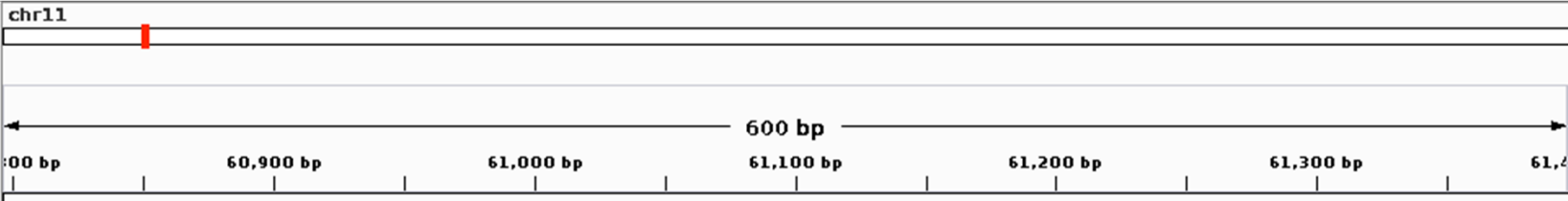
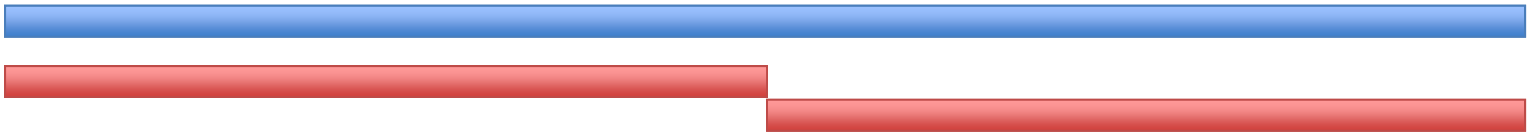
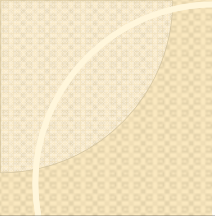
Overlap with restriction digest sites



Design primers and PCR amplify a region encapsulating the SNP



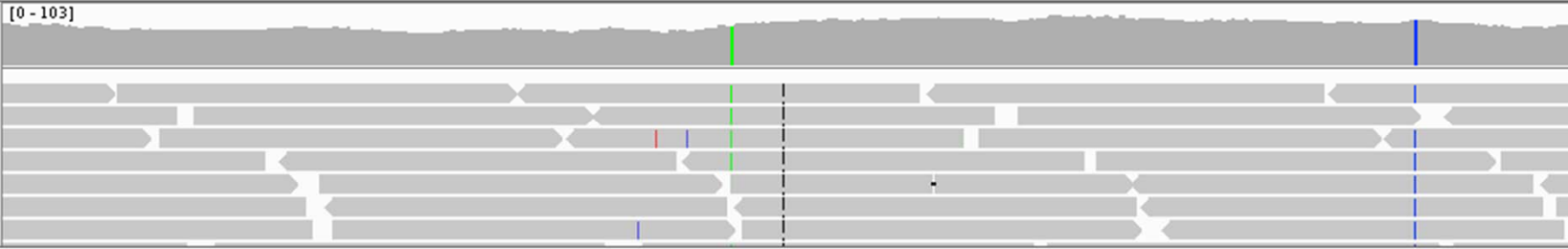
Digest PCR product and run on gel to find different banding patterns



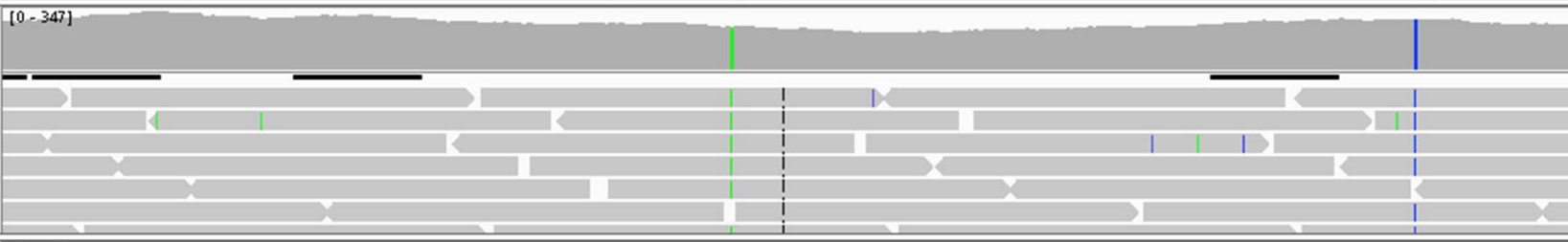
Xbal



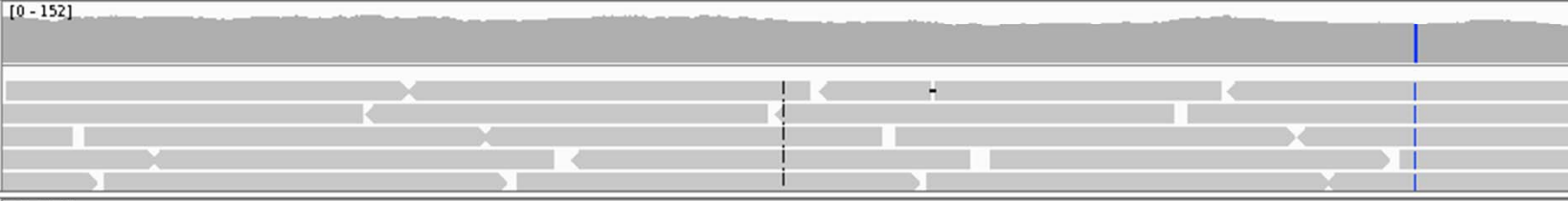
House Ale



High Gravity



Belgian Wit



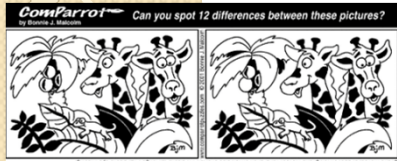
Belgian Trappist



Sequence →



Diagnostic Test Implementation

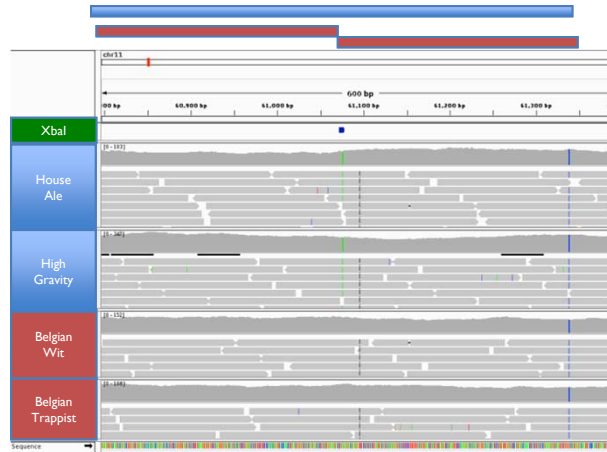


Find the strain-identifying SNPs

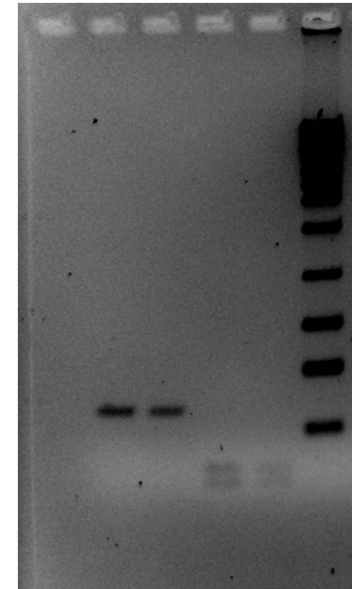
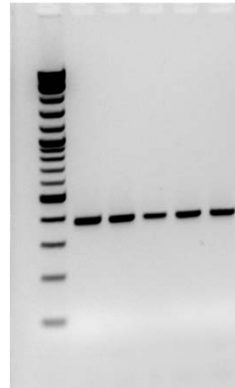


5'... TCTAGA... 3'
3'... AGATCT... 5'

Overlap with restriction digest sites



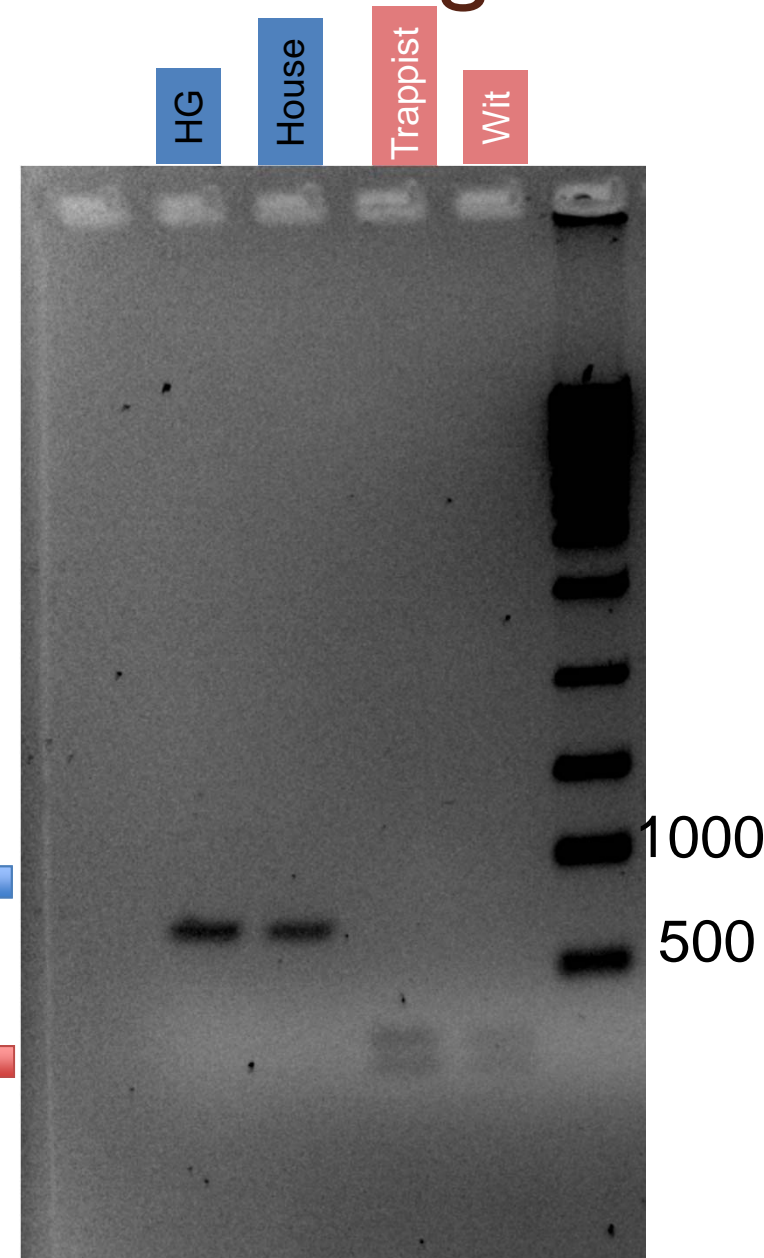
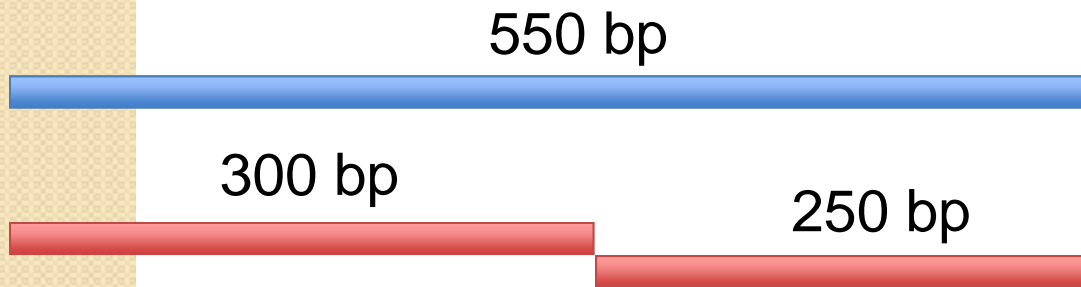
Design primers and PCR amplify a region encapsulating the SNP



