Genetics of phenotypic variation

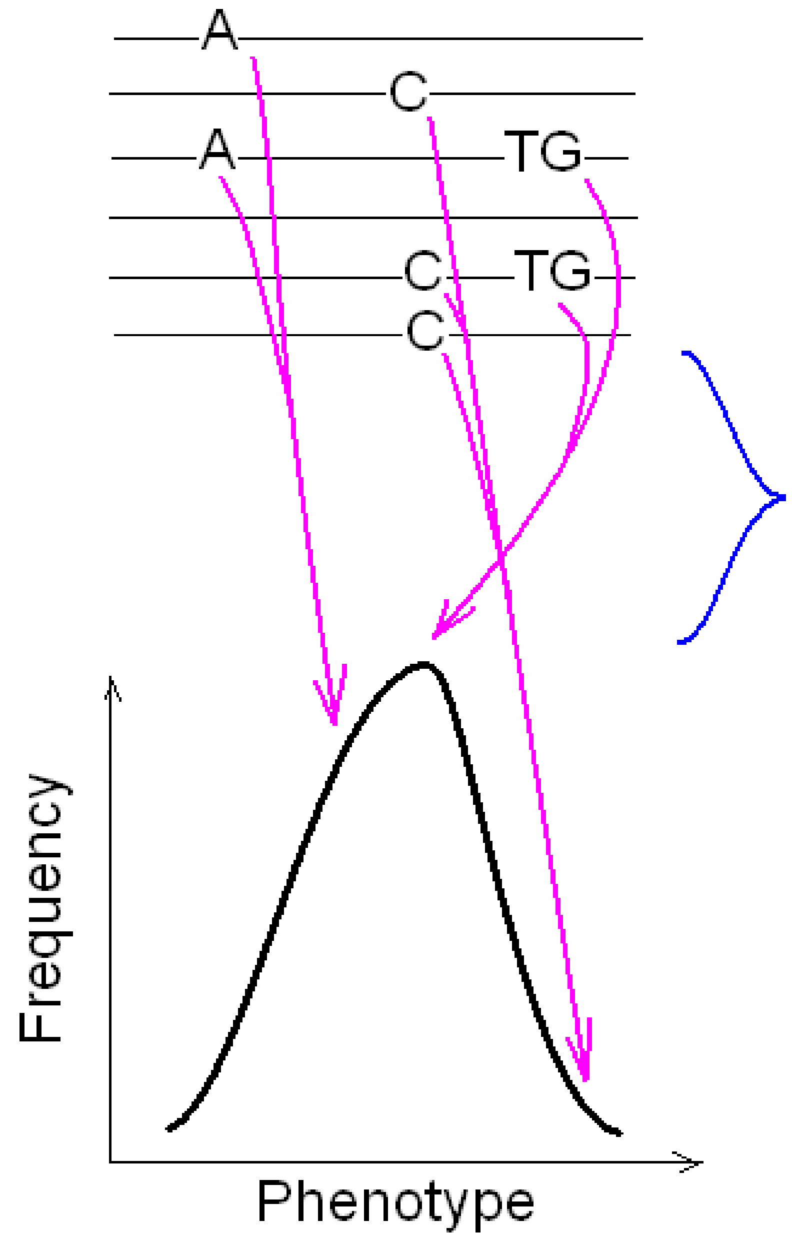
Tom Turner & Eric von Wettberg
USC&UCD
Peter Chang & Brian Dilkes
USC UCD
Ryan Bickel & Artyom Kopp
USC UCD
Lauren McIntyre, Marta Wayne & Larry Harshman
UFG UNL

Intro

- Genetics of local adaptation of *A. lyrata*, QTL → model selection;
- Is outcrosser (*A. arenosa*) better than selfer (*A. thaliana*)?

- Gene-level association studies, *bric-a-brack* in *D. melanogaster*?
- How variation in pathways scales into phenotypic variation, InR/TOR pathway in *D. melanogaster*.

