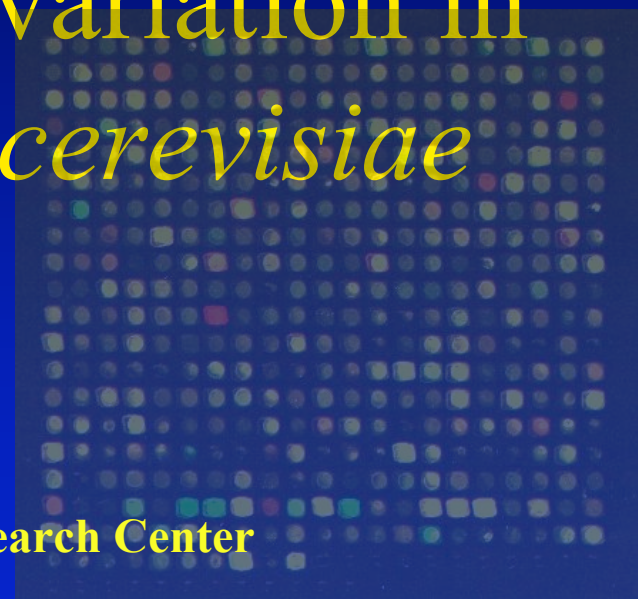
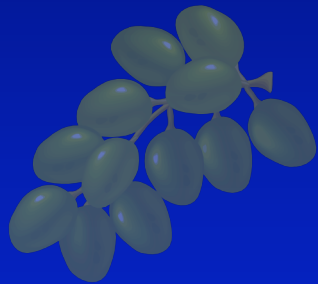


Local regulatory variation in *Saccharomyces cerevisiae*



Rachel Brem
Fred Hutchinson Cancer Research Center
Seattle, WA



A G A T T C G A C T T G A C G G T A C C A T
A G C T T C G A T T T G G C G G T A T C A T

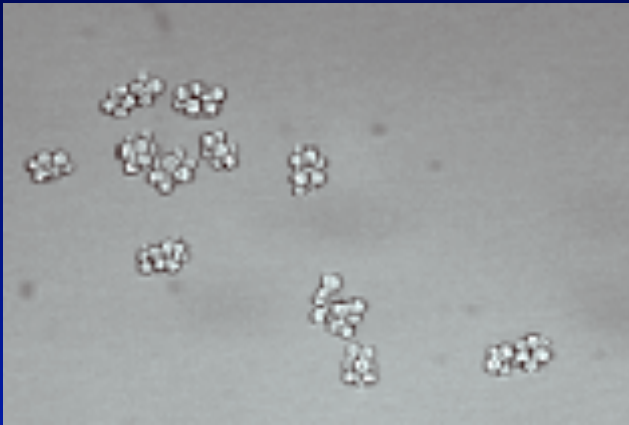
Why natural genetic variation?

- Therapeutic targets
- Diagnostics
- Agricultural breeding
- Evolutionary theory

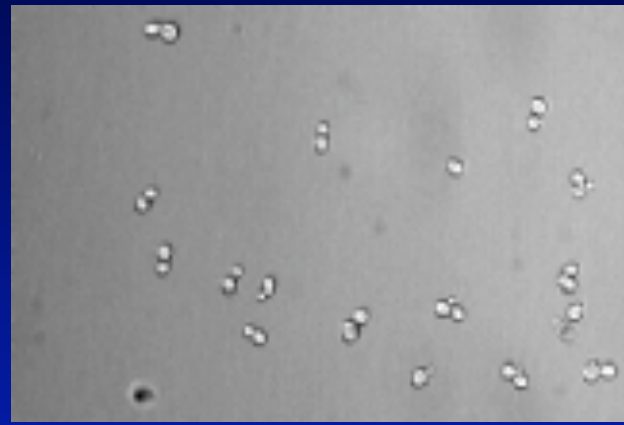


The challenge: genetic complexity

Genetically diverse yeast strains



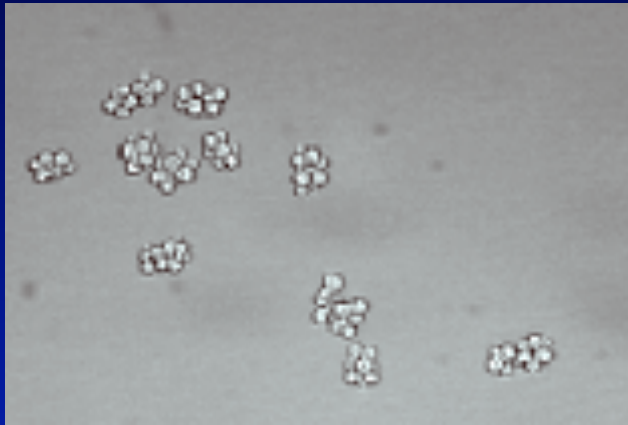
RM, vineyard



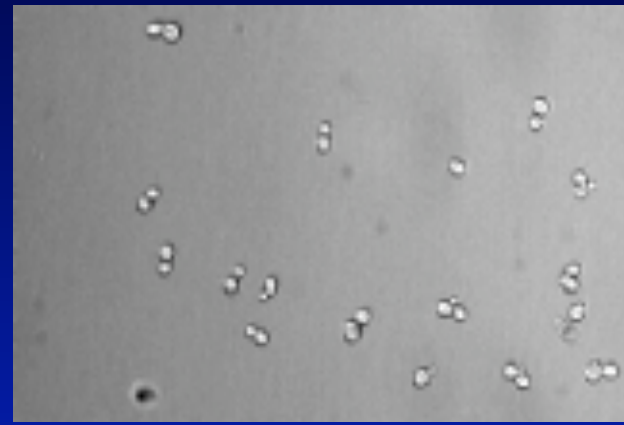
BY, lab

Differ at 1:100 nucleotides

Genetically diverse yeast strains

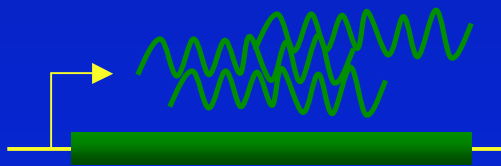


RM, vineyard

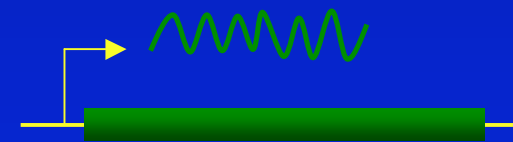


BY, lab

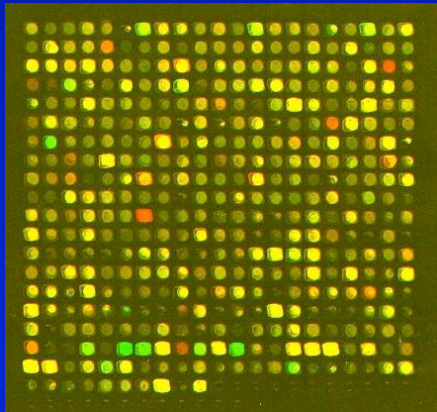
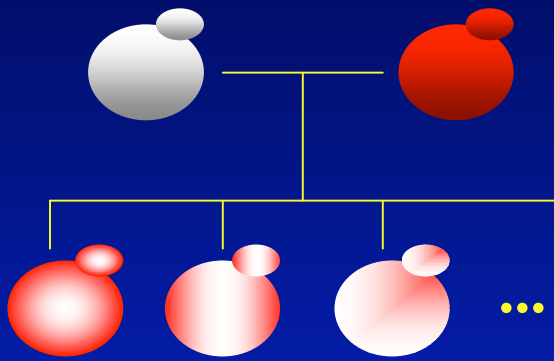
Differ at 1:100 nucleotides