Digest PCR product and run on gel to find different banding patterns

Diagnostic Test Identifies Ale vs. Belgian

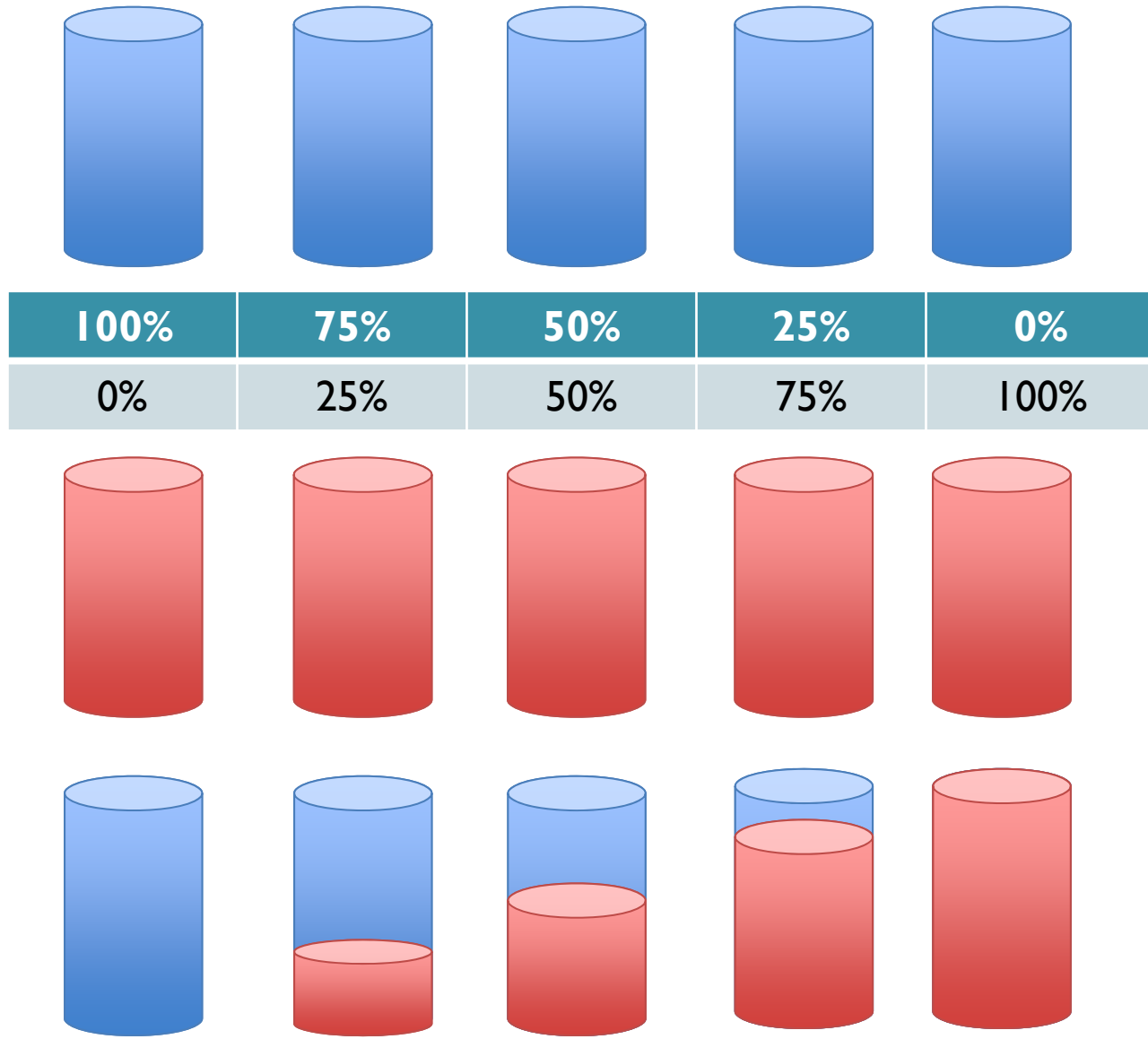
- Tested to make sure they all amplify a single band, then ran the digestion across all four strains.
- As expected, only Belgian PCR products get cut