Intro

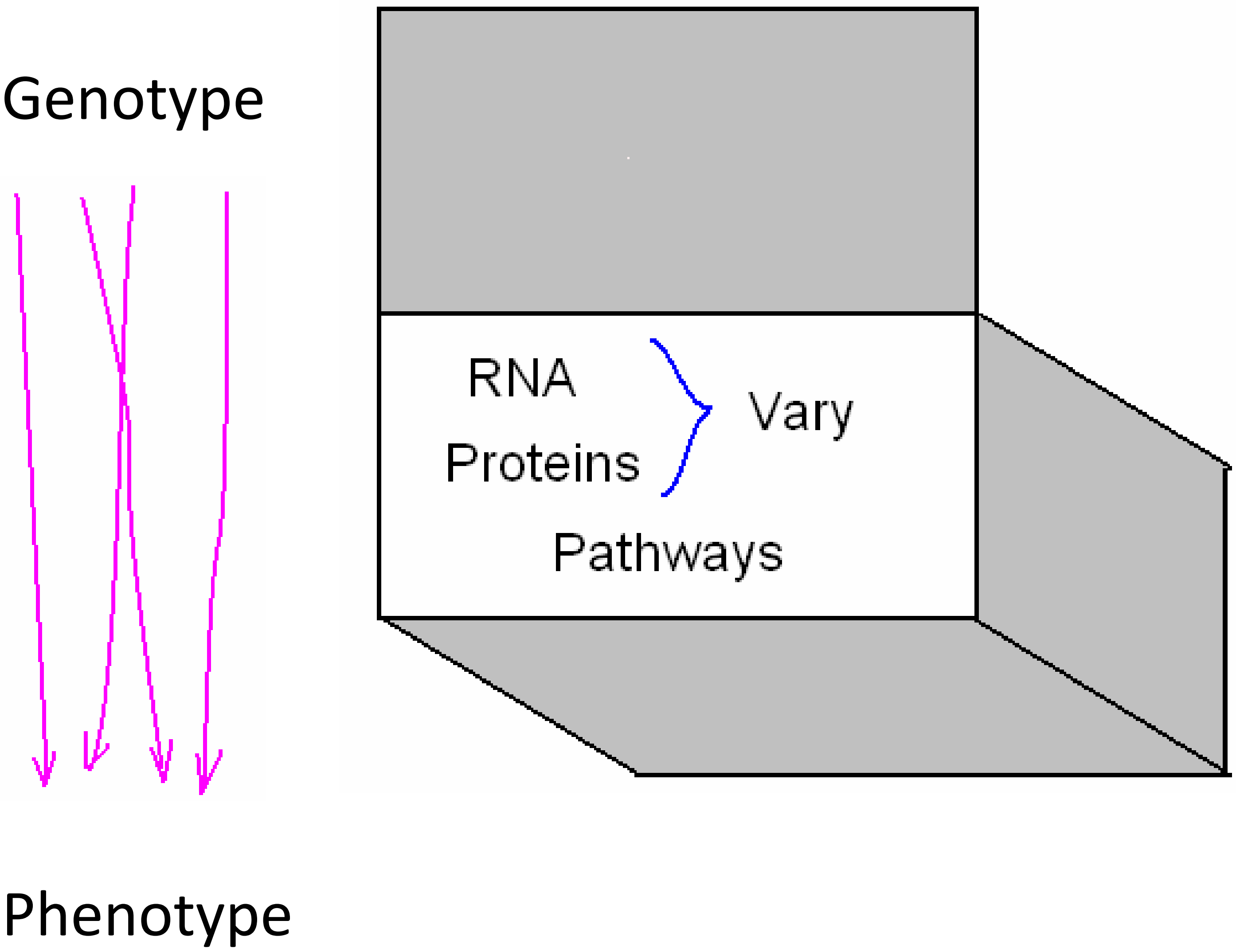


Molecular Evolution

Quantitative
Genetics

Evolution of Phenotype

Intro



Genotype

Phenotype

RNA

Proteins

Vary

Pathways

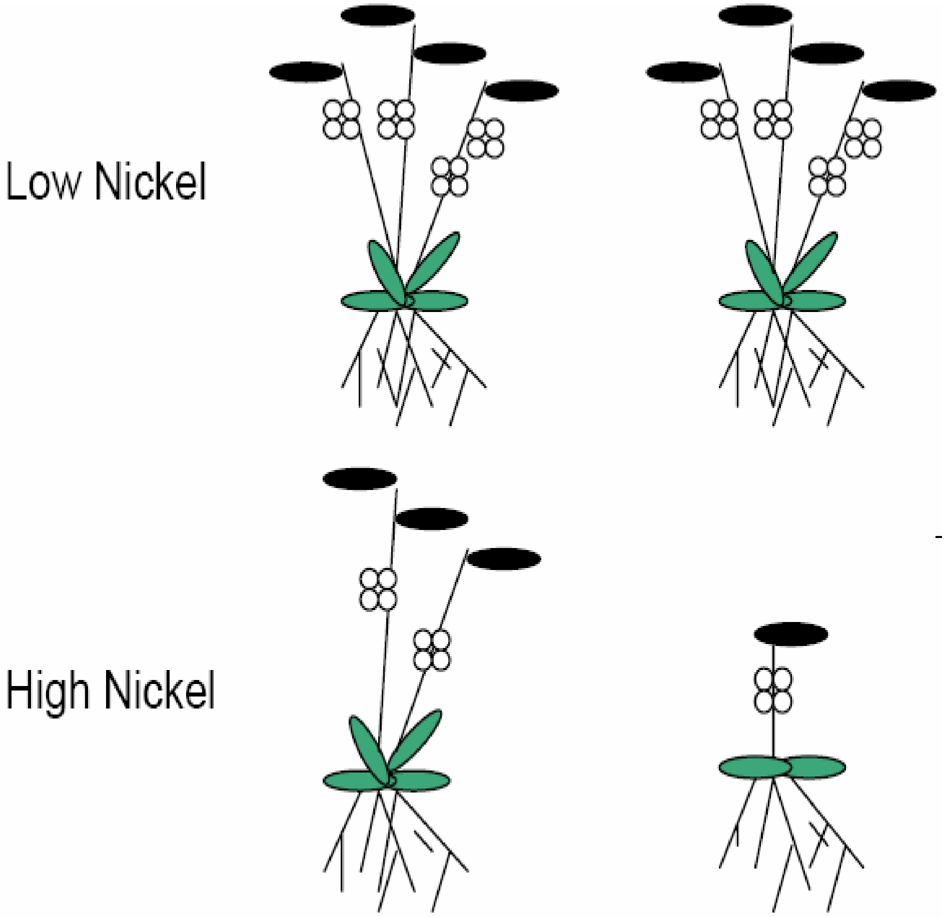
Intro

- Genetics of local adaptation of *A. lyrata*, QTL → model selection;
- Is outcrosser (*A. arenosa*) better than selfer (*A. thaliana*)?

- Gene-level association studies, *bric-a-brack* in *D. melanogaster*?
- How variation in pathways scales into phenotypic variation, InR/TOR pathway in *D. melanogaster*.

Arabidopsis lyrata adaptation to serpentine soils, how many genes?

Sequence divergence, how much does it matter?



Arabidopsis lyrata

~95% seq. identity

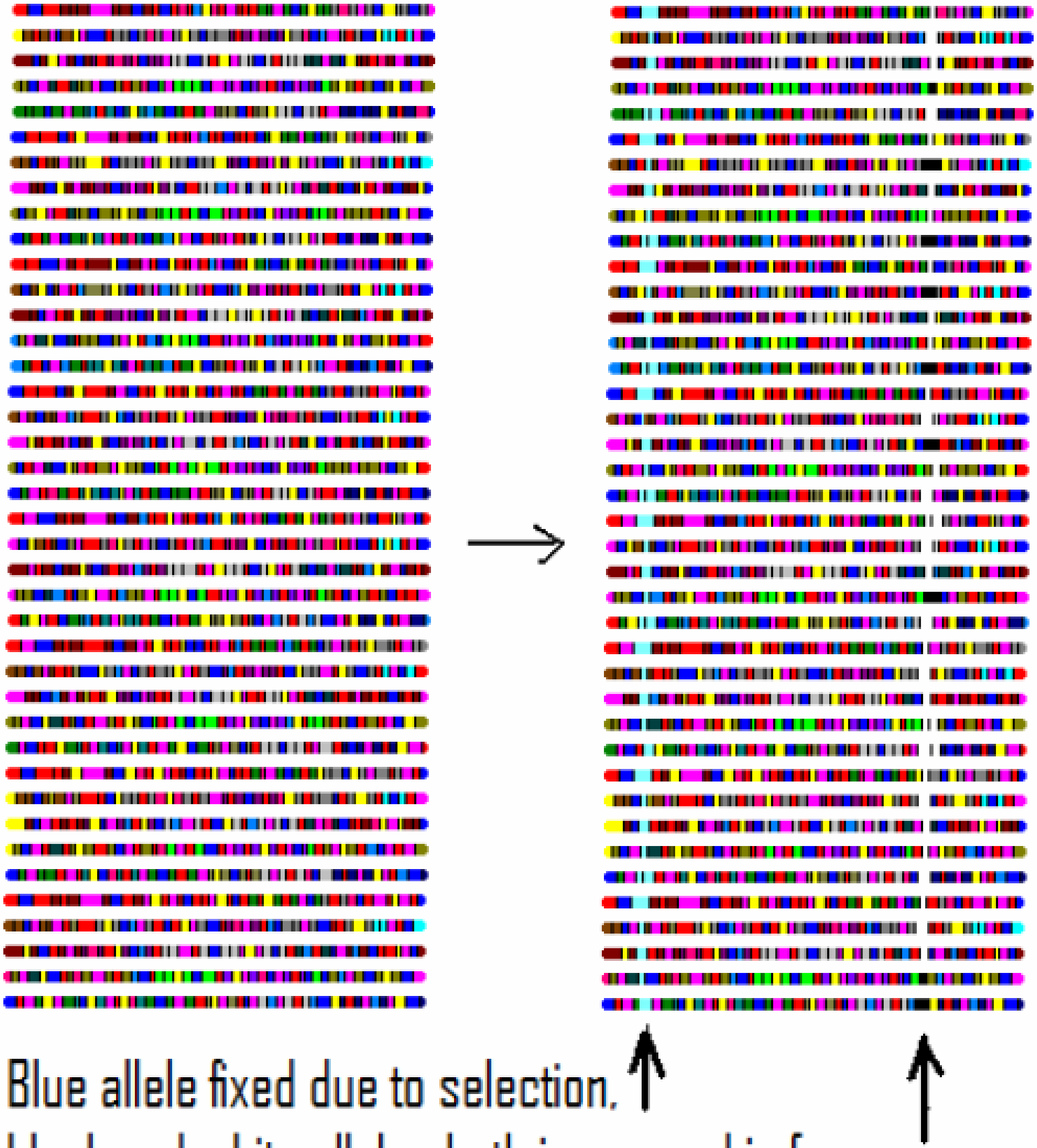
~tiling Arabidopsis array:

~3M PM/MM features

2 populations on normal

2 on serpentine X 3 chips

→ 12 arrays



Blue allele fixed due to selection,
black and white alleles both increased in frequency

Soil adaptation: steps

95% sequence identity →

→ 1 b out of 25 are divergent

2853369 PMs on the array →

→ 371642 are matches

28 (p<0.05 Bonferroni);

72 (FDR<1%).

+ windows → a couple of hundreds significantly diverged genes!