**Natural variation
in gene expression**

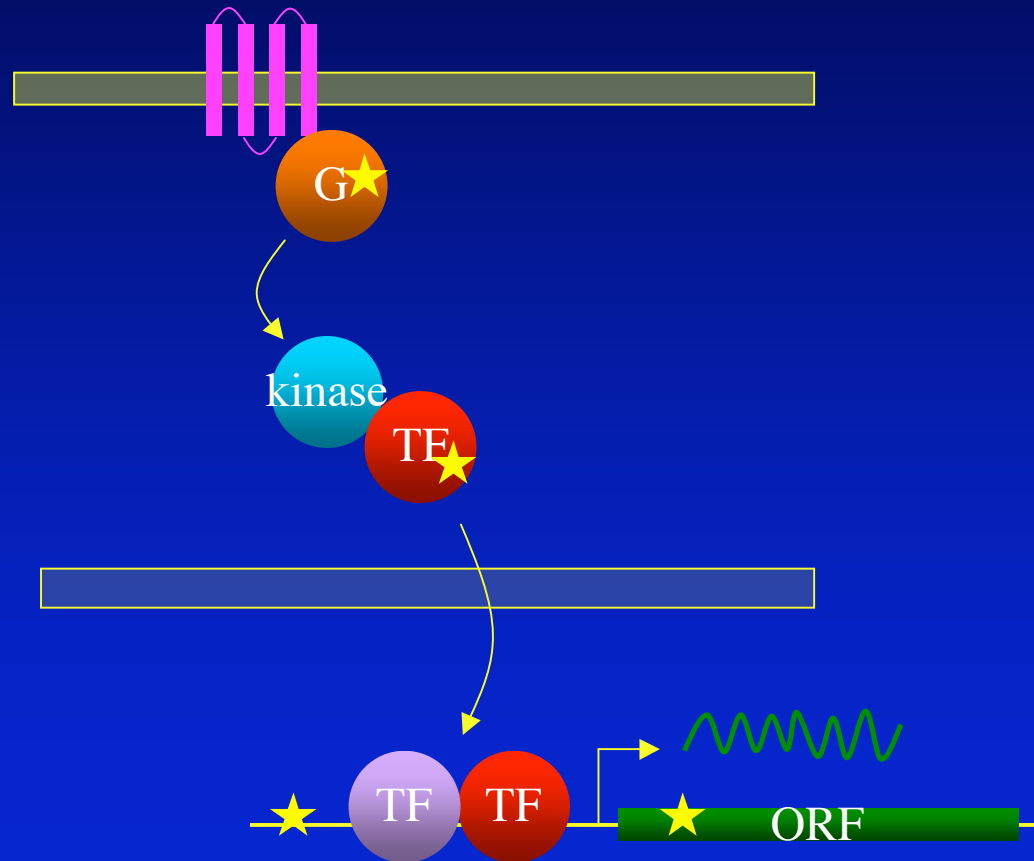


Why expression, why yeast?

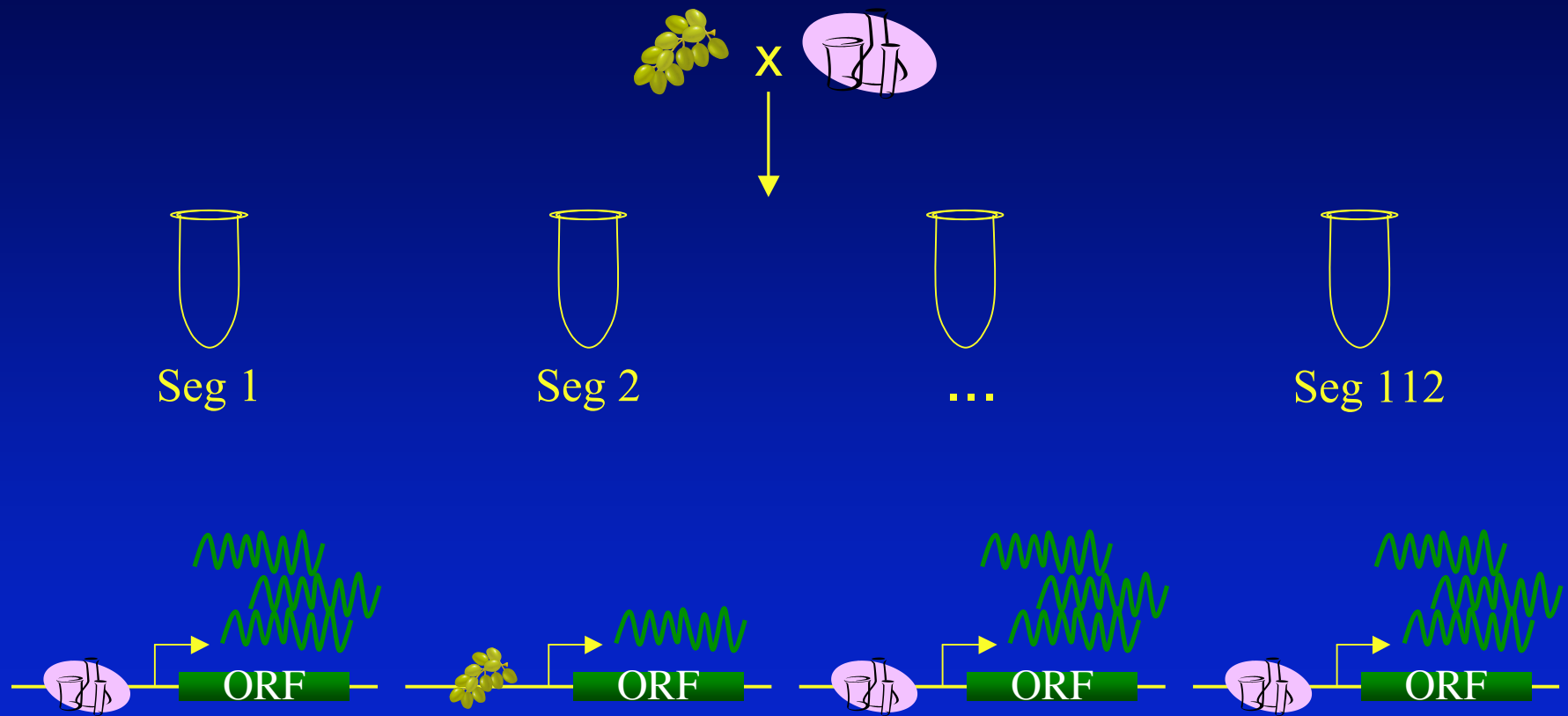


- Medical genetics of transcriptional reg.
- Links in regulatory networks
- Map loci of modest effect
- Down to single SNPs
- Test hypotheses in vivo
- 1000's of QTLs