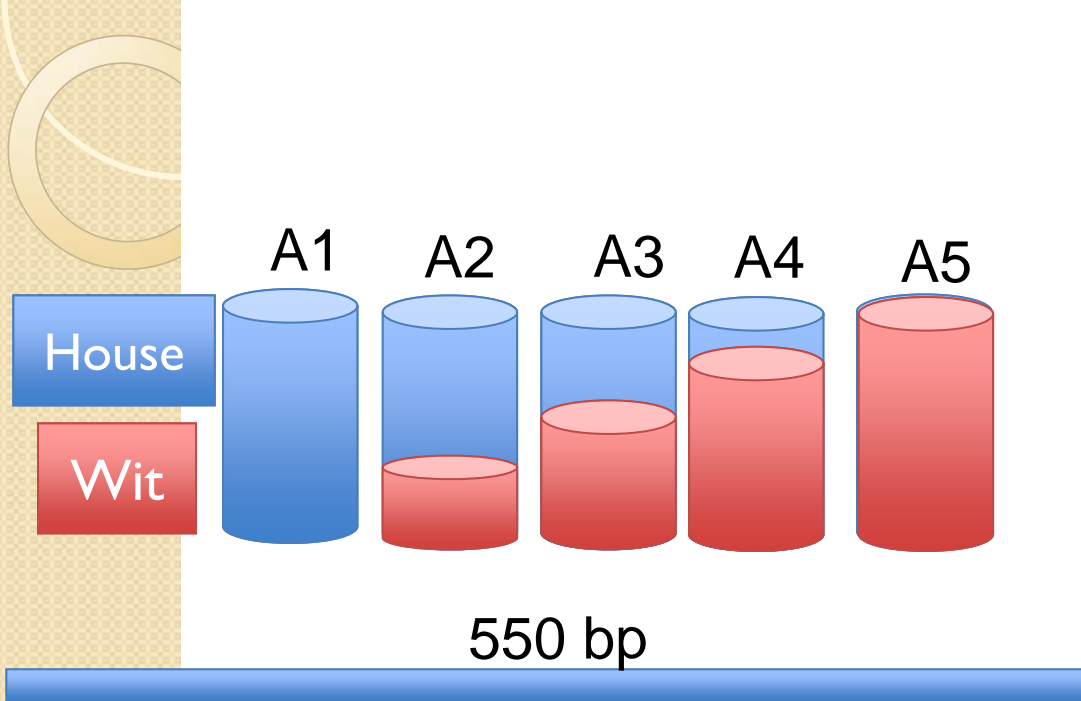
Mimicking Contamination Levels

From 1mL, 1.0 OD cultures

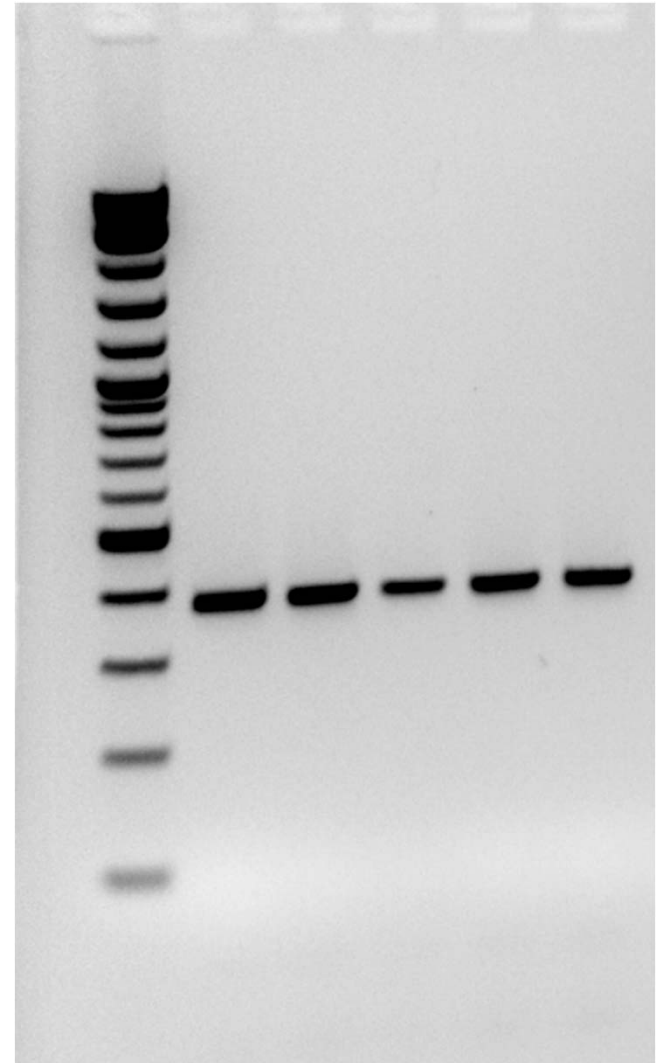
House
Wit



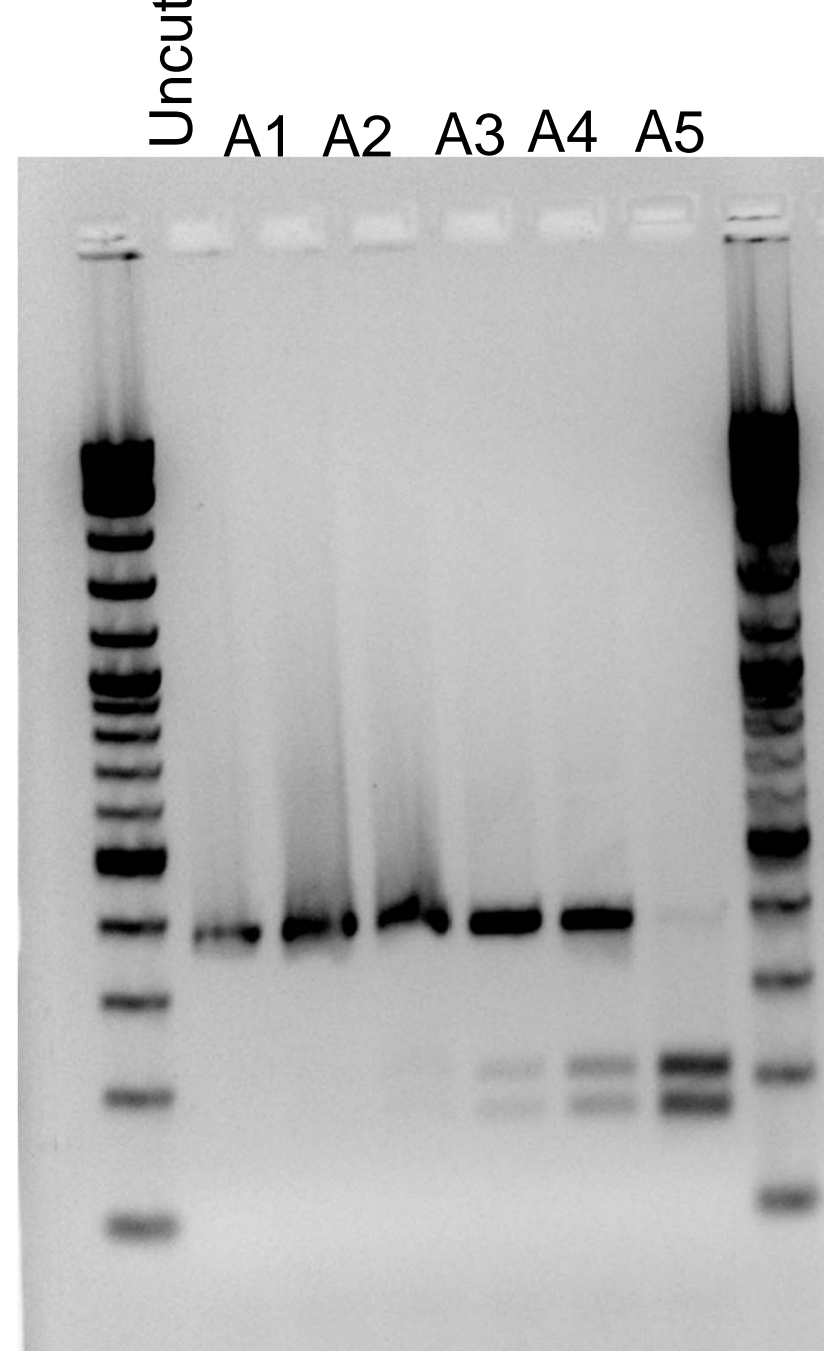
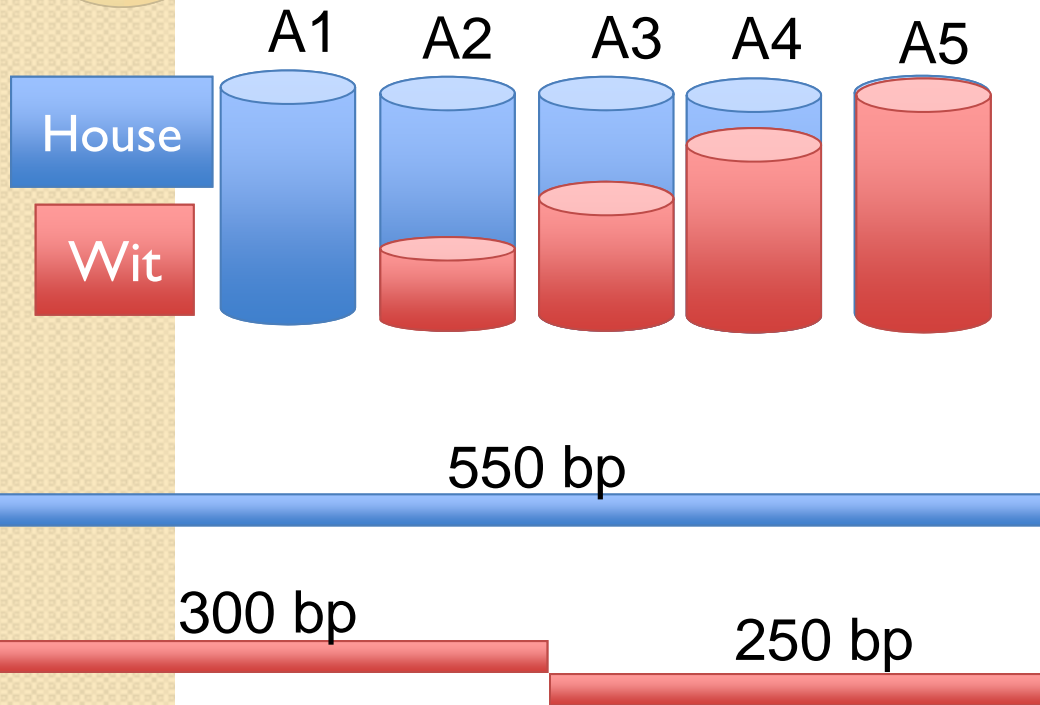
PCR to confirm pre-digestion product



A1 A2 A3 A4 A5

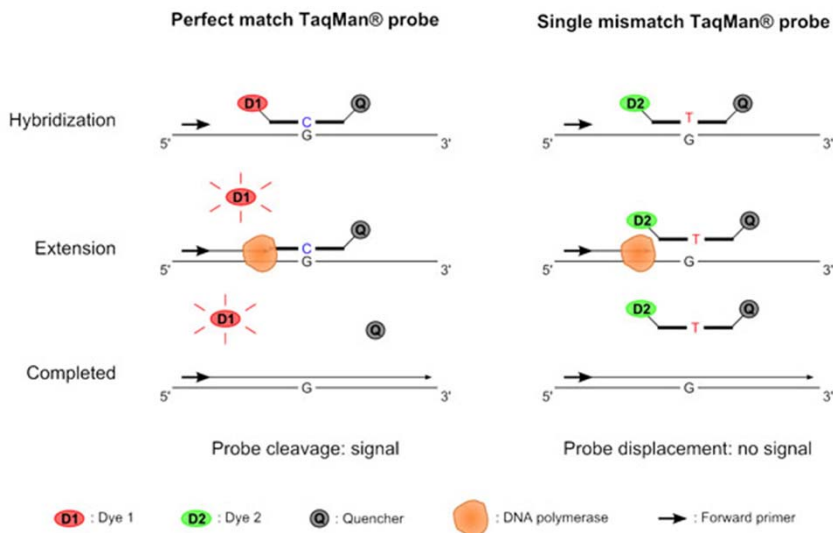


Digest reveals contamination at 25%



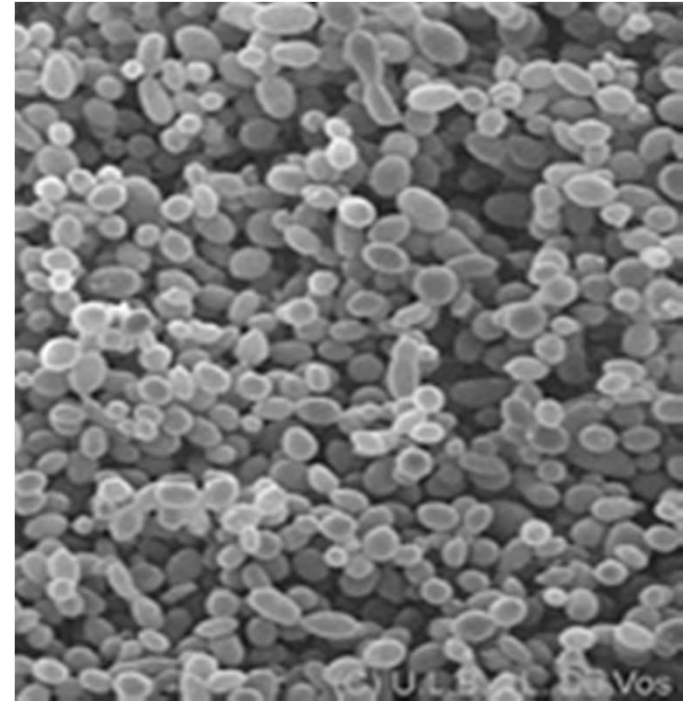
Practical Uses

- This test will be performed in addition to plating, during propagation and fermentation.
- PCR and electrophoresis
- Creation of strain-specific PCR probes for use in qPCR



Potential Downside

- Brewery-specific
- Will not detect all yeast contaminants
- Rigidity in recipe design and formulation



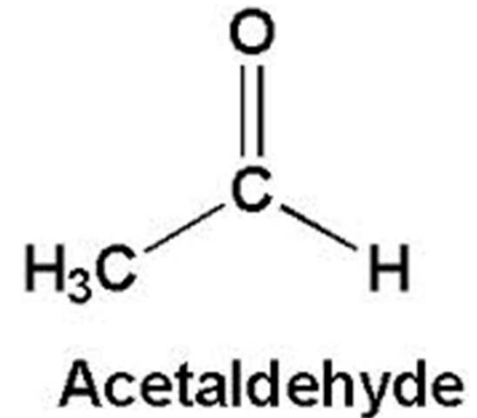
Practicality at Yeast Supplier Level

- Potential coverage of hundreds of strains
- Ensure purity and consistency for customers
- Development of a quick, quantitative assay for strain identification
- Troubleshooting