Reasonable? Overrepresented GO categories:

- Membrane proteins ($p<0.001$)
- Transporters ($p<0.001$)
- Potassium ion transport ($p=0.006$)
- Cellular calcium ion homeostasis ($p=0.008$)
- Cation transport ($p=0.03$).

cax7 → 8/23 (70b) are polys in granitic

/ fixed in serpentine soils.

→ At least a couple hundred genes are involved in adaptation to soil types.

- Genetics of local adaptation of *A. lyrata*, QTL → model selection;

- 200 genes are divergent → ☹
- 30000 – 200 are NOT diverged → ☺

(Solexa-resequenced to 60x, all holds up)

Imagine that we do RILs from 2 populations with KNOWN divergent genes →

Mapping exercises → Assigning effects and probabilities (priors: allele frequency differences, known gene functions etc.)

Intro

- Genetics of local adaptation of *A. lyrata*, QTL → model selection;
- Is outcrosser (*A. arenosa*) better than selfer (*A. thaliana*)?

- Gene-level association studies, *bric-a-brack* in *D. melanogaster*?
- How variation in pathways scales into phenotypic variation, InR/TOR pathway in *D. melanogaster*.

Is outcrosser (*A. arenosa*) better than selfer (*A. thaliana*)?

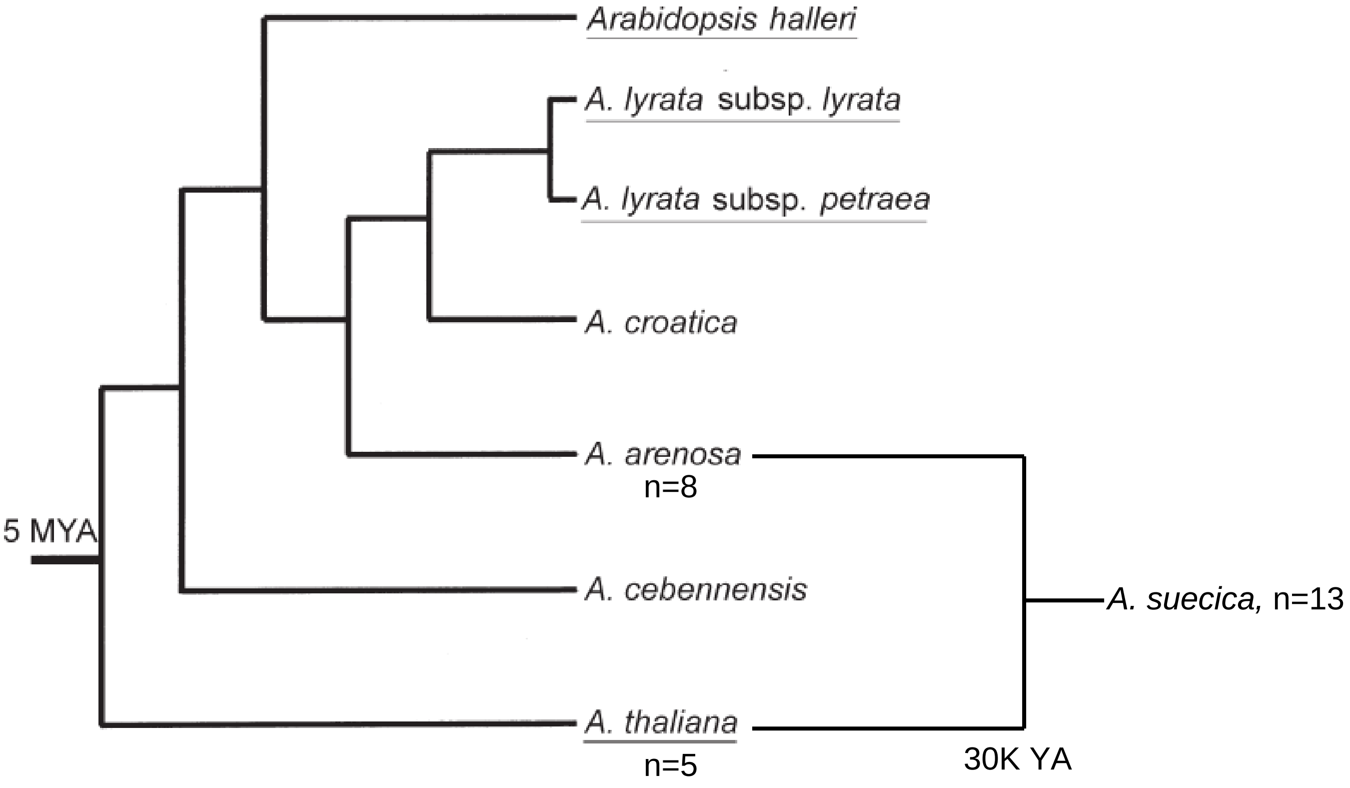
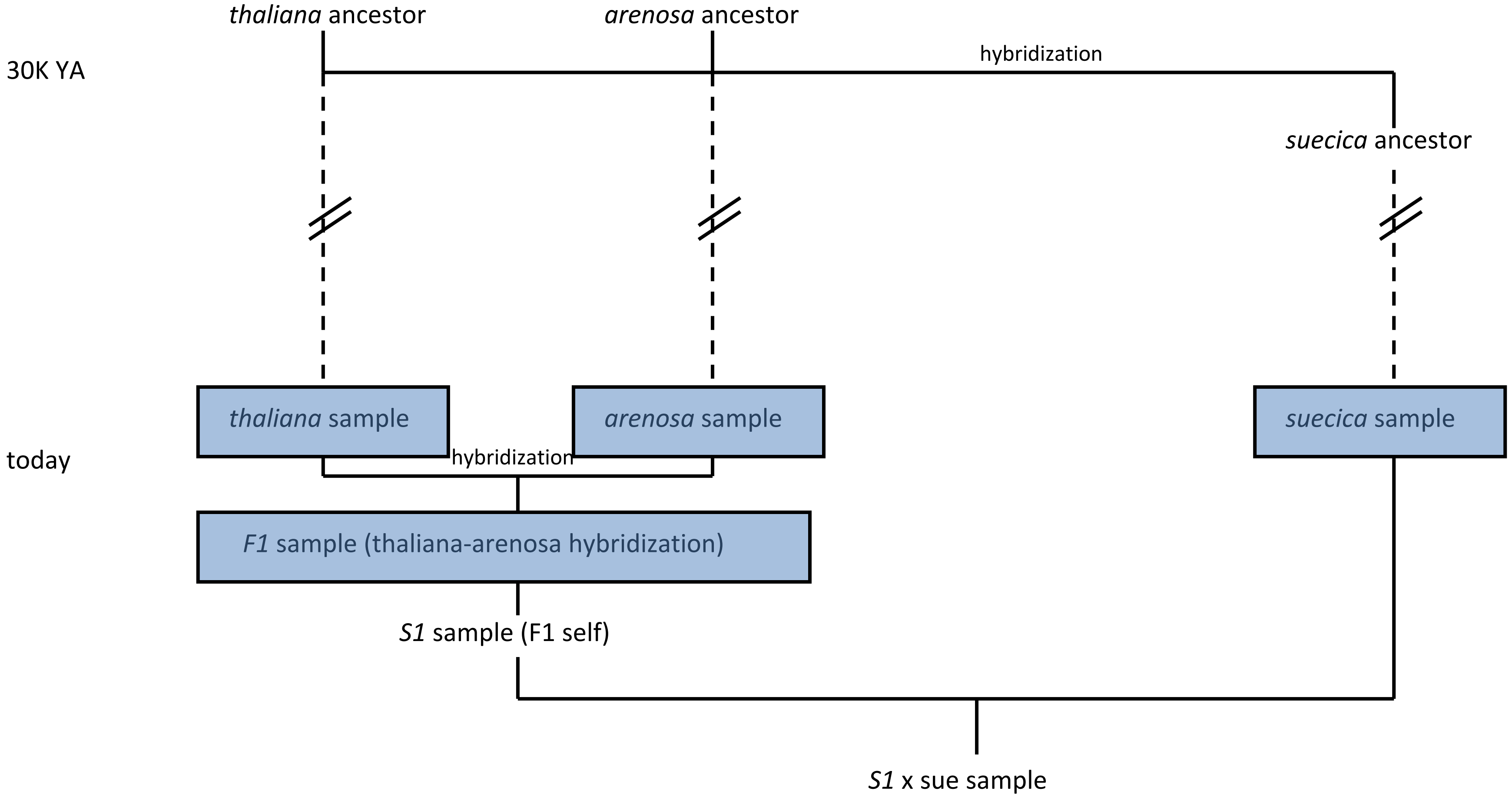


Figure 1. Phylogenetic relationships within *Arabidopsis*.