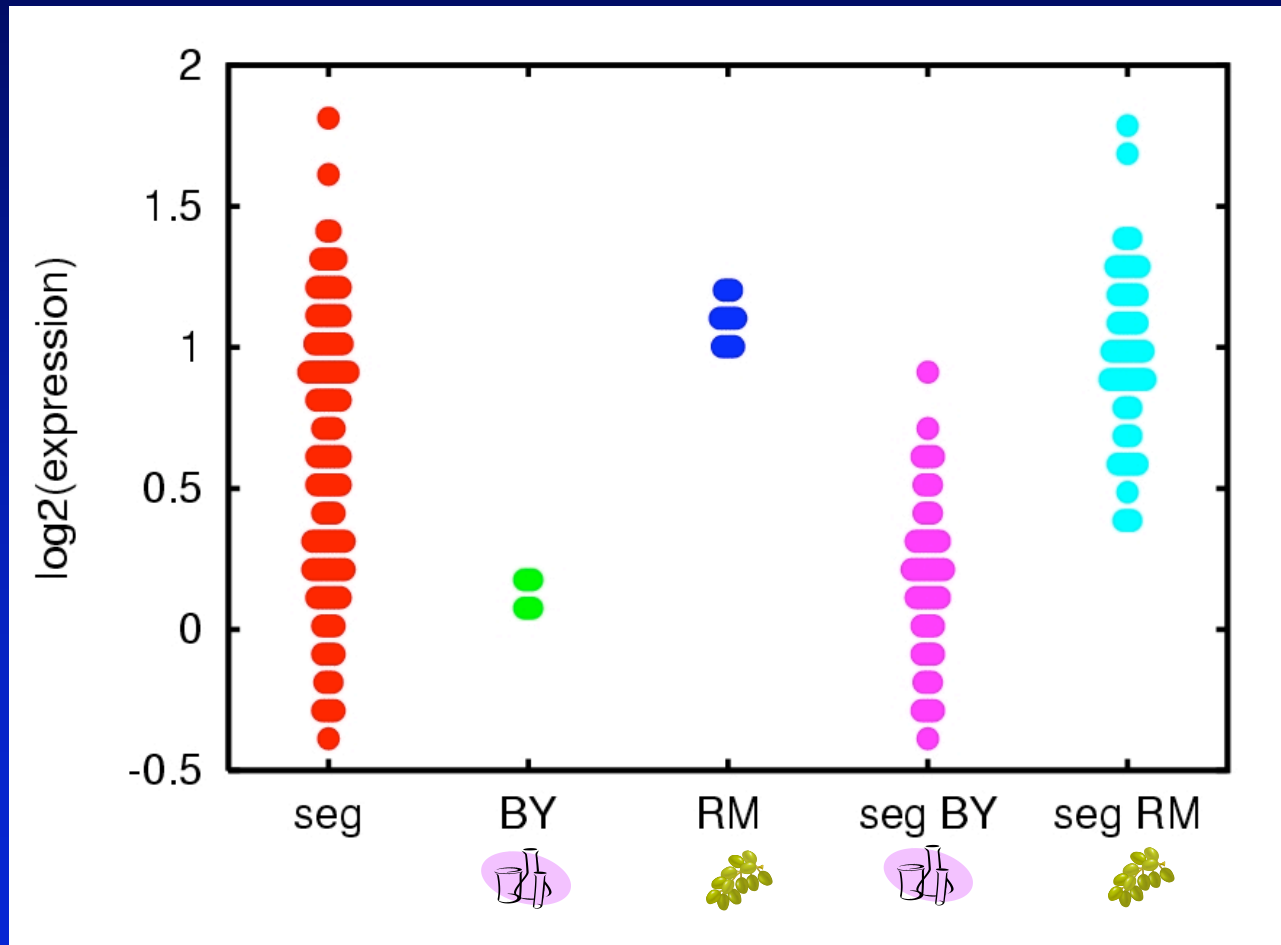
Local vs. nonlocal variation



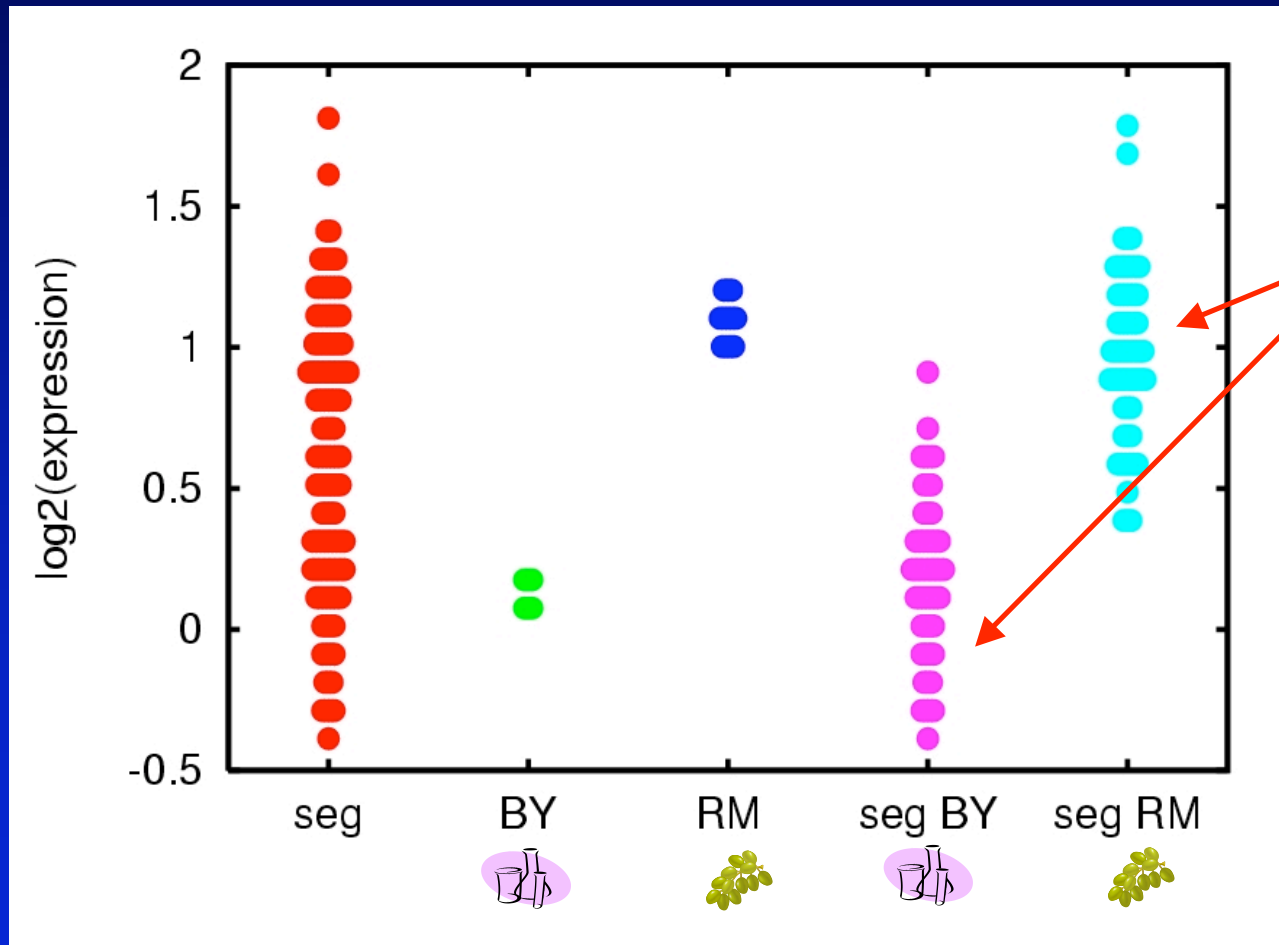
Linkage test for local variation



Linkage test for local variation



Linkage test for local variation

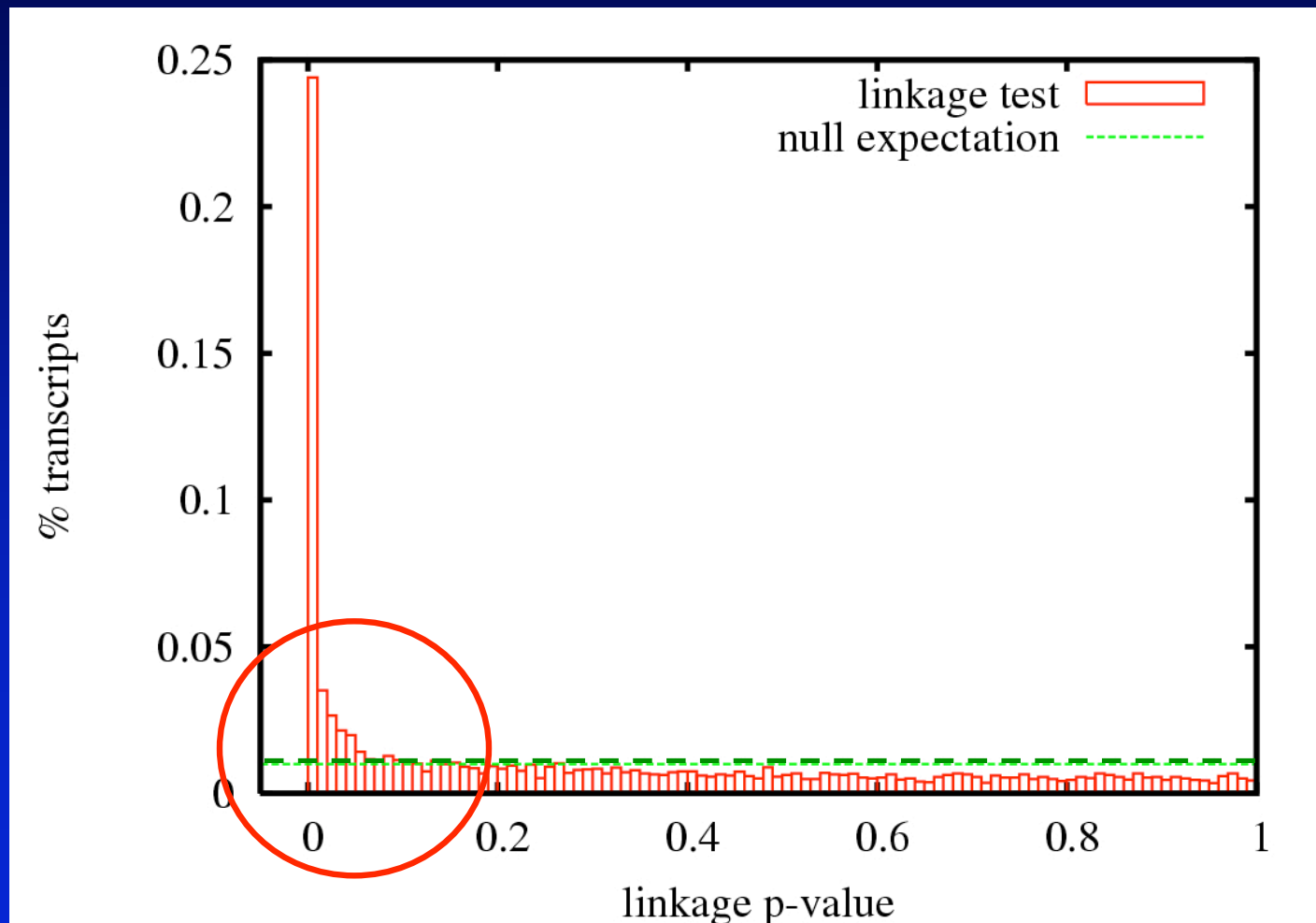


Wilcoxon test
p-value;
permutations
for false
positive count

1428 transcripts show evidence for local
variation at $p < 0.0122$

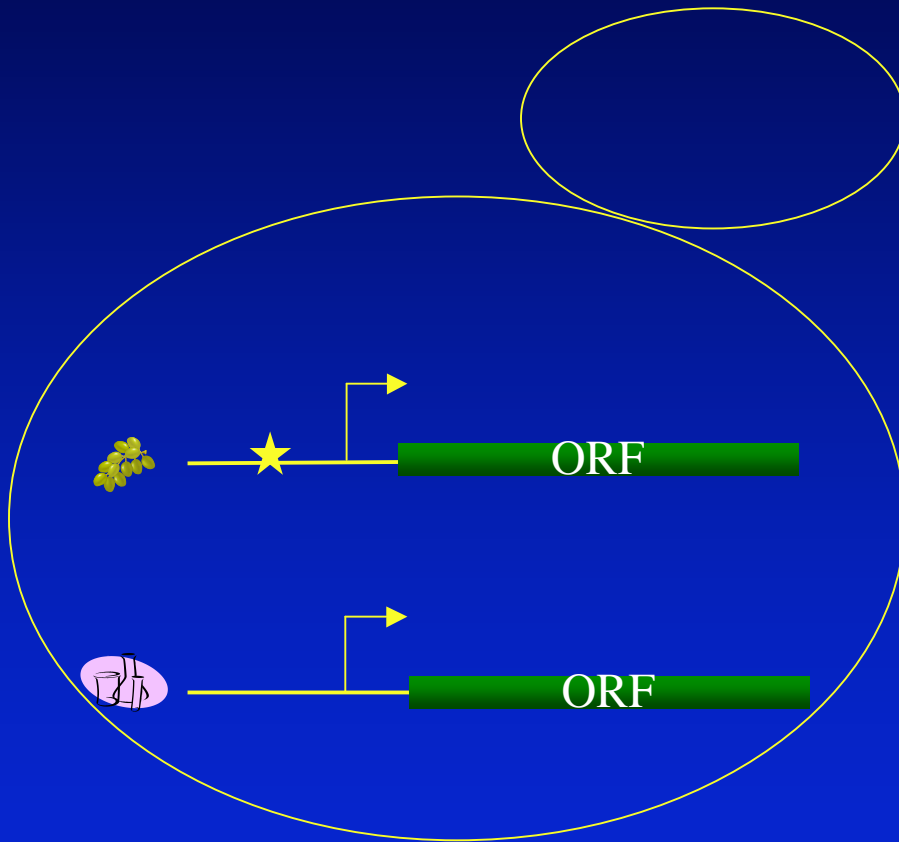
71 null transcripts by chance (FDR = 0.05)