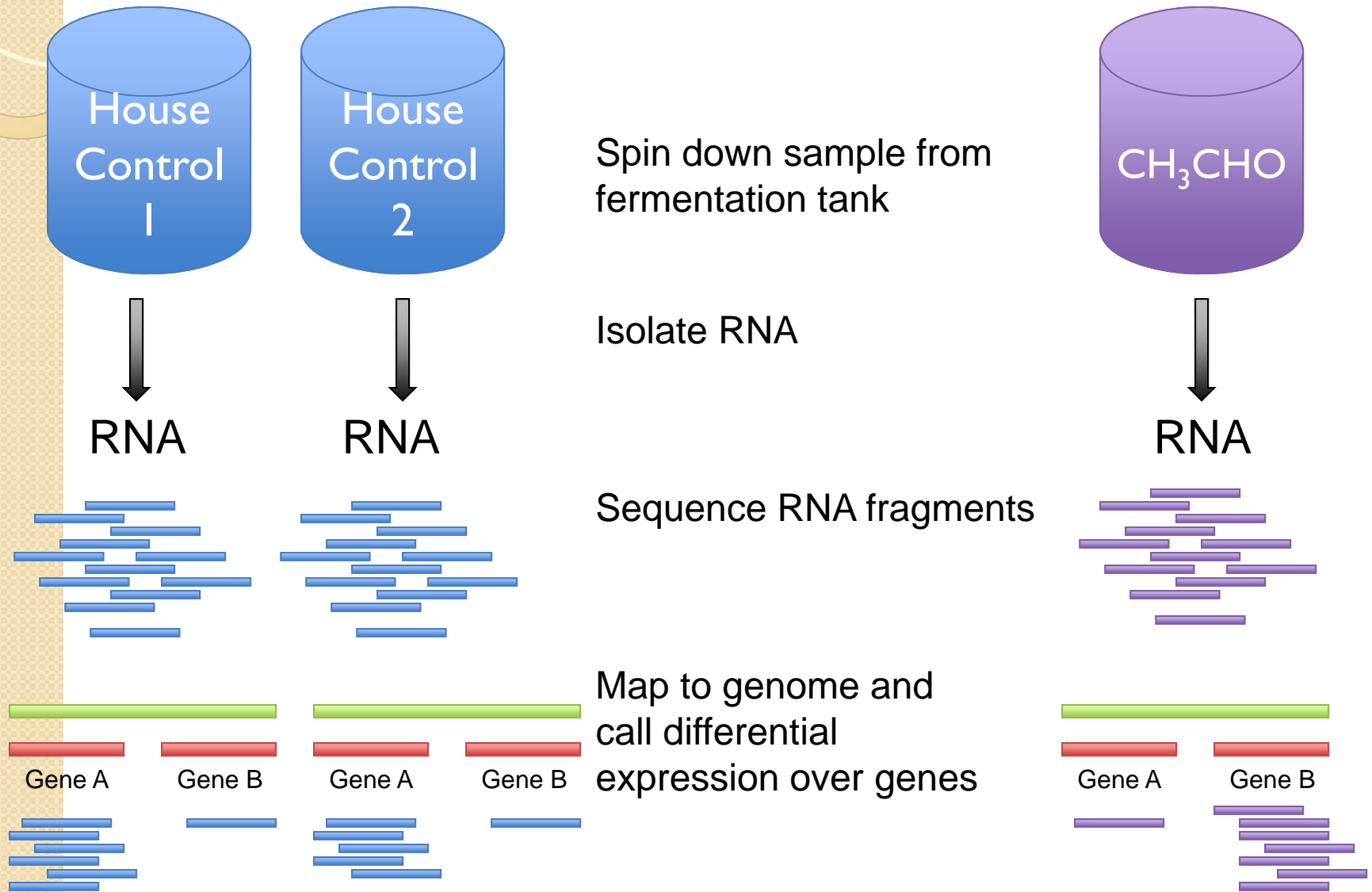


Other Possibilities for using NGS

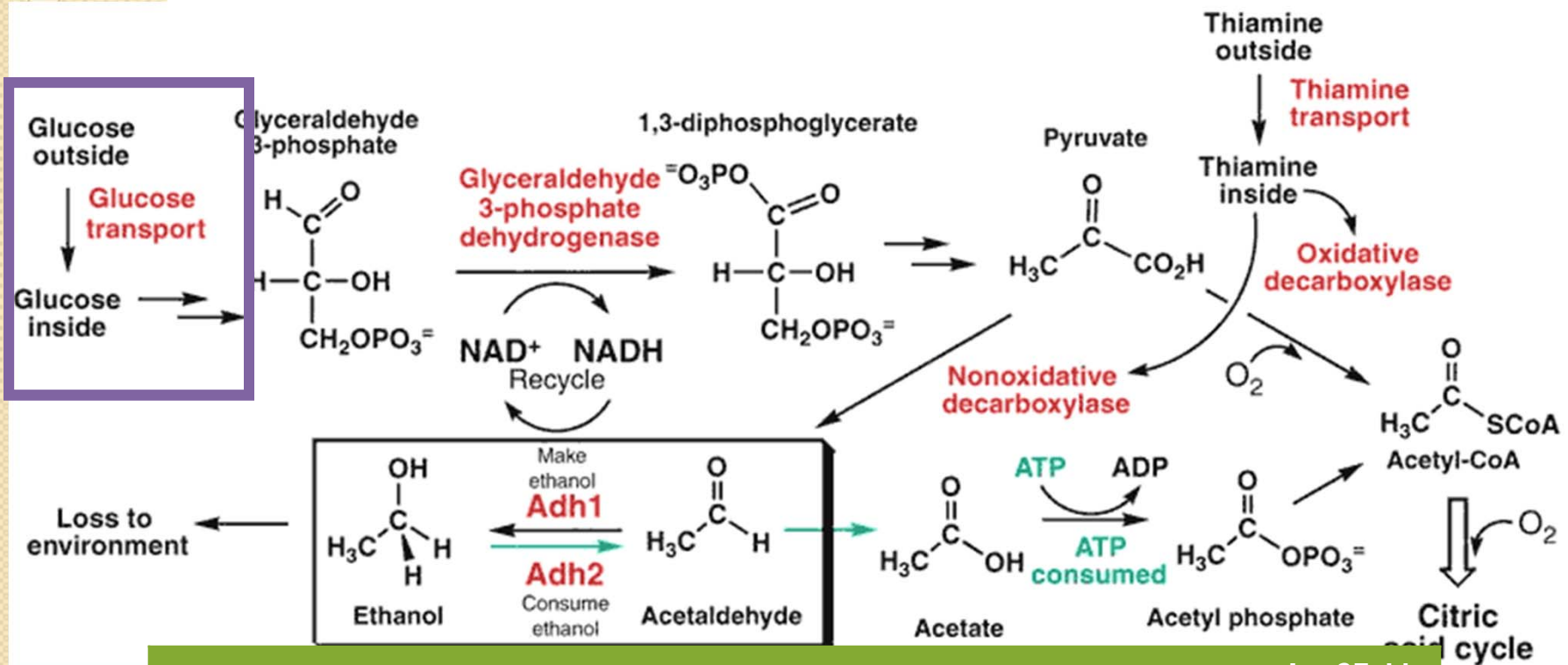
- Next generation sequencing can go beyond genomic DNA analysis
 - RNA-seq gives us a snapshot of the yeast transcriptional profile
- Poor yeast performance and off flavors
 - RNA sequencing for up or down regulation of individual genes
 - Potential changes in process control