Fan et al, 2007



thaliana ancestor

arenosa ancestor

30K YA

hybridization

suecica ancestor

thaliana sample

arenosa sample

suecica sample

today

hybridization

F1 sample (thaliana-arenosa hybridization)

S1 sample (F1 self)

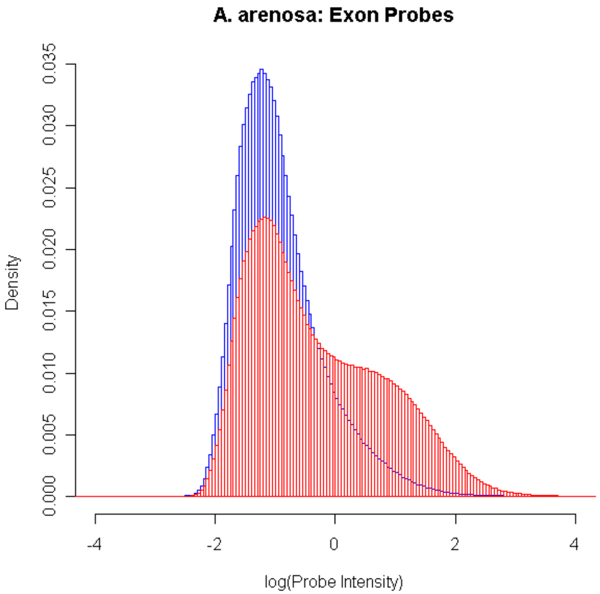
S1 x sue sample

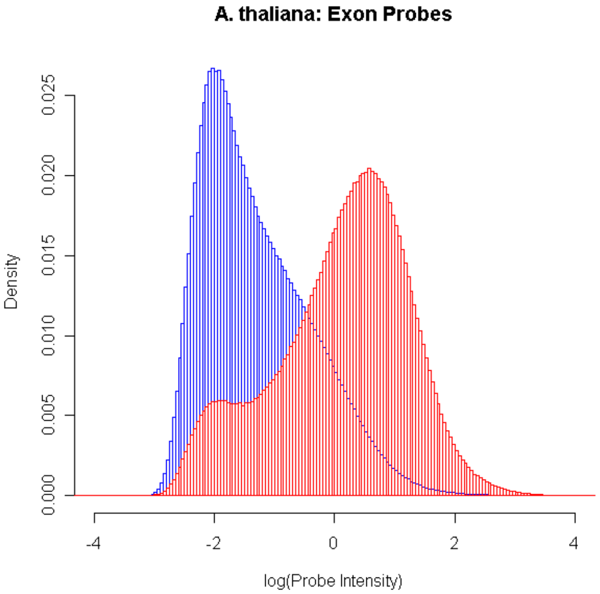
Question 1: Which allele has been maintained in *A. suecica*?

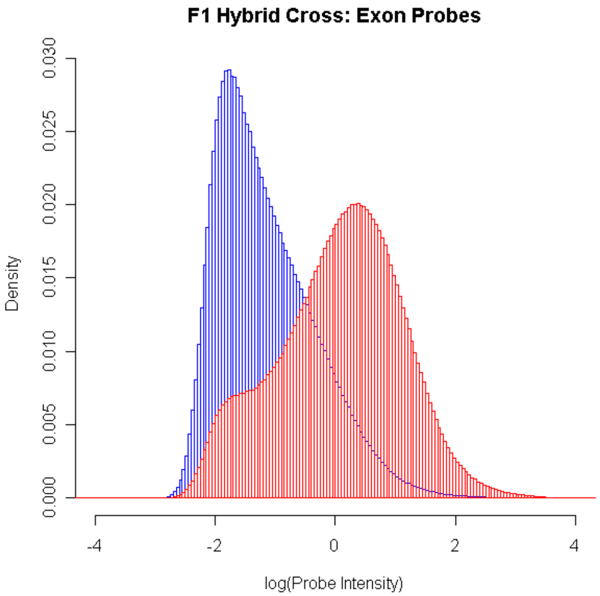
Question 2: Which allele has been silenced in *A. suecica*?

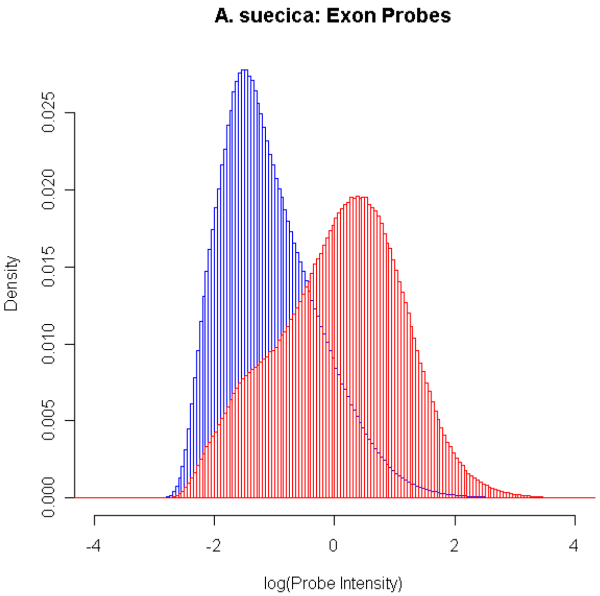
A. thaliana chips hybed with:

PM Probes
MM Probes

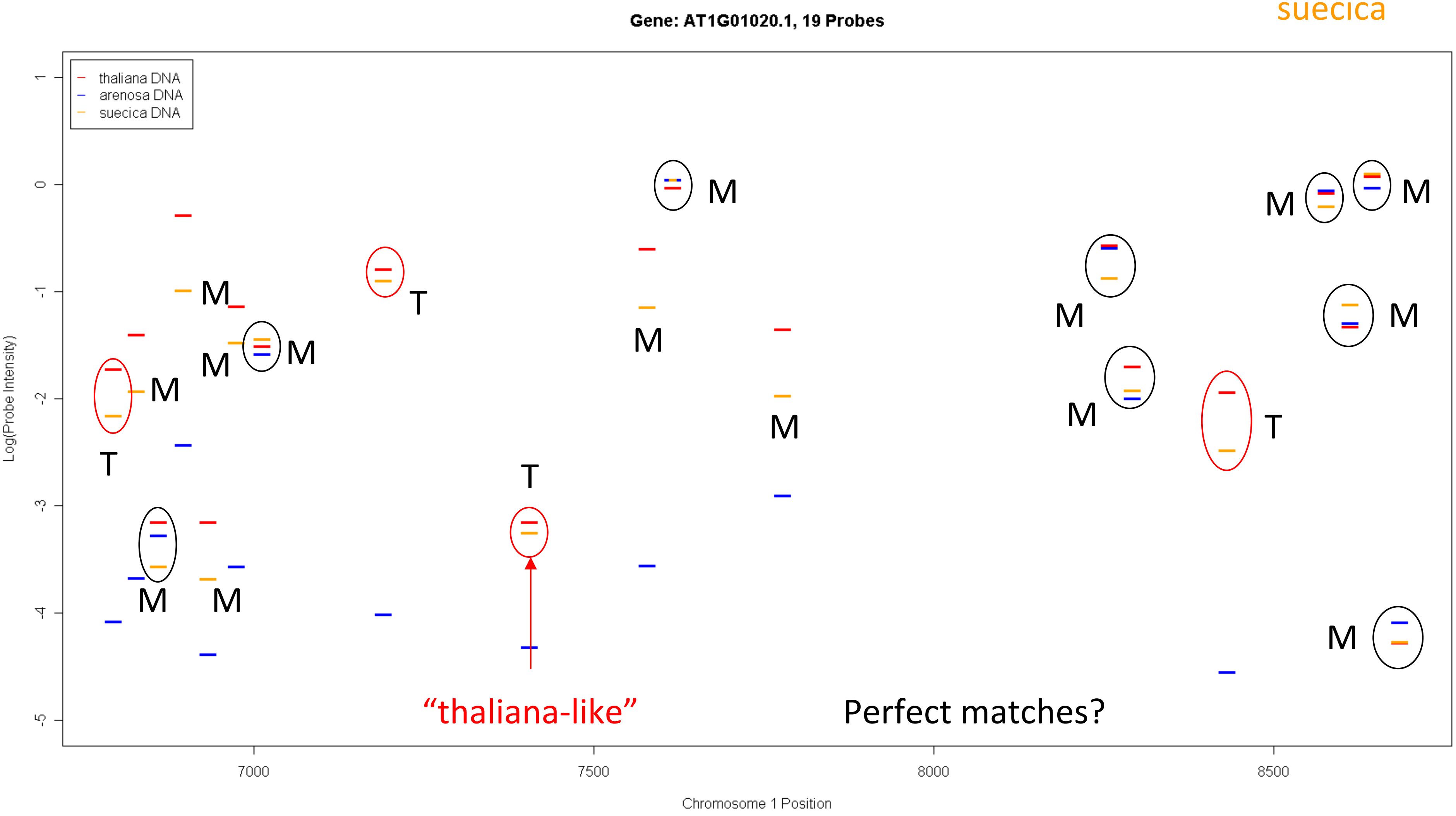

A. arenosa: Exon Probes


A. thaliana: Exon Probes


F1 Hybrid Cross: Exon Probes


A. suecica: Exon Probes

thaliana
arenosa
suecica



TMMMMMMTTMMMMMTMMMM → Probes Missing in arenosa: 4
Probes Missing in thaliana: 0

Probe Distribution

Arabidopsis Tiling Array At35b		Overall	
		Matches	Total
Number of Probes in Exons	F	601443	3039991
	R	611591	3039991
Number of Exons Covered	F	75448	78873
	R	76177	79197
Number of Genes Covered	F	15647	16018
	R	15519	15744
Average Number of probes/exon	**F**	**7.97**	
	R	**8.03**	
Average Number of exons/gene	**F**	**4.82**	
	R	**4.91**	

What is retained?

DNA	*A. suecica*	Genes
Deletions	*arenosa*-biased	180
	thaliana-biased	45
RNA	*arenosa*-biased	939
Silencing	*thaliana*-biased	383

Is outcrosser (*A. arenosa*) better than selfer (*A. thaliana*)?

In F1 → silencing is unbiased and ‘random’ (among progeny, tissues) Chen et al. 2006.

In *A. suecica* → *A. arenosa* alleles are:

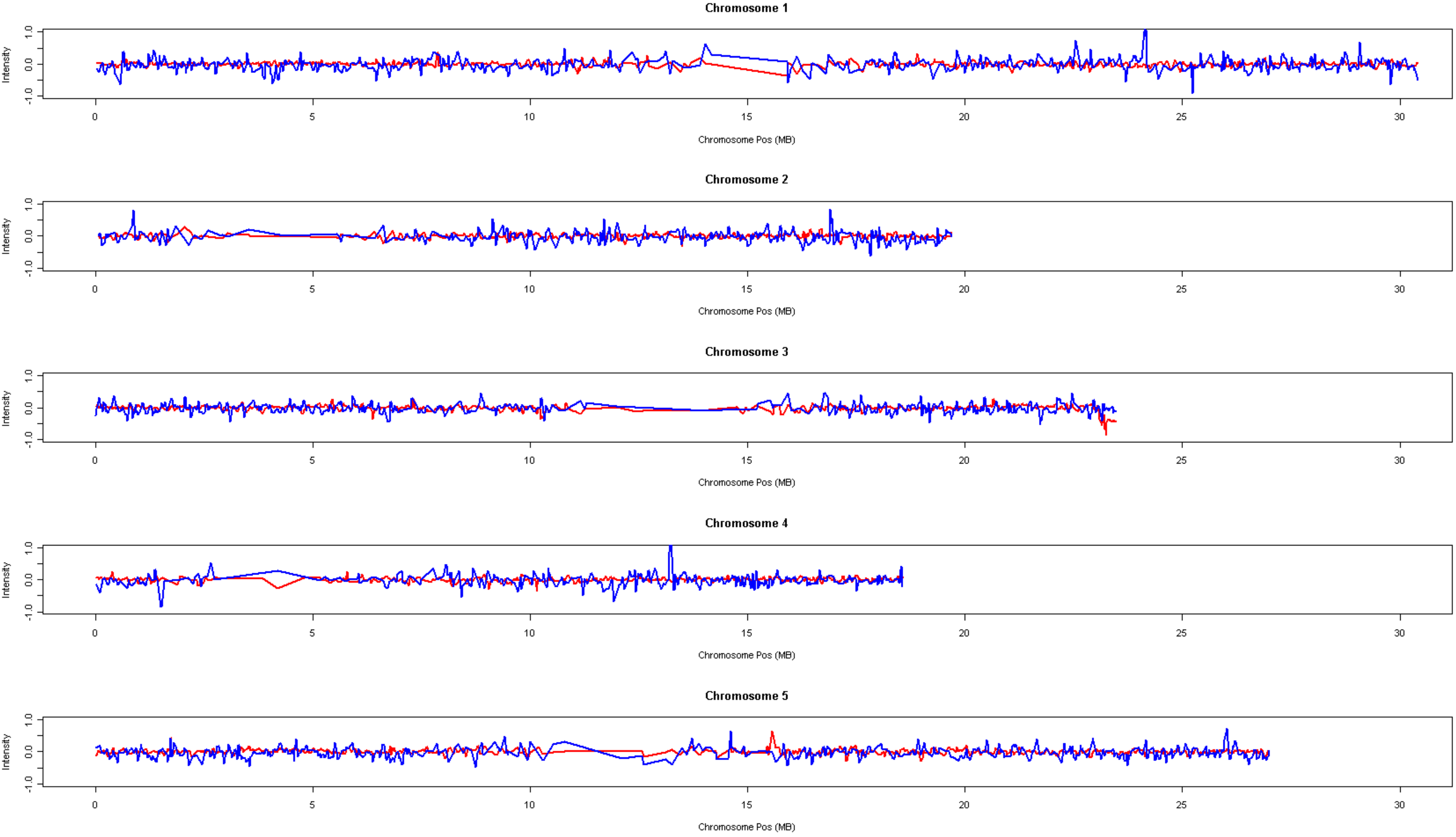
- more frequently retained;
- more frequently utilized.

Is silencing ‘region-specific’?