True prevalence of local variation ~50%

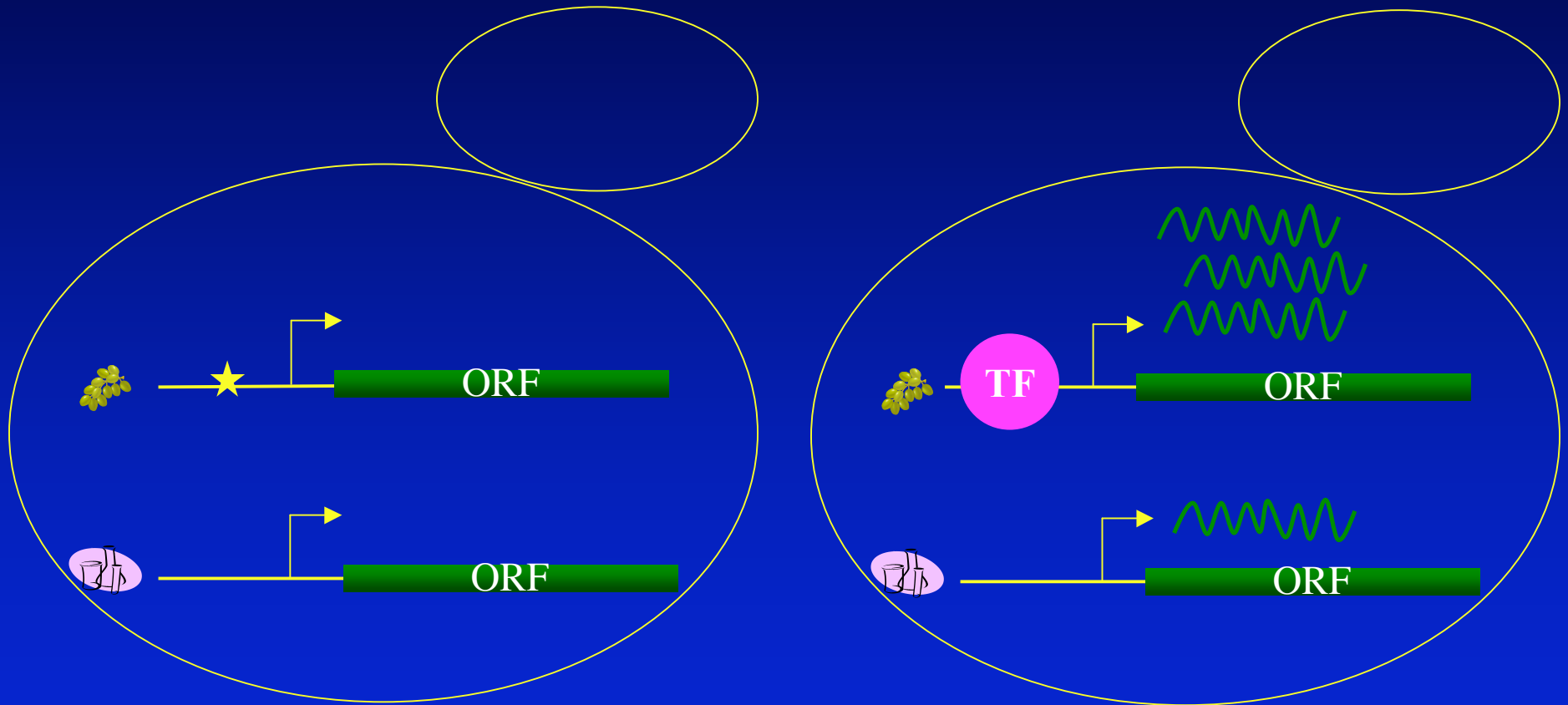


QVALUE
estimate:
49% of
tests are
true
positives

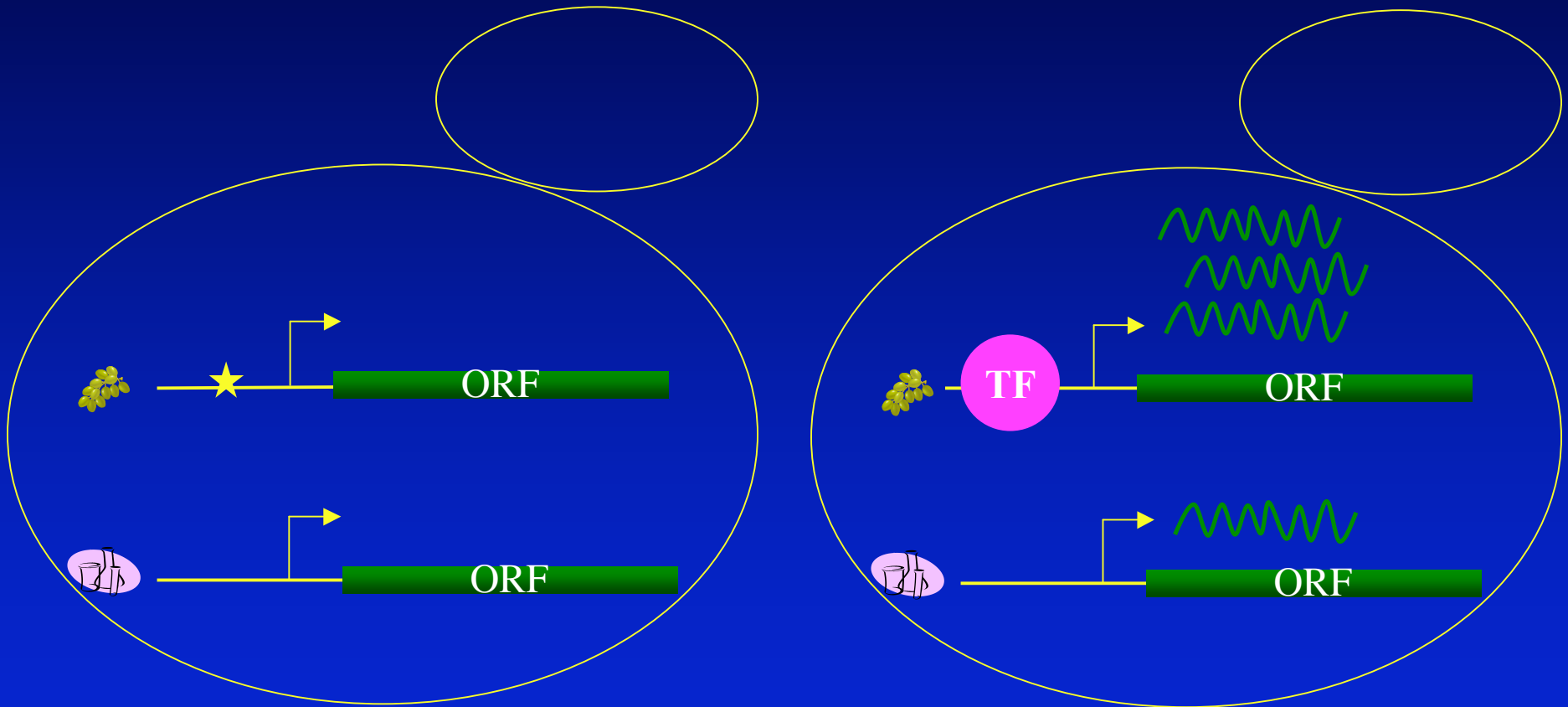
Allele-specific expression



Allele-specific expression

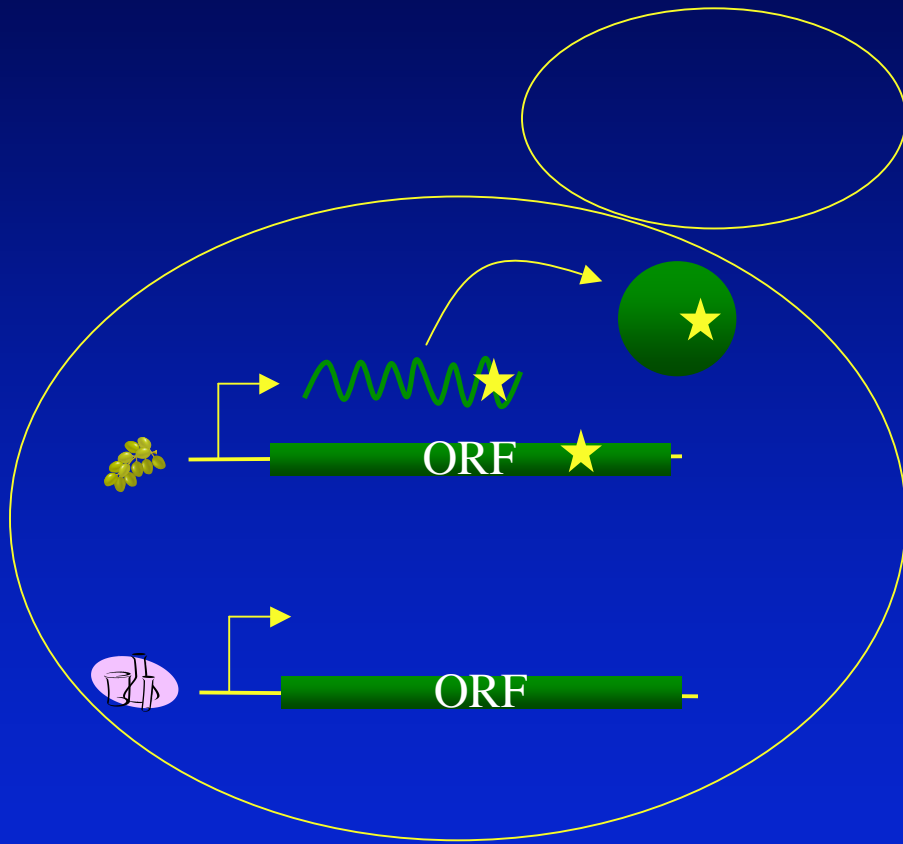


Allele-specific expression