Finding Differentially Expressed Genes

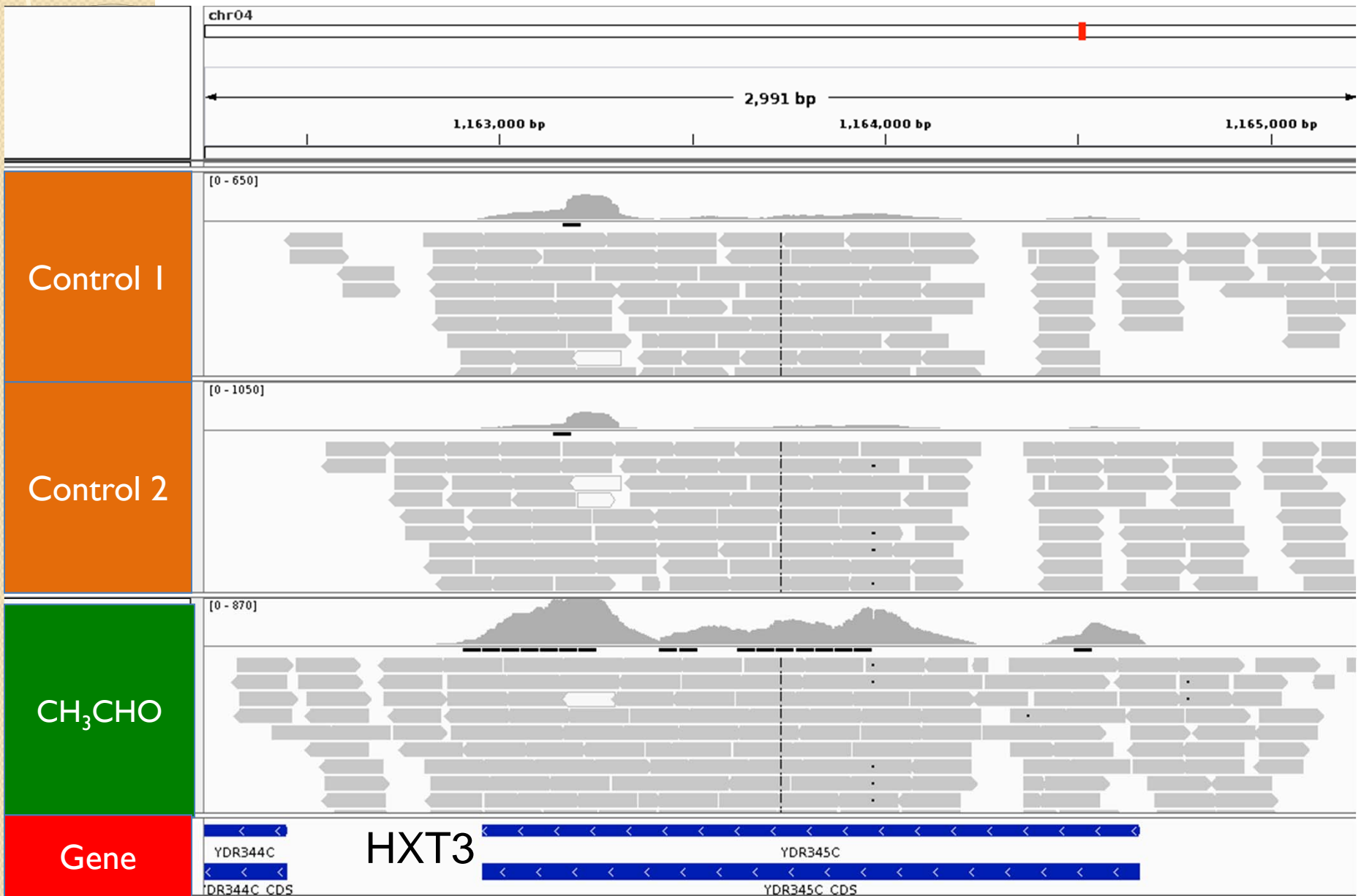


Differential Expression In the Metabolic Pathway

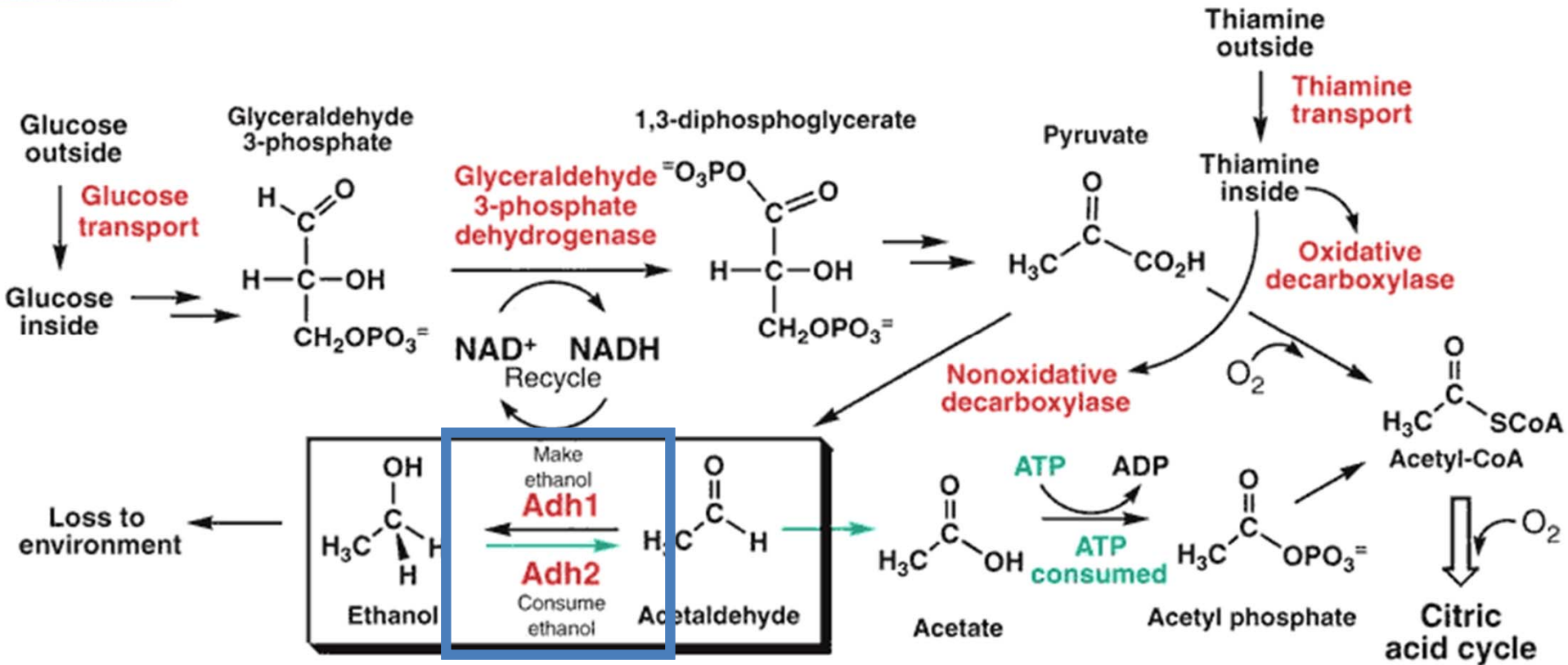


Gene	Role	Control Expression	CH ₃ CO Expression	Log2Fold Change(CH ₃ CO/Control)
HXT2	Hexose transporter induced in low glucose	62.5629	183.009	1.54854
HXT3	Hexose transporter active in low or high glucose	87.0462	407.17	2.22578
HXT4	Hexose transporter induced in low glucose	37.1486	123.099	1.72844

HXT3

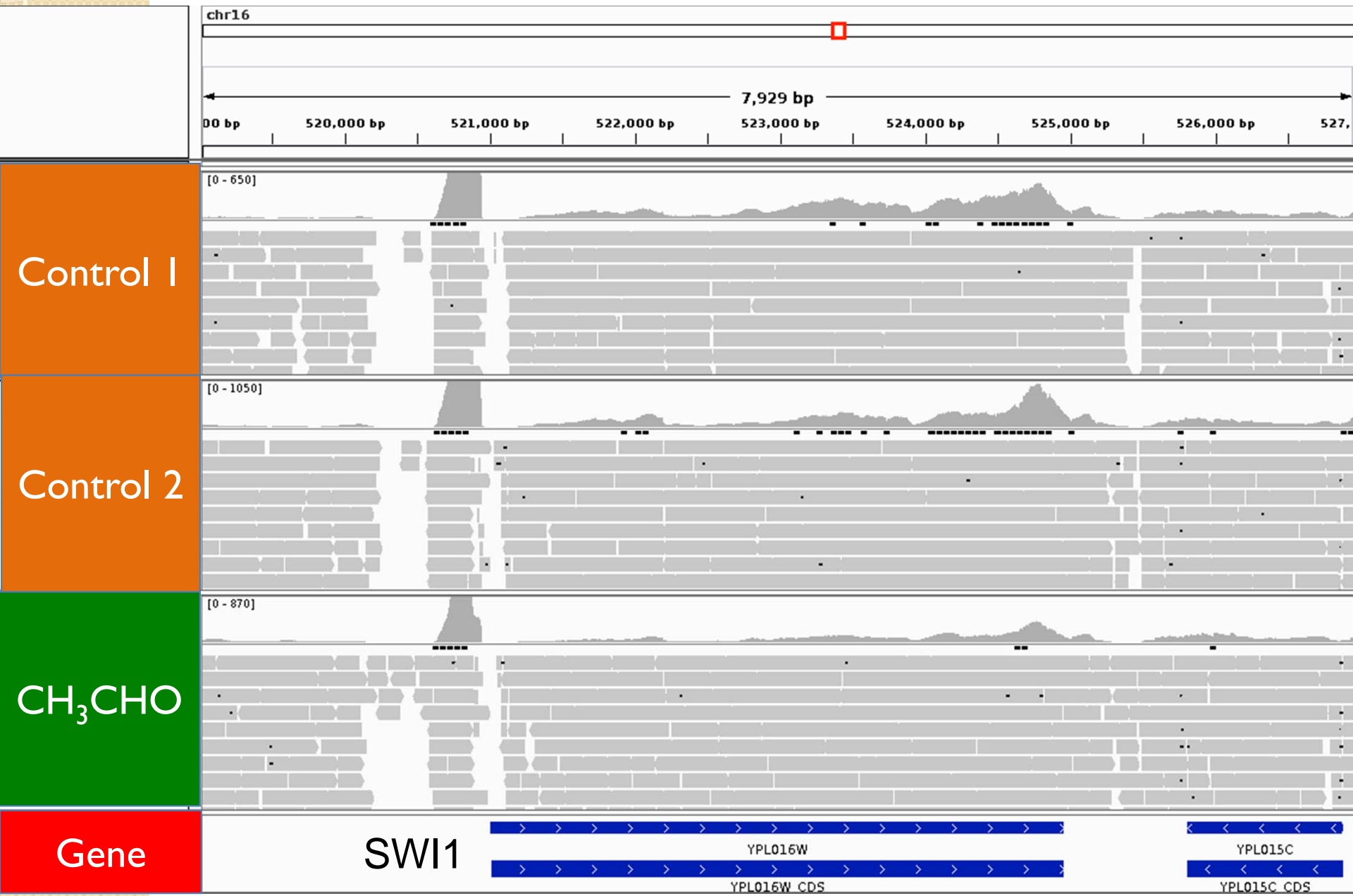


Metabolic Overview



Gene	Role	Control Expression (FPKM)	CH_3CO Expression (FPKM)	Fold Change
SWII	Controls expression of ADH1/ADH2	255.942	85.727	-1.578

SWI1



Ongoing work



- RNAseq has given us a unique look at this confounding problem
- Differential expression surrounding glucose metabolism could hint at different carbon levels present in the media
- Has potential to influence brewing practices in the future



Industry Recommendations

- Be aware of potential yeast cross contamination if using more than one strain
- Consider the use of Next Generation Sequencing to address QA/QC issues
- Seek out collaborations with academic institutions
- Be open to the exchange/publication of information that may be applicable to other brewers

Presentation Recap

- Bacterial or wild yeast contamination are not the only types of contamination in a brewery
- Next Gen Sequencing as a potential tool
- Development of brewery-specific yeast purity assay for Avery
- Potential use of NGS data extends to yeast suppliers as well
- Other ongoing applications in brewing science



Acknowledgements



Dowell Lab

Robin Dowell

BioFrontiers Institute Next-Gen
Sequencing Facility

Jim Huntley

Avery Brewing Company

Avery Production Staff



BioFrontiers researcher Robin Dowell, Avery microbiologist Dan Driscoll and Next-Gen Sequencing Facility Director Jim Huntley pose with Avery beer at BioFrontiers. Photo by Casey Cass.



Questions?