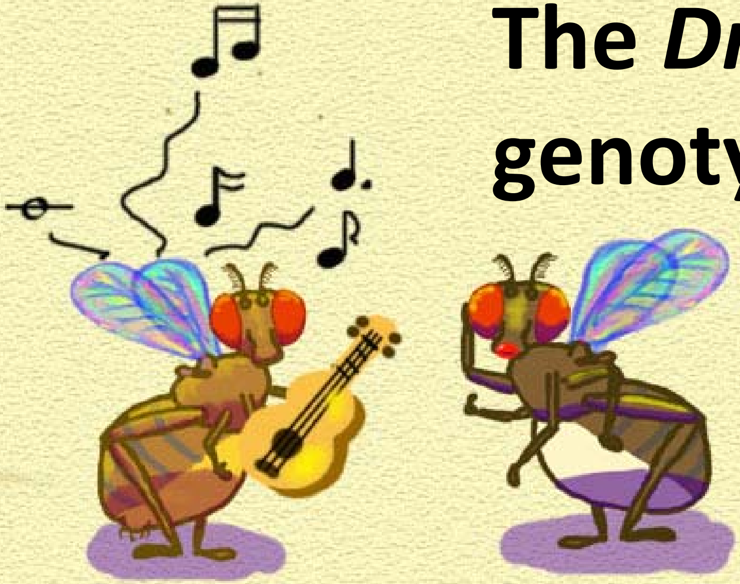
Is silencing ‘pathway-specific’?

Is silencing reflective of ‘population genetics’?

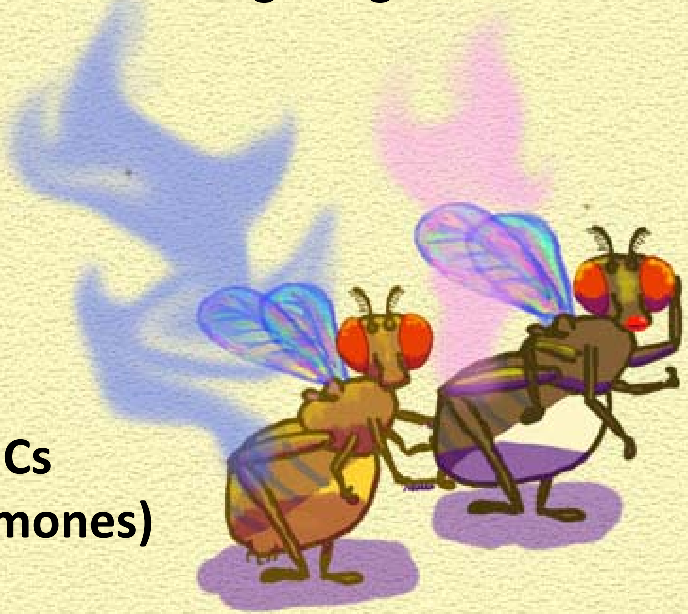
Is outcrosser (*A. arenosa*) better than selfer (*A. thaliana*)?



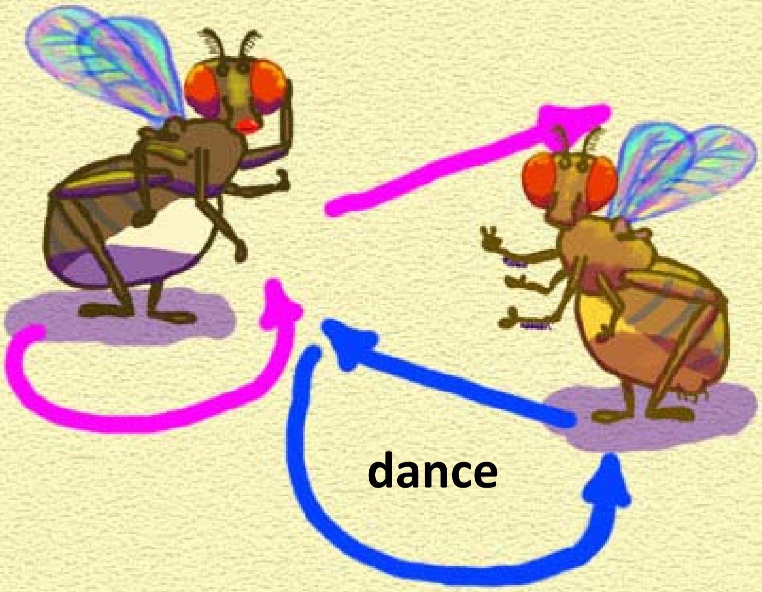
The *Drosophila* model for the genotype to phenotype map



wing song

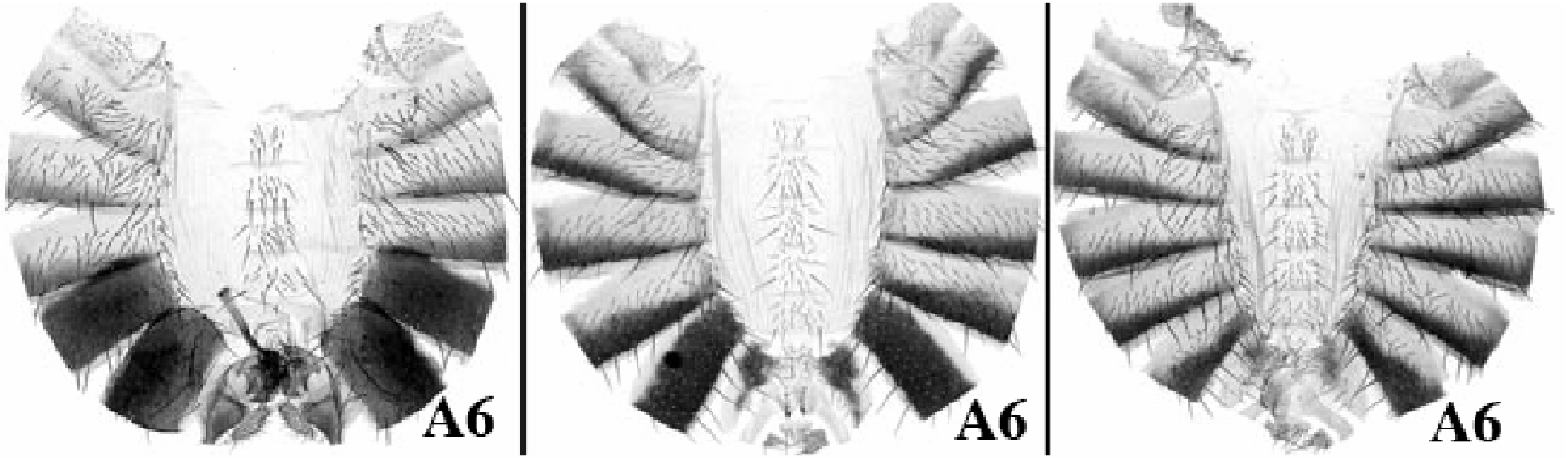


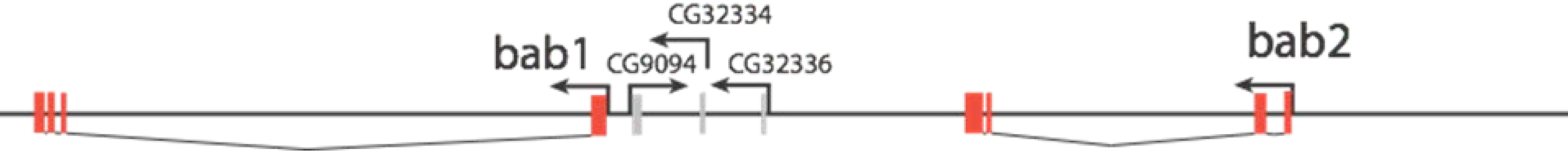
CHCs
(pheromones)