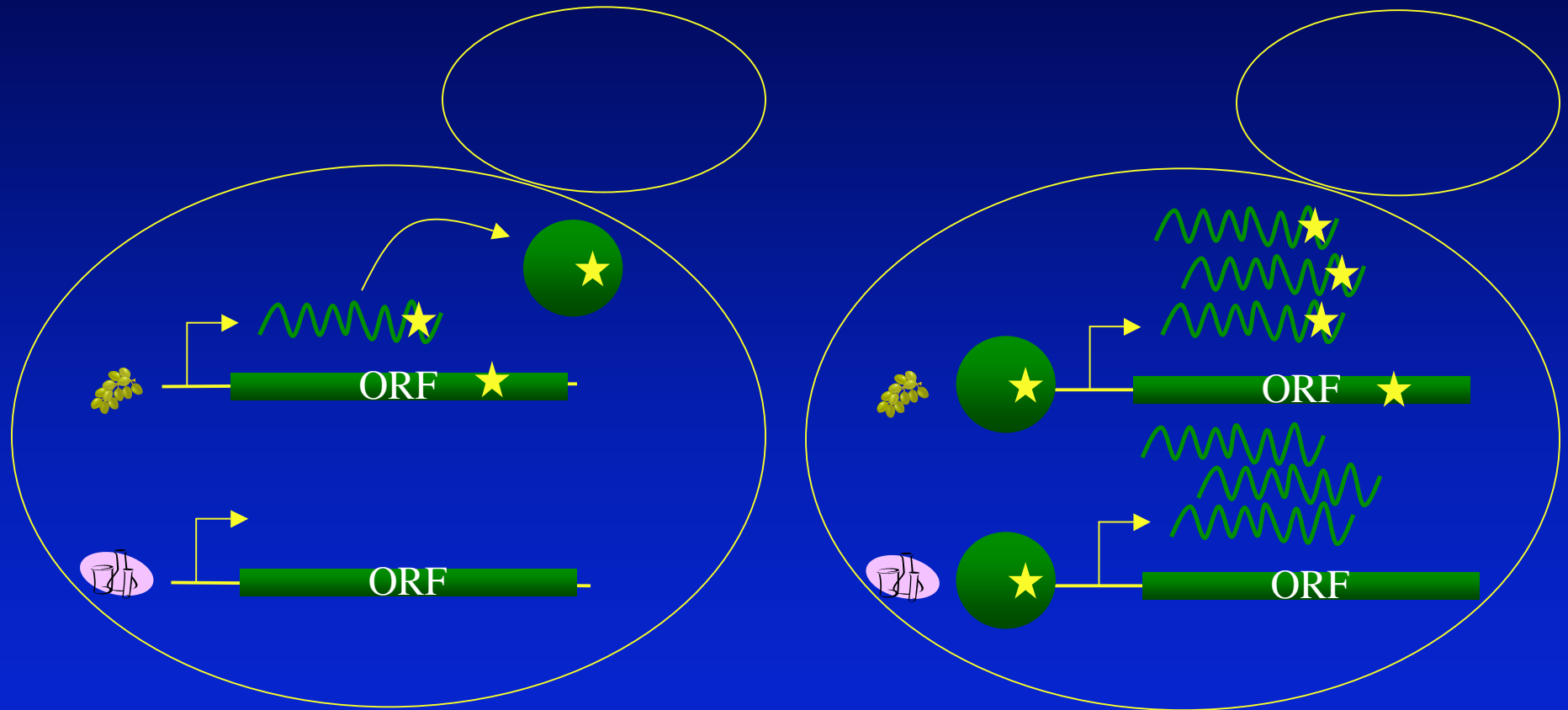


TaqMan: transcript from each copy labeled with different dye

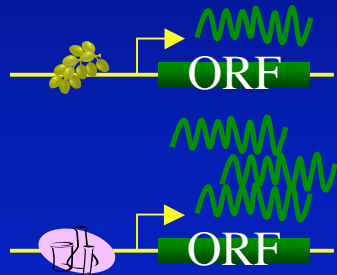
Non-allele-specific expression



Non-allele-specific expression



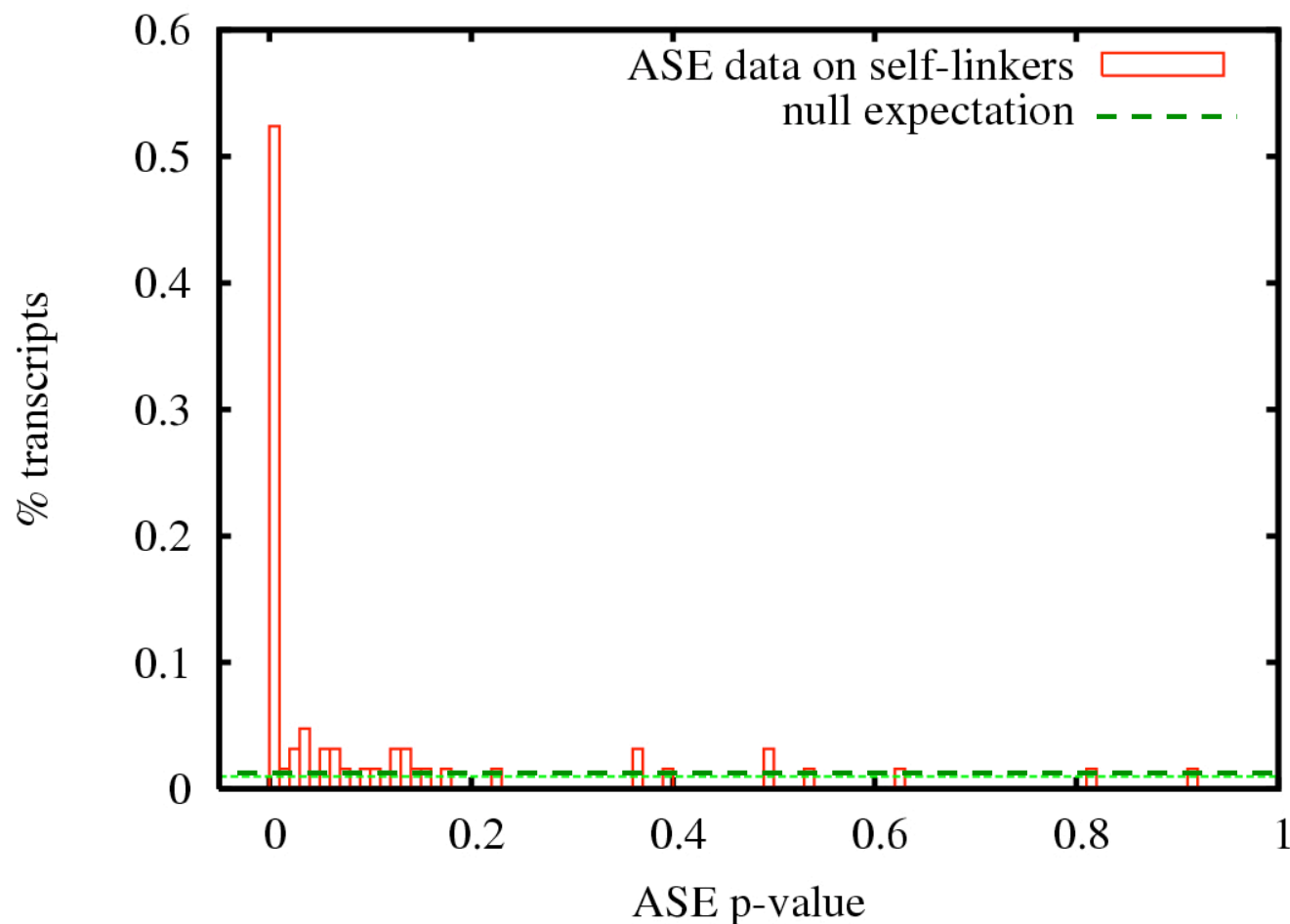
>50% of local variation is allele-specific