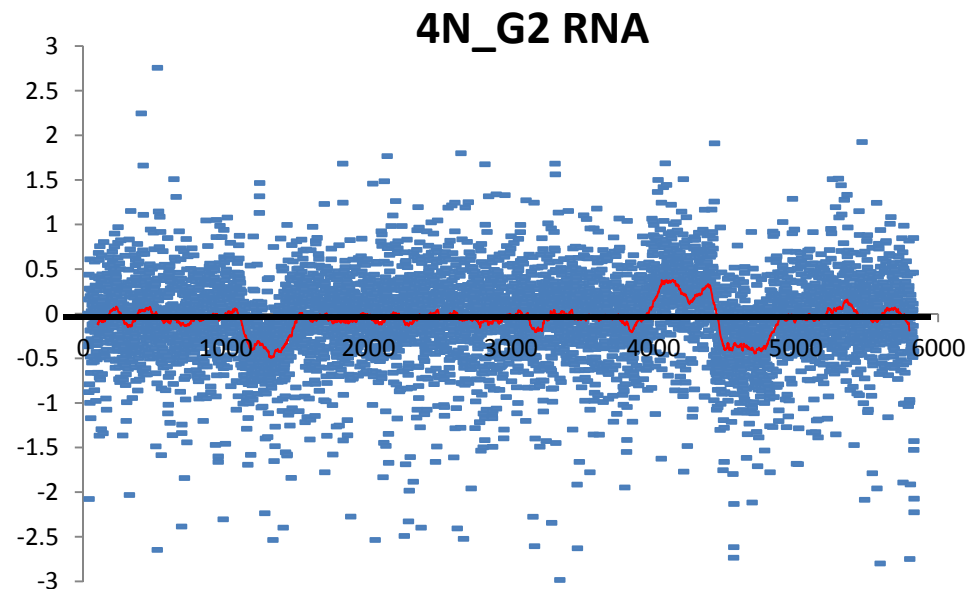


Exploratory RNA Seq

- ⦿ Acetaldehyde buildup
 - Matched gravity between samples, burped off bottom of fermentation tank, pelleted cells, isolated RNA.
 - One Squashy sample
 - Two Control samples
- ⦿ RNAseq provides genomic coverage to check for mutations between samples
- ⦿ Attempting to see if expression patterns emerge from comparing the squashy and control samples.

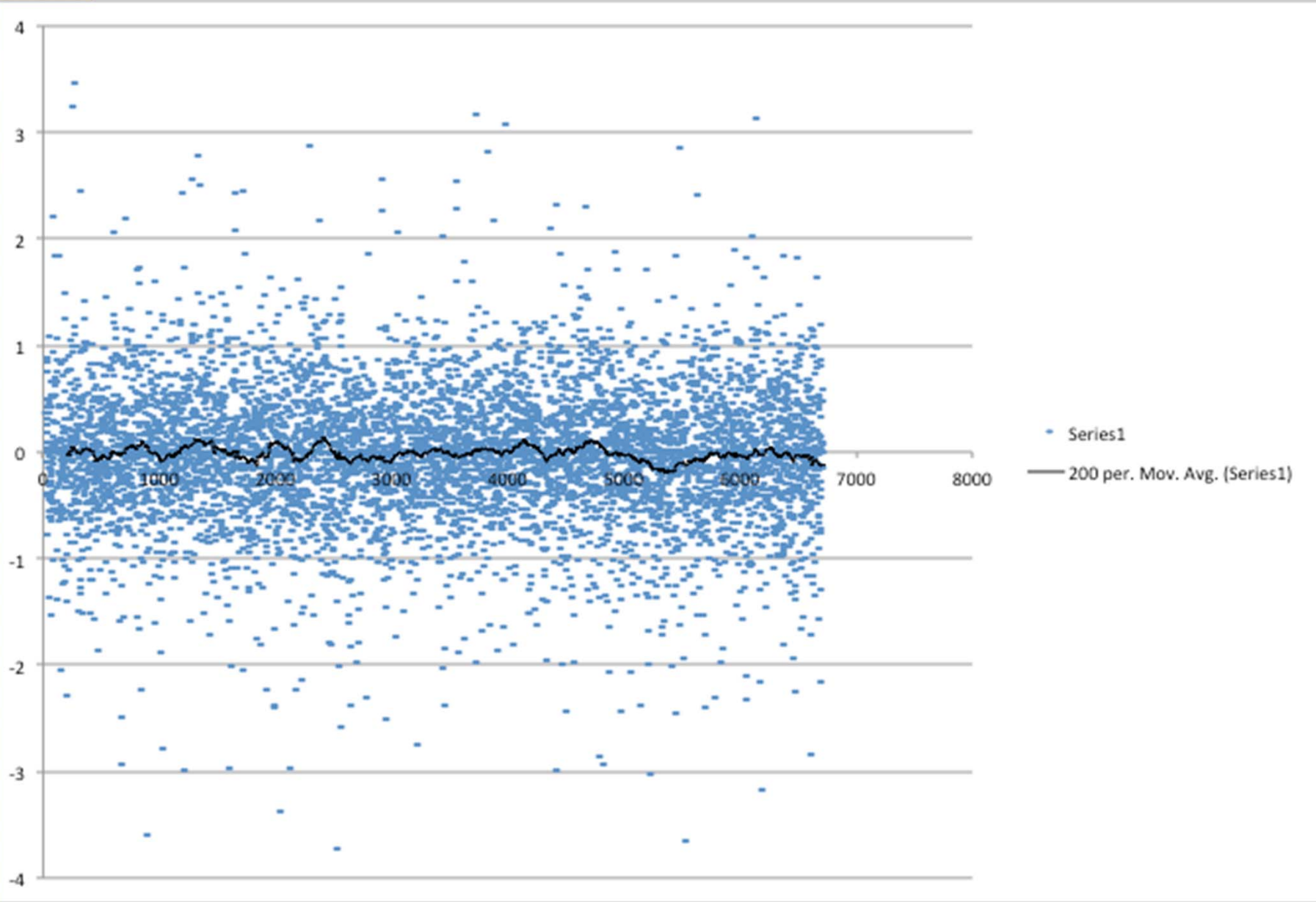
Genomic Differences from RNAseq

- Wanted to see if the problem in the fermentation was caused by a mutation that swept the population
 - Chromosomal copy number changes
 - SNPs

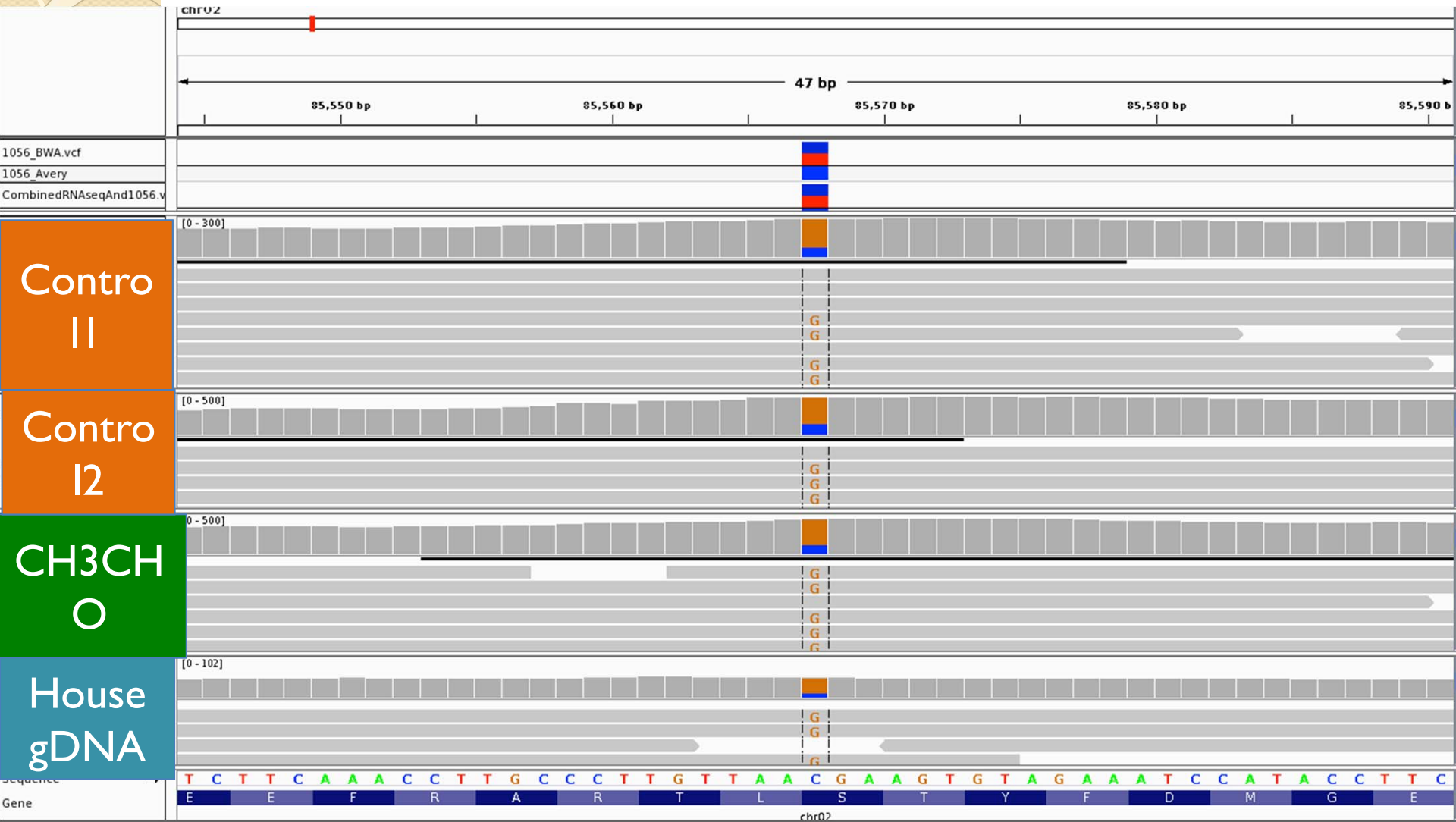


No large chromosomal changes

Log2Fold Change (Squashy/Control)