dance

Building a genotype to phenotype map

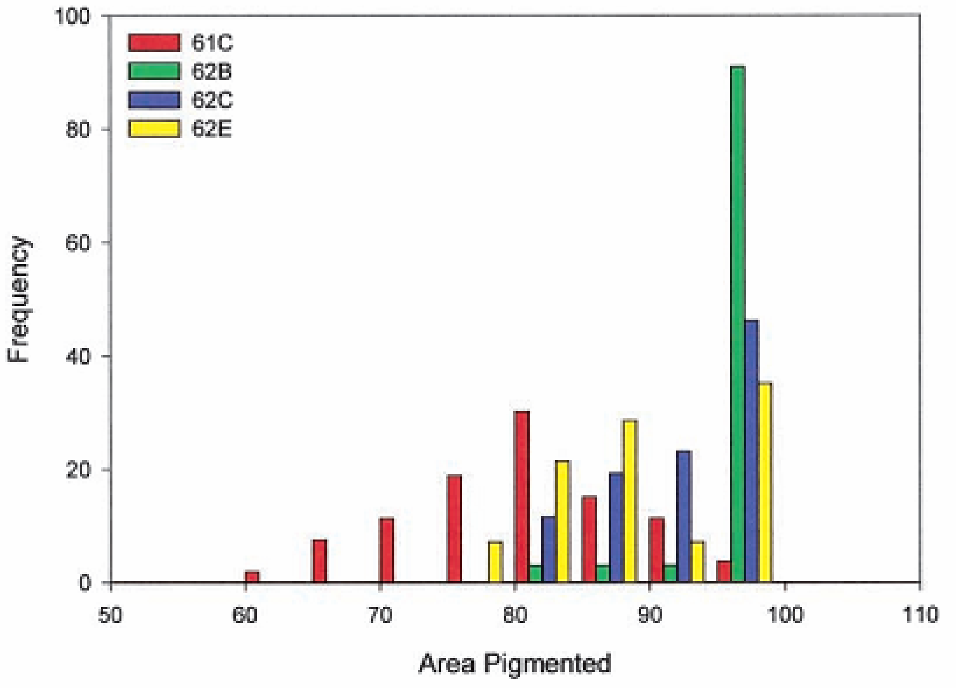


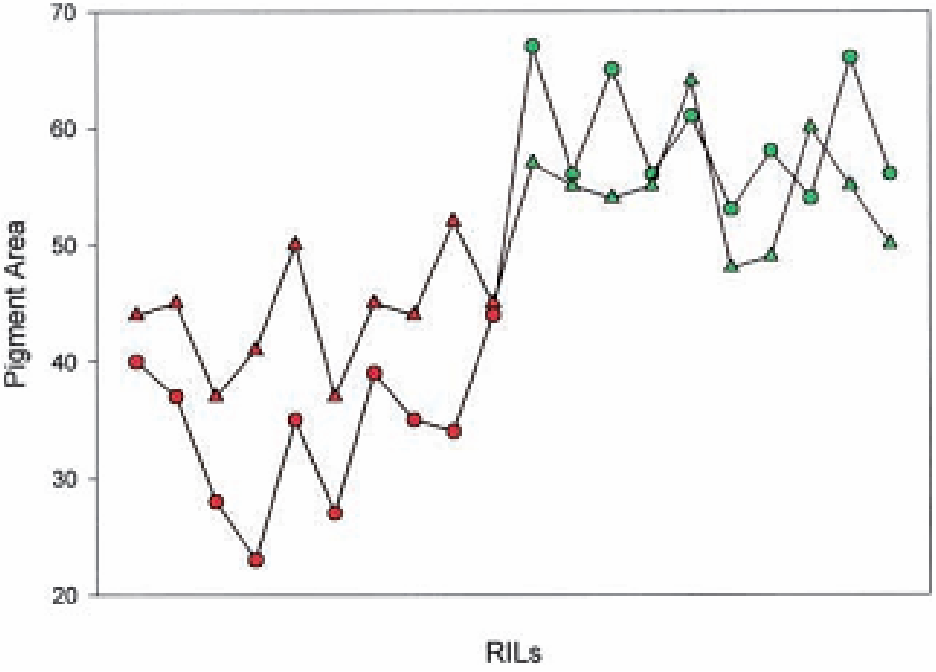


148 kb region → 40 SNPs / 6 indels per kb

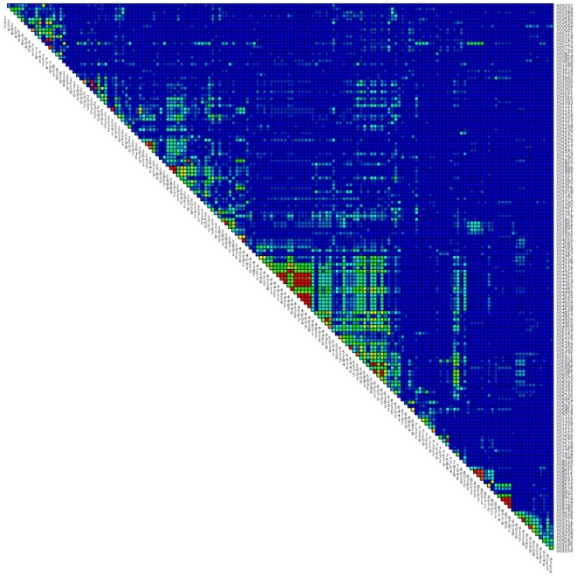
→ ~6000 SNPs and ~900 indels

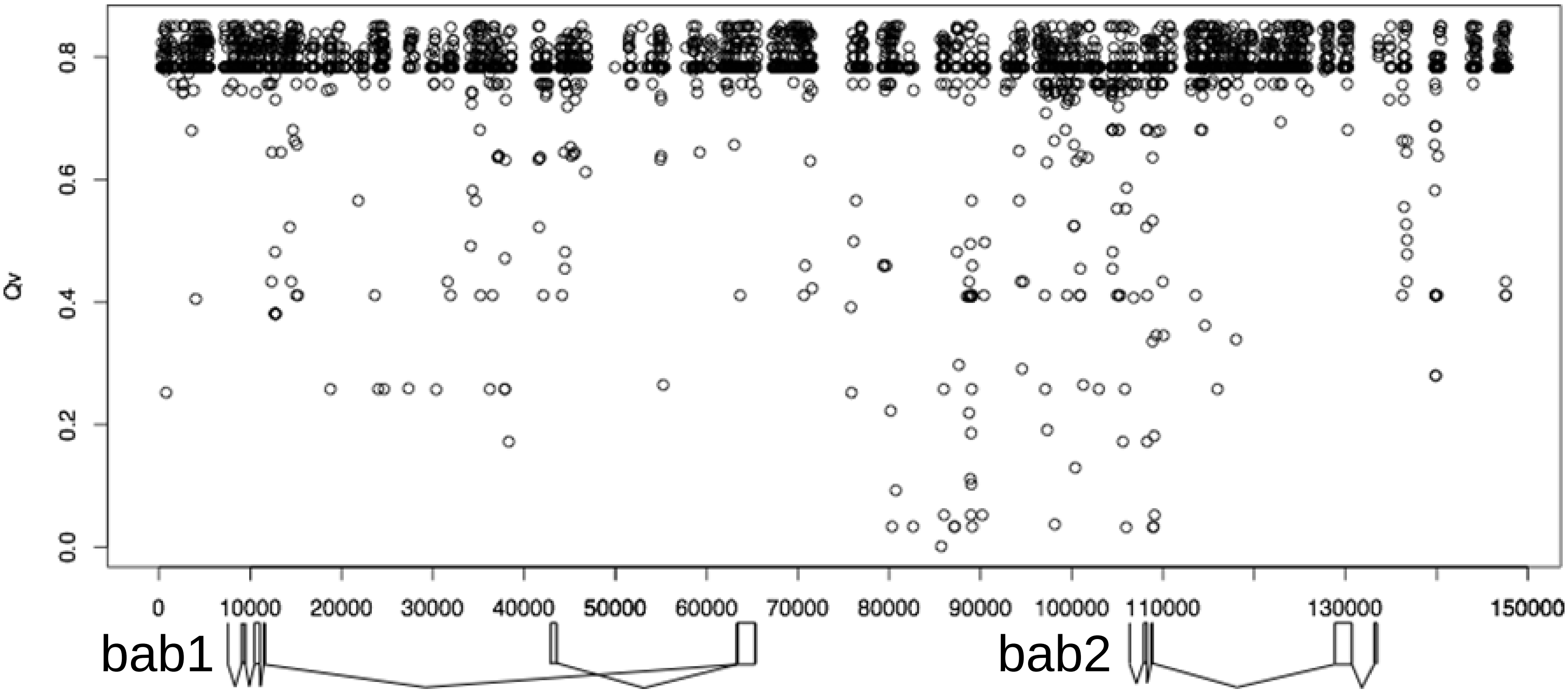
Effects of alleles are large, and the evidence for a gene is clear-cut