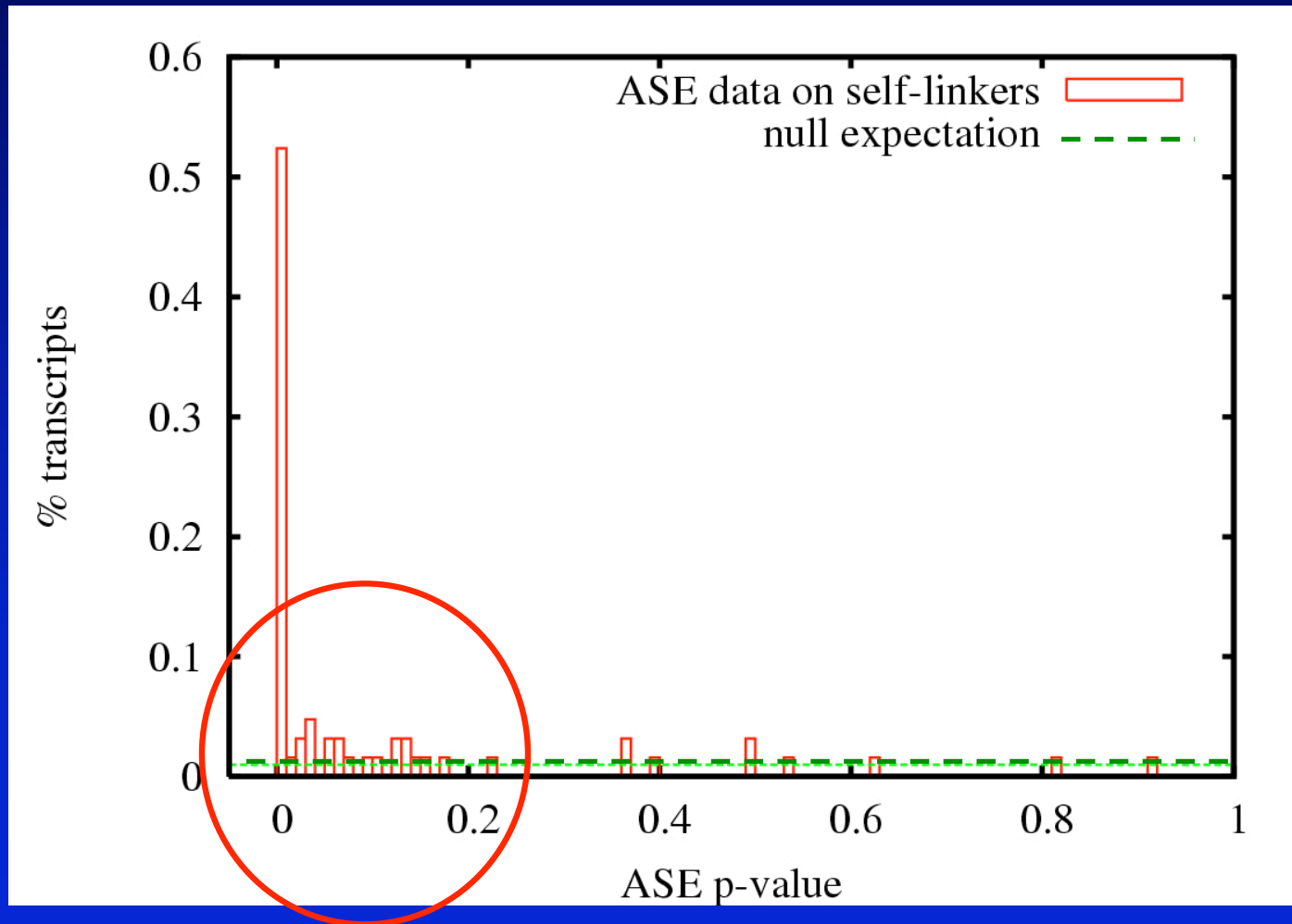
	Tested	ASE $p < 0.05$
Transcripts, local variation	77	44 (57%)
Transcripts, no local variation	16	2 (13%)

James Ronald, Jackie Whittle

True prevalence of allele-specific effect $\sim 80\%$

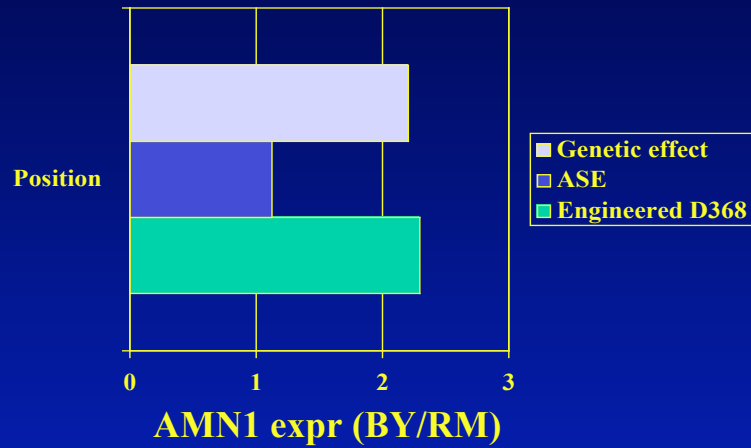


True prevalence of allele-specific effect ~80%

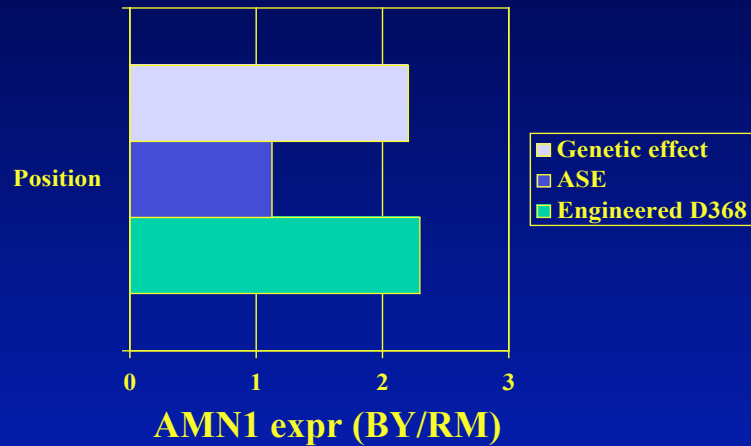


QVALUE
estimate:
78% of
tests are
true
positives

AMN1: local variation through feedback



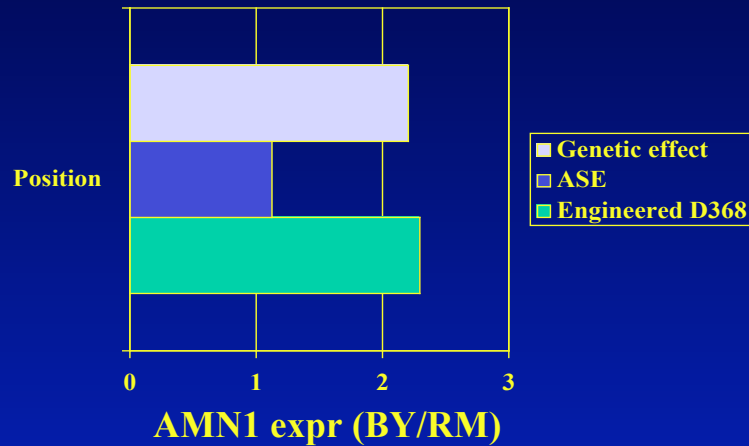
AMN1: local variation through feedback



	351	370
	DNFLLRLSQS IPNLKHL	DLR
	DNFLLRLSQS IPNLKHL	DLR
<i>S. paradoxus</i>		D
<i>S. mikatae</i>		D
<i>S. bayanus</i>		D
<i>S. servazzii</i>		D
<i>Z. rouxii</i>		D
<i>K. lactis</i>		D
<i>P. angusta</i>		D