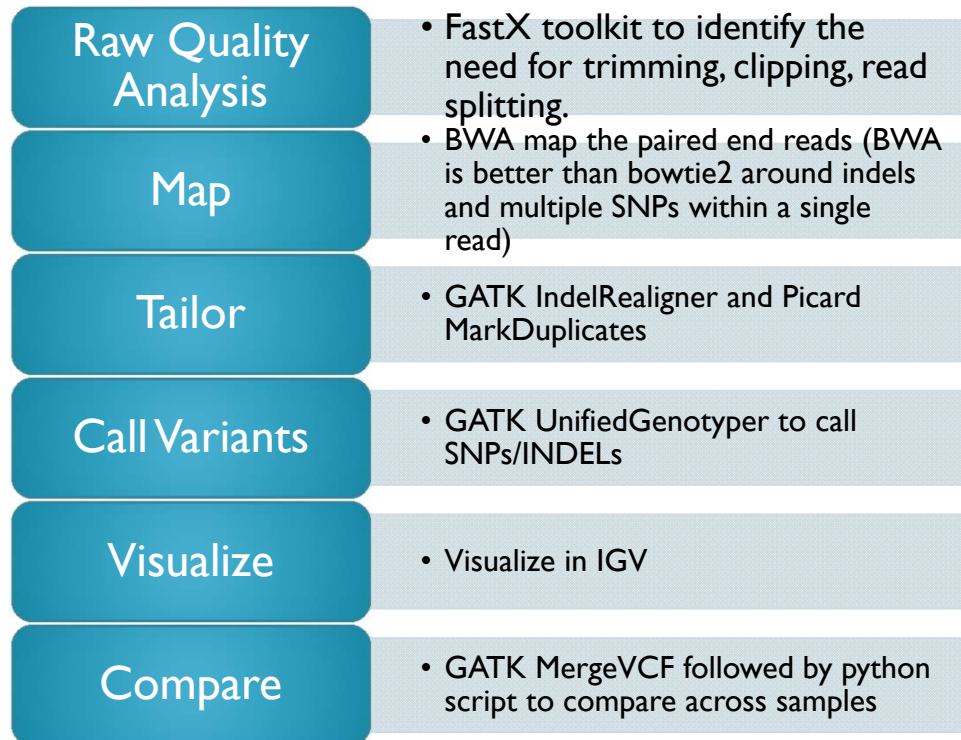


Allelic frequencies match up



No standout SNPs called uniquely to Squashy sample

- SNP calling pipeline on Control1, Control2, Squashy1
- Intersect SNPs, define unique SNPs, examine Squashy-unique set
- No convincing SNPs present in dataset



Calling Differential Expression

- Ran the Cufflinks pipeline to get differential expression on the RNA-seq data
- 108 differentially expressed genes (q-value < .05)
- 285 differentially expressed genes (q-value < 0.1)
- GO primarily enriched for cell-cycle related genes (samples weren't matched in growth phase)

SNP calling pipeline

Raw Quality Analysis

- FastX toolkit to identify the need for trimming, clipping, read splitting.