AMN1

AMN1: local variation through feedback



351 370

DNFLLRLSQS IPNLKHL^DLR

DNFLLRLSQS IPNLKHL^DLR

S. paradoxus D

S. mikatae D

S. bayanus D

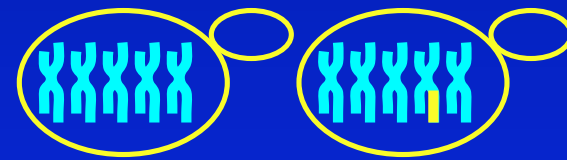
S. servazzii D

Z. rouxii D

K. lactis D

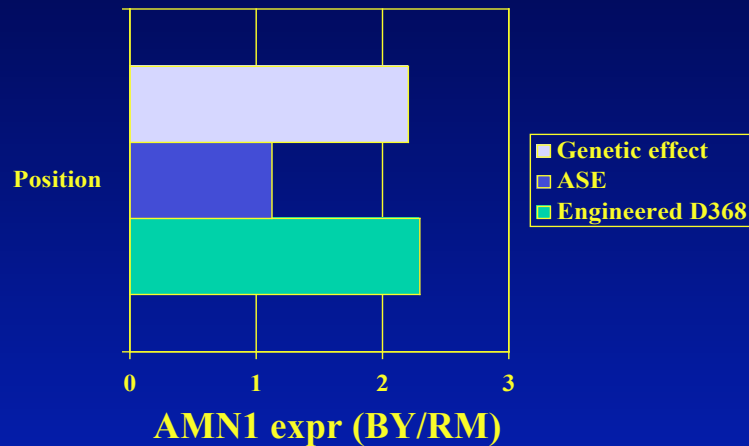
P. angusta D

AMN1



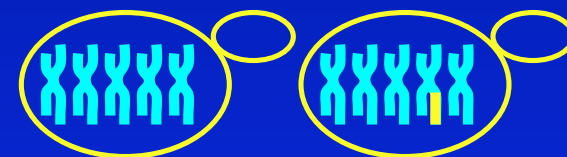
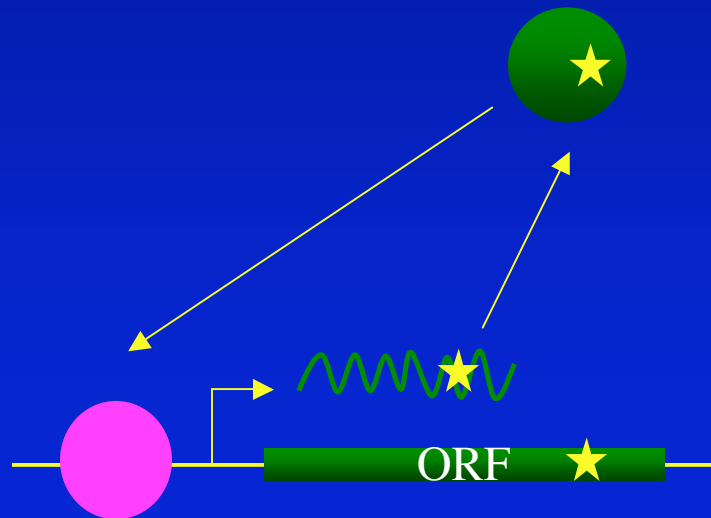
Introduced allele in
isogenic background

AMN1: local variation through feedback



	351		370
	DNFLLRLSQS	IPNLKHL	DLR
	DNFLLRLSQS	IPNLKHL	DLR
<i>S. paradoxus</i>			D
<i>S. mikatae</i>			D
<i>S. bayanus</i>			D
<i>S. servazzii</i>			D
<i>Z. rouxii</i>			D
<i>K. lactis</i>			D
<i>P. angusta</i>			D

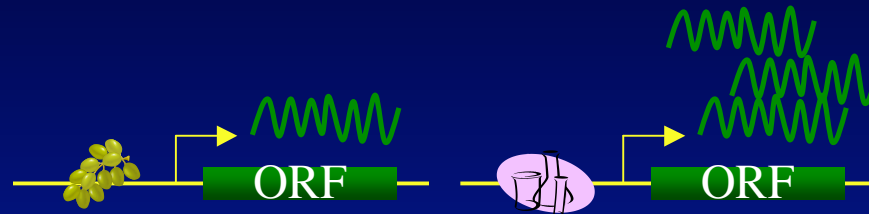
AMN1



Introduced allele in isogenic background

Sequence analysis of local variation

Assemble genes with local variation



Align parent sequence

	-20		30
RM	accgaatata	--cacacata	ATGGCATCAACCGATTTCTCCAAGATTGAA
BY	aacgaatata	ttcacacata	ATGGCATCCACCGATTTCTCCAAGAATGAA

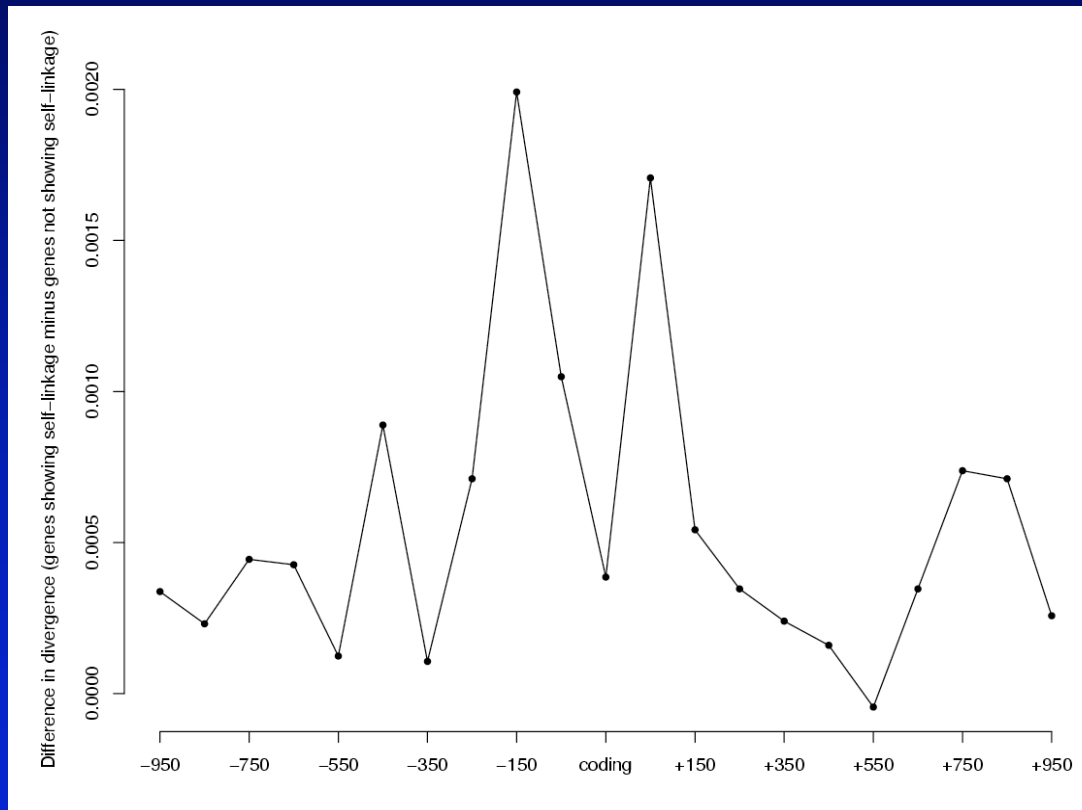
Count polymorphisms

	-20		30
RM	accgaatata	--cacacata	ATGGCATCAACCGATTTCTCCAAGATTGAA
BY	aacgaatata	ttcacacata	ATGGCATCCACCGATTTCTCCAAGAATGAA

Find spatial distribution

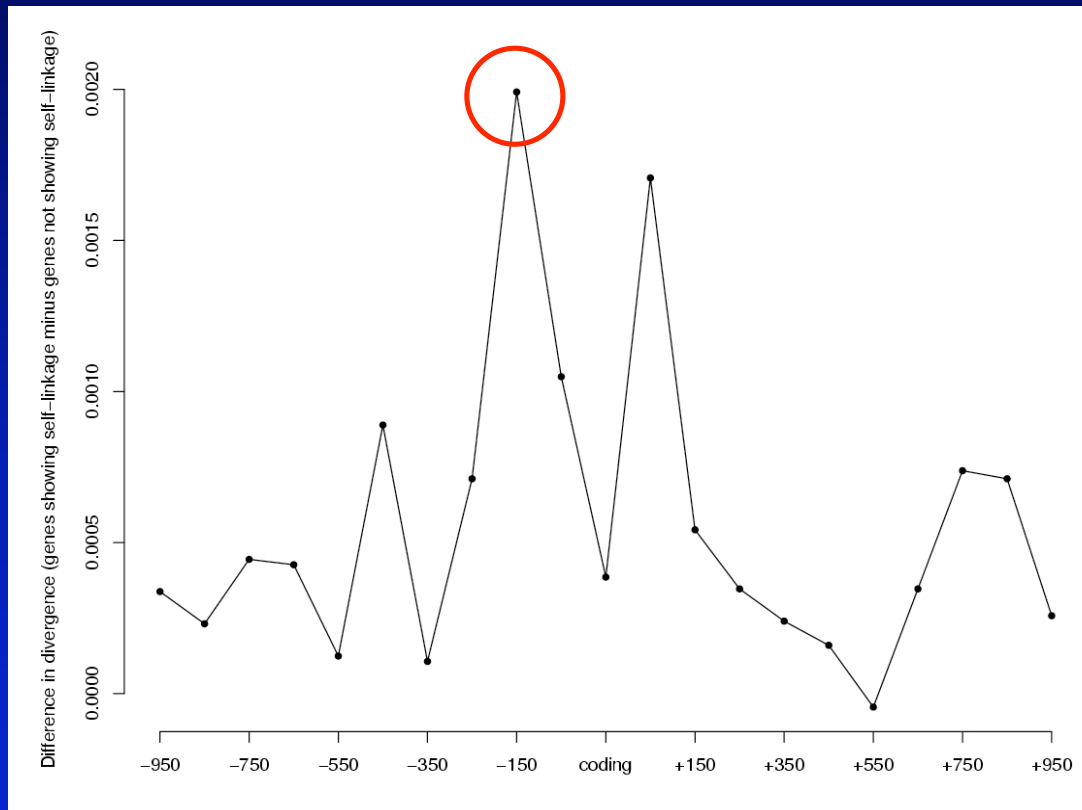
	-20				30
RM	accgaatata	--cacacata	ATGGCATCA	ACCGATTTCTCCAAGATTGAA	
BY	aacgaatata	ttcacacata	ATGGCATCC	ACCGATTTCTCCAAGAATGAA	

SNP sites, transcripts under local variation



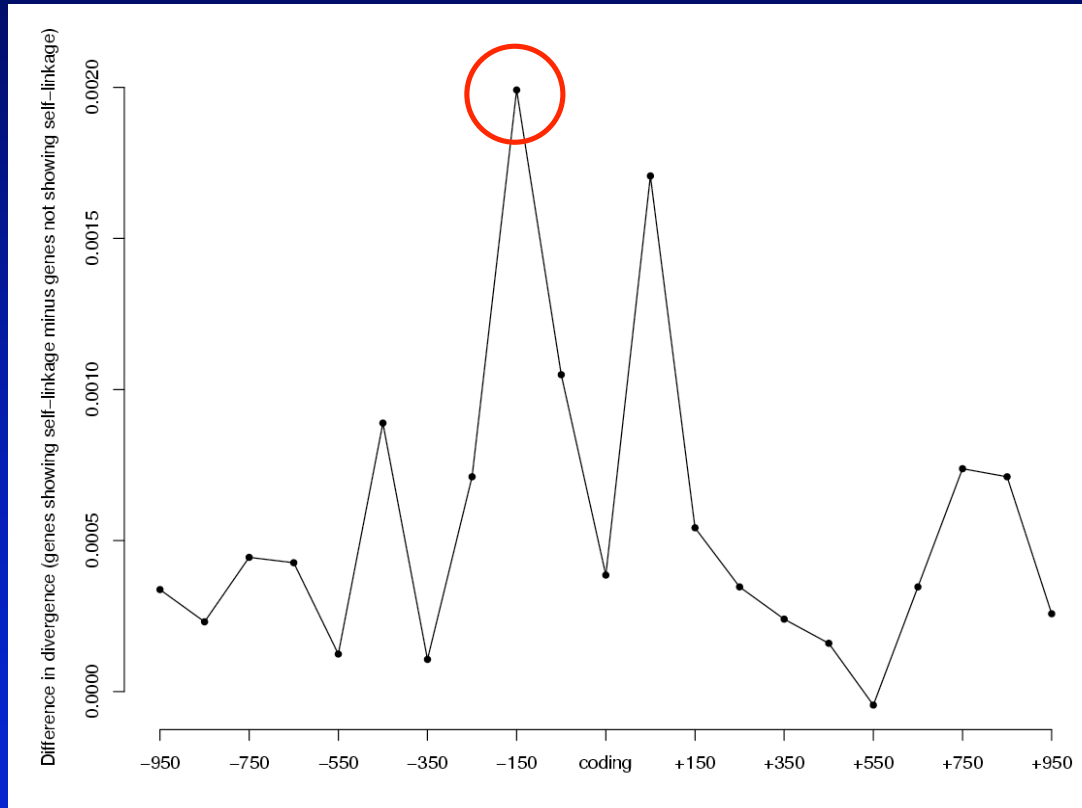
James Ronald

SNP sites, transcripts under local variation



James Ronald

SNP sites, transcripts under local variation

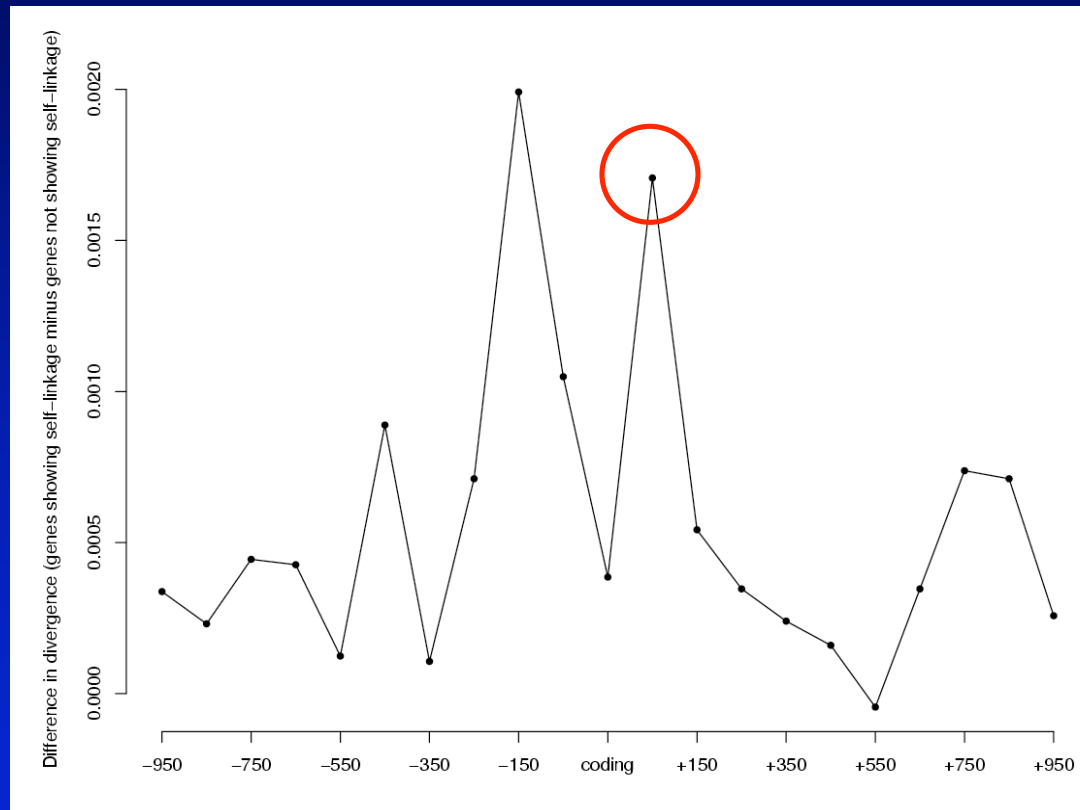


	TF sites	non-TF sites
genes, local variation	121 SNPs	240 SNPs
genes, no local variation	20 SNPs	1493 SNPs

OR = 3.73
 $p < 0.0005$

James Ronald

SNP sites, transcripts under local variation

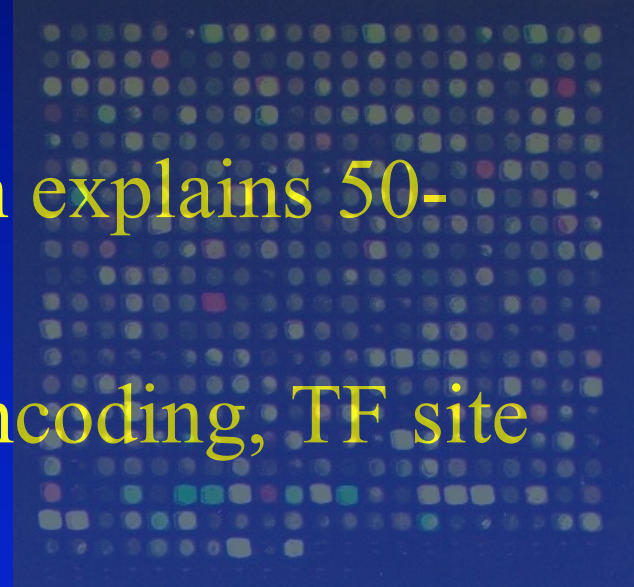


Mutations in 3' UTR
control transcript levels?

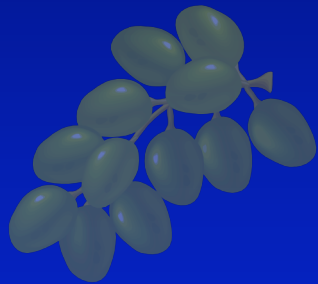
James Ronald

Conclusions

- Local regulatory variation affects 25-50% of yeast transcripts
- Allele-specific mechanism explains 50-80% of these phenotypes
- Evidence strongest for noncoding, TF site variation



Acknowledgements



James Ronald

Jackie Whittle

Leonid Kruglyak