Map

- BWA map the paired end reads

Tailor

- GATK IndelRealigner and Picard MarkDuplicates

Call Variants

- GATK UnifiedGenotyper to call SNPs/INDELs

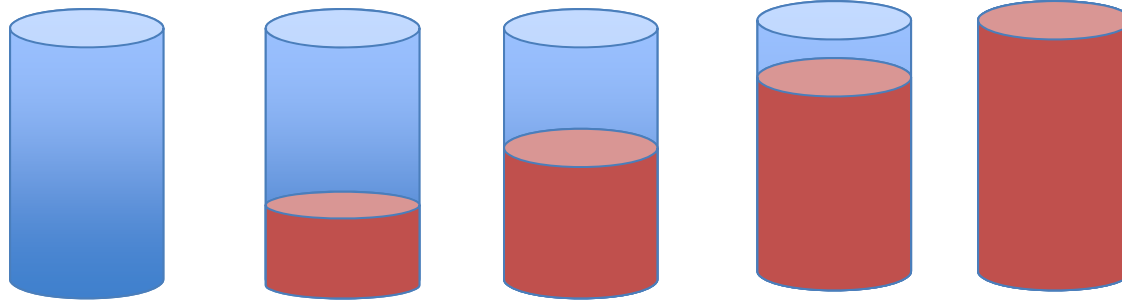
Visualize

- Visualize in IGV

Compare

- GATK MergeVCF followed by python script to compare across samples

“Colony”-PCR prep



Centrifuge 2 min. to pellet cells



20mM NaOH, 10 minutes at 95°C



Quick Spin again to pellet cells



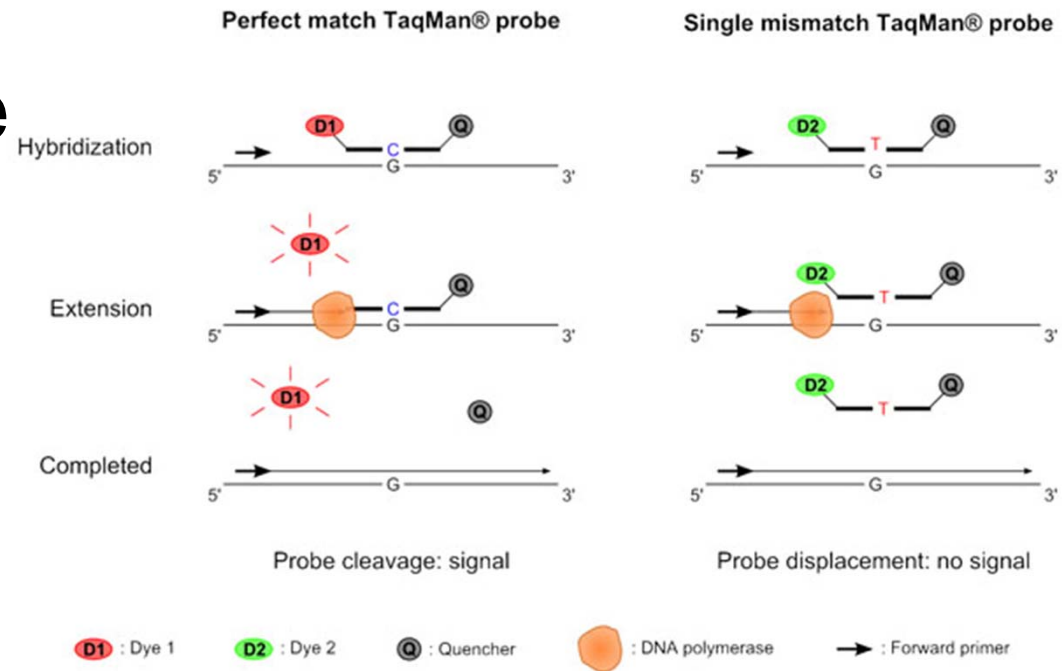
PCR using 1:10 dilution of supernatant



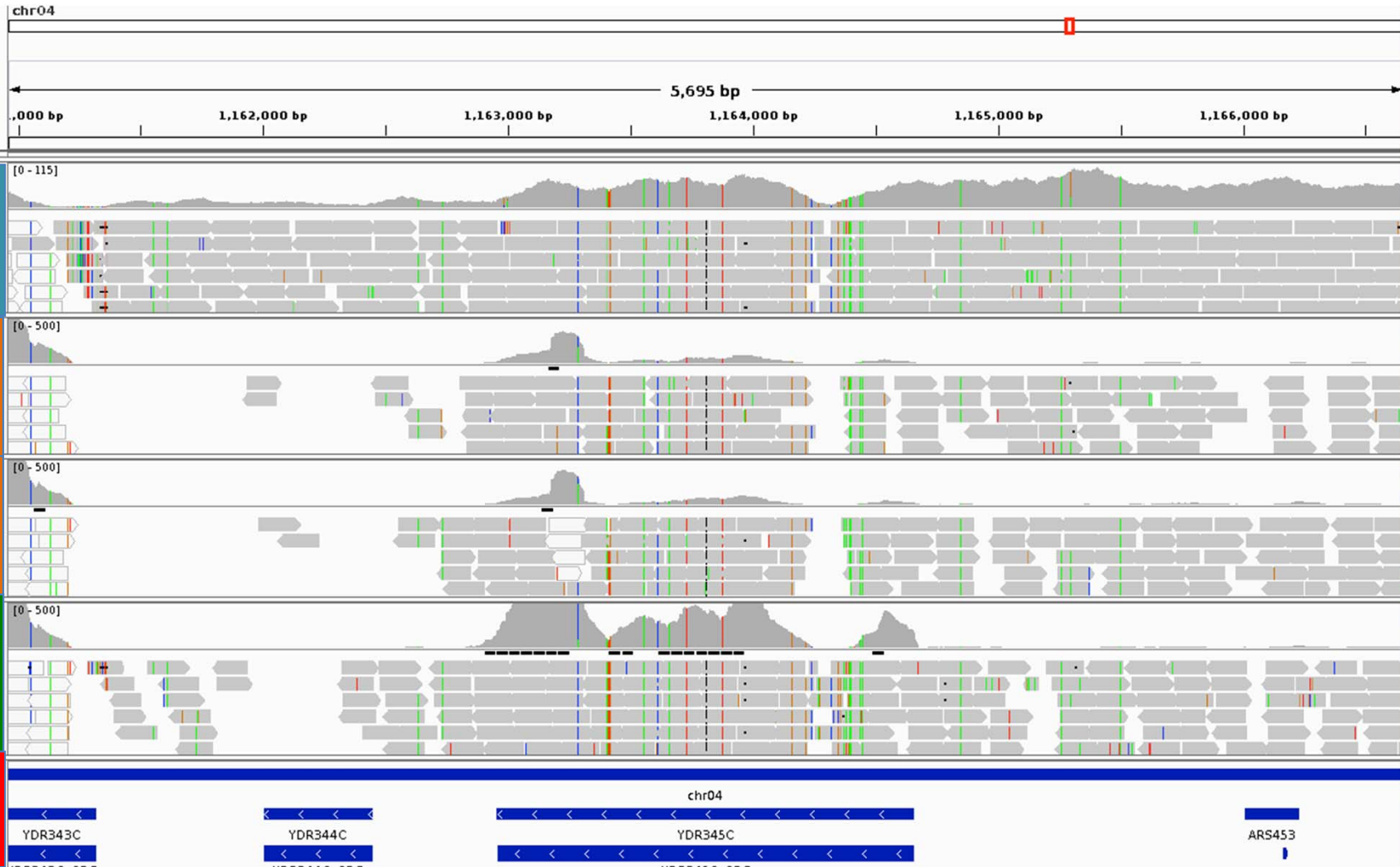
Restriction Digest PCR products

With the use of a qPCR machine

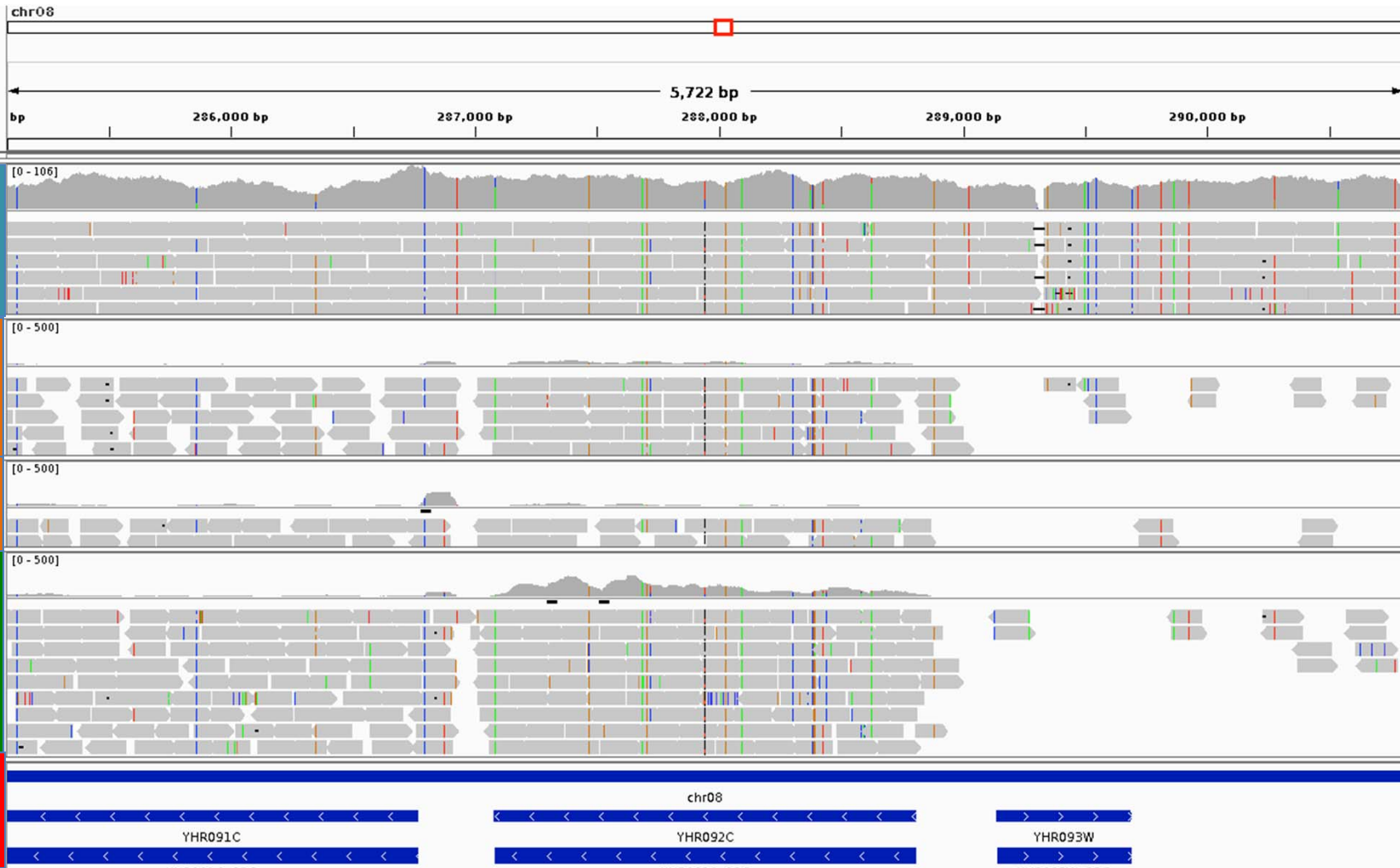
- Alternatively use a qPCR machine with probes designed over SNP dense regions (higher sensitivity, more expensive overhead)



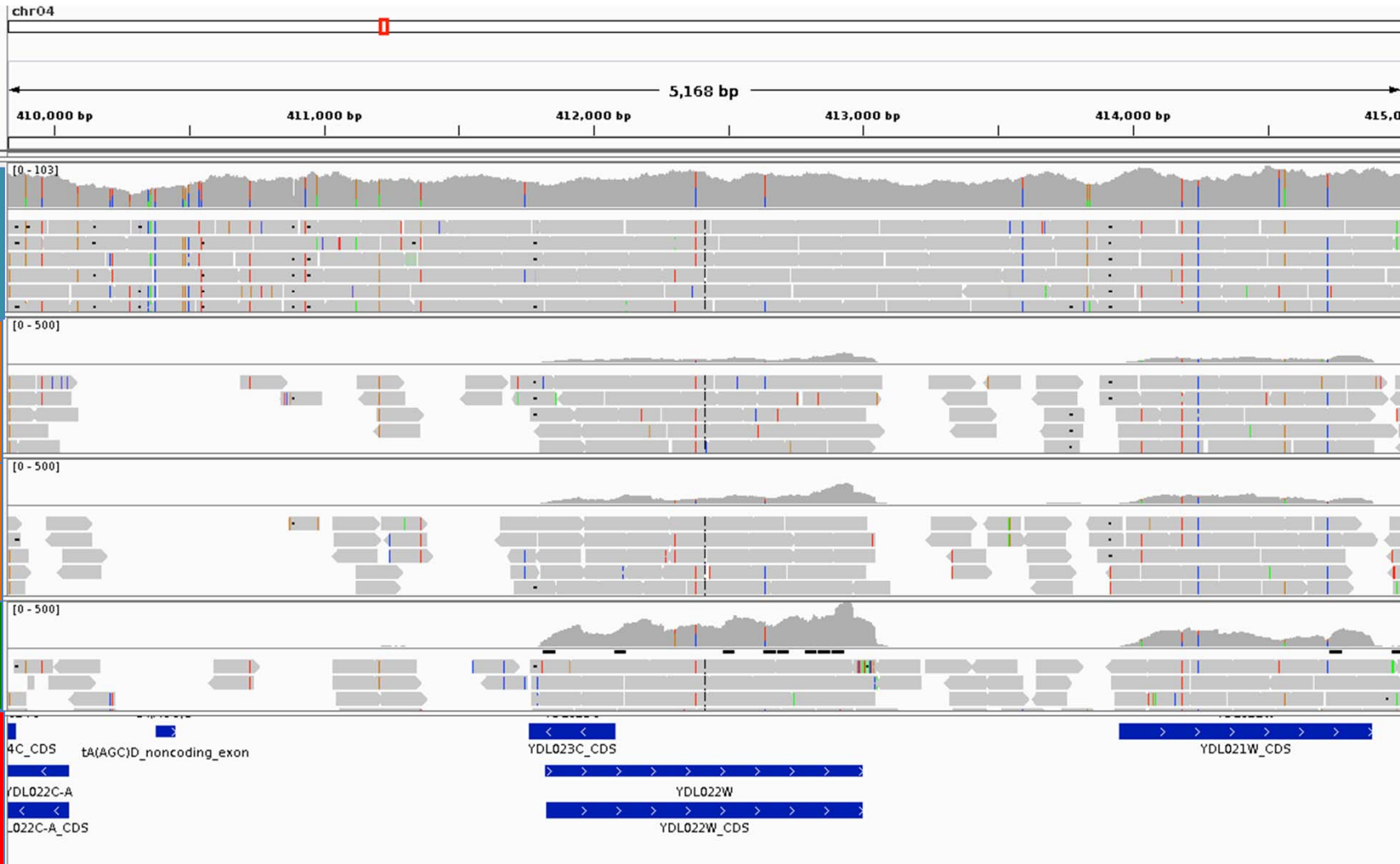
HXT3



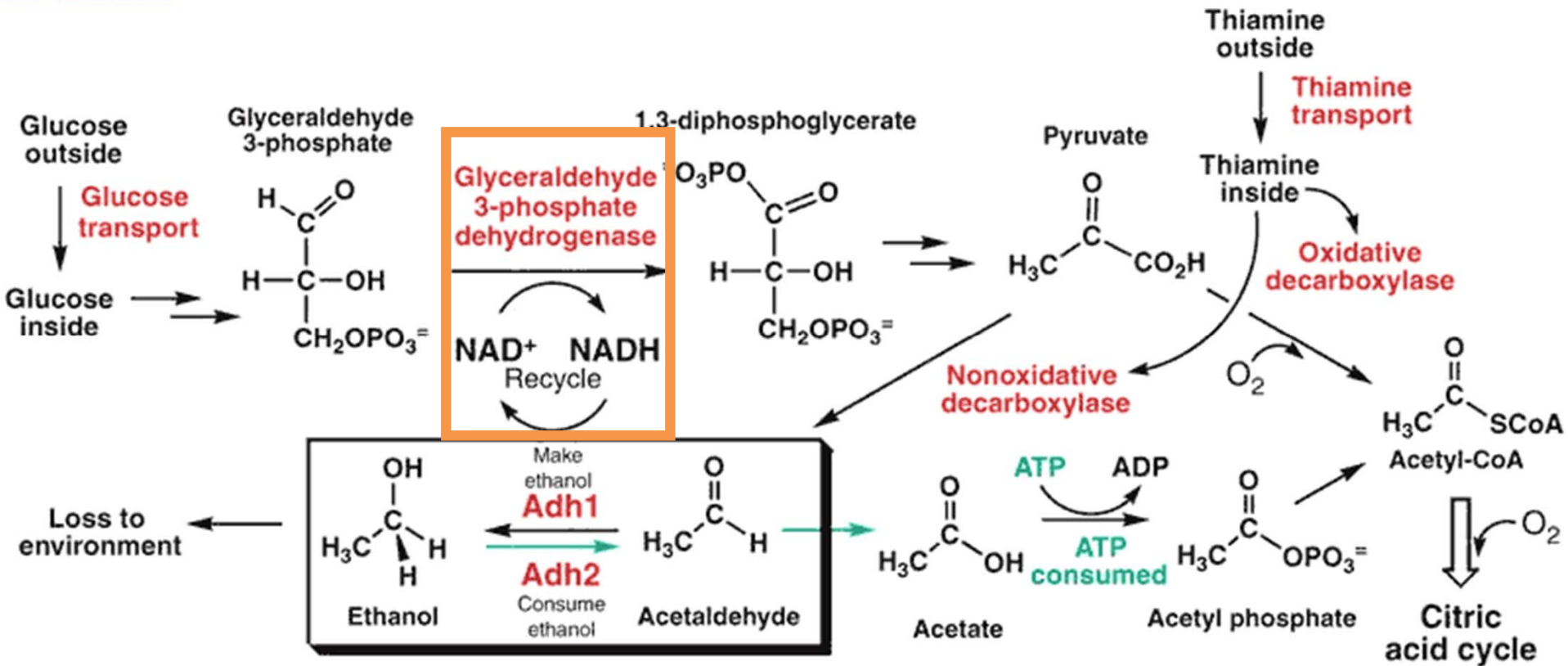
HXT4



GPD1



Metabolic Overview



Gene	Role	Control Expression (FPKM)	CH ₃ CO Expression (FPKM)	Log ₂ Fold Change(CH ₃ CO/Control)
GDPI	Glyceraldehyde Dehydrogenase	72.9632	188.643	1.37041