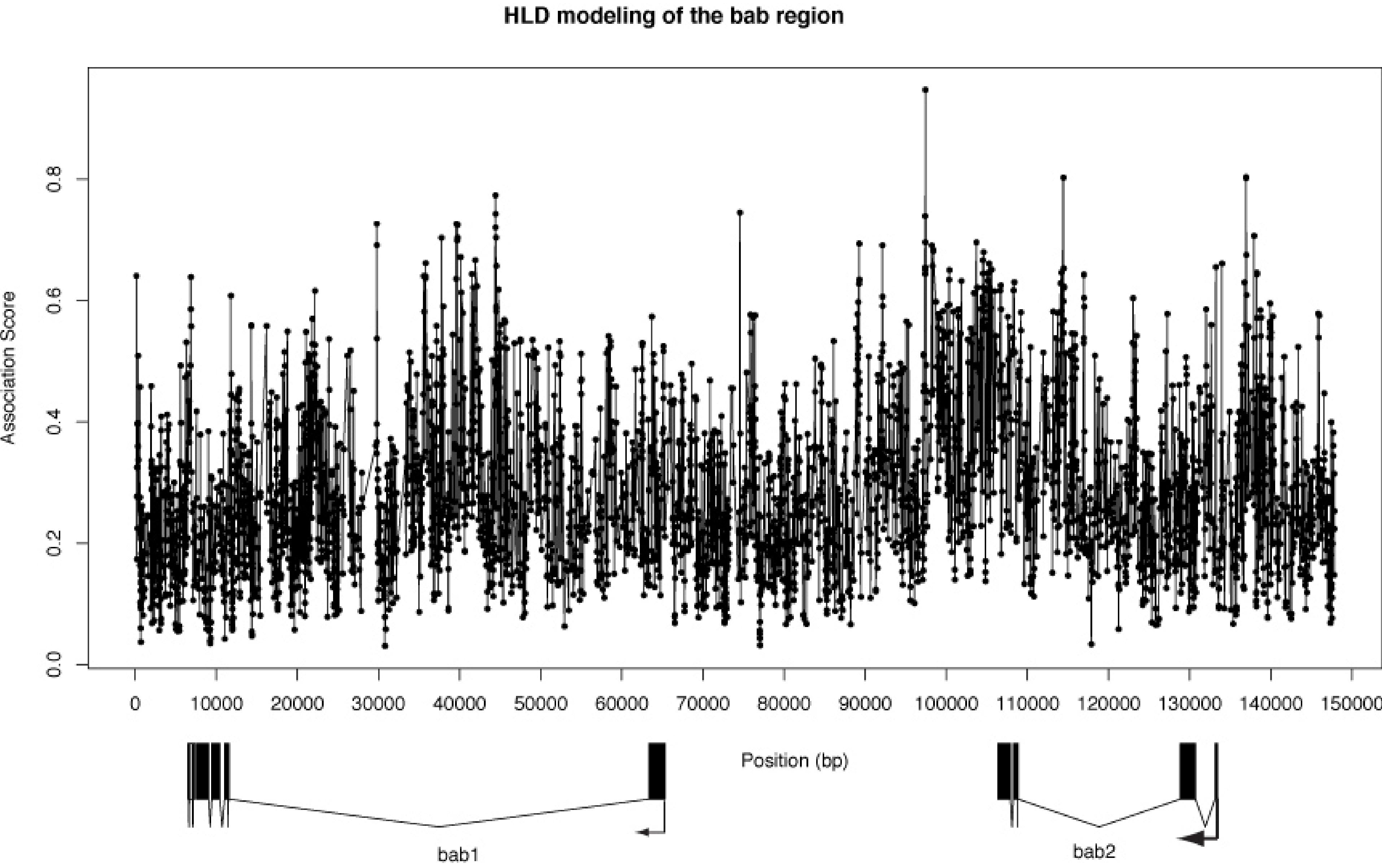




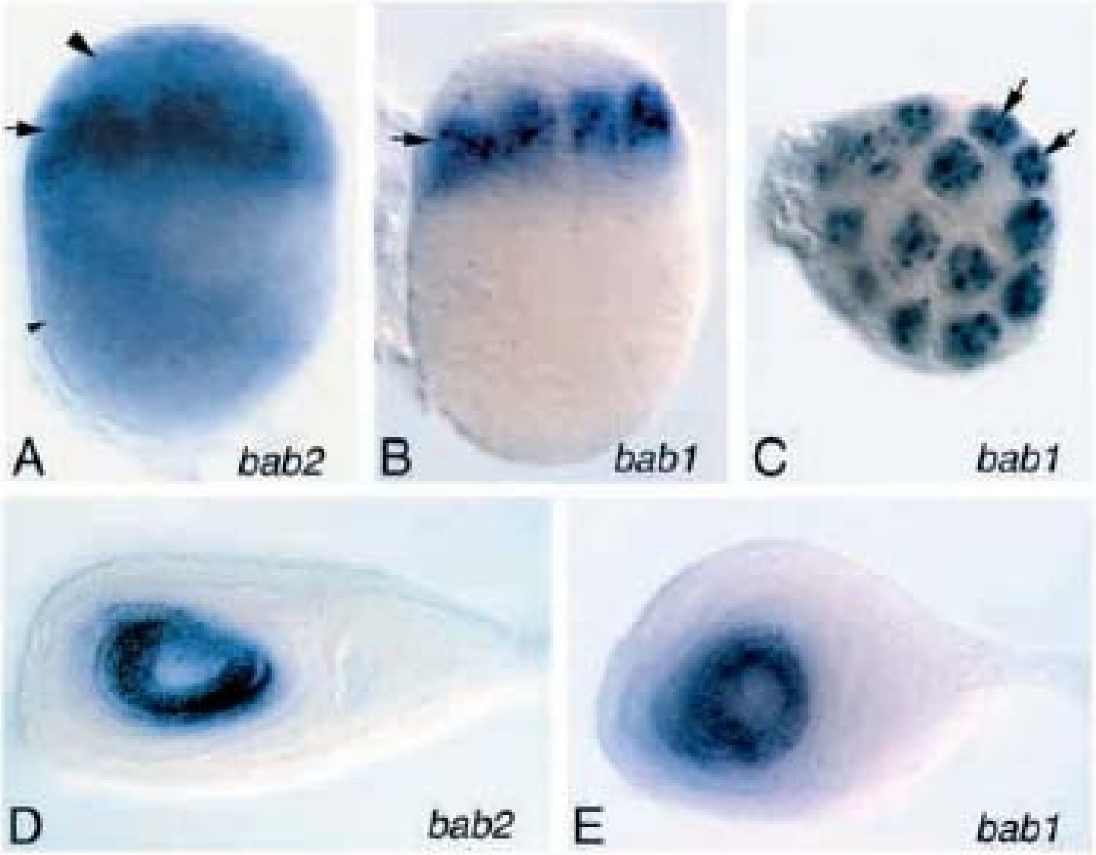
Low LD (200 b) enables precise association inferences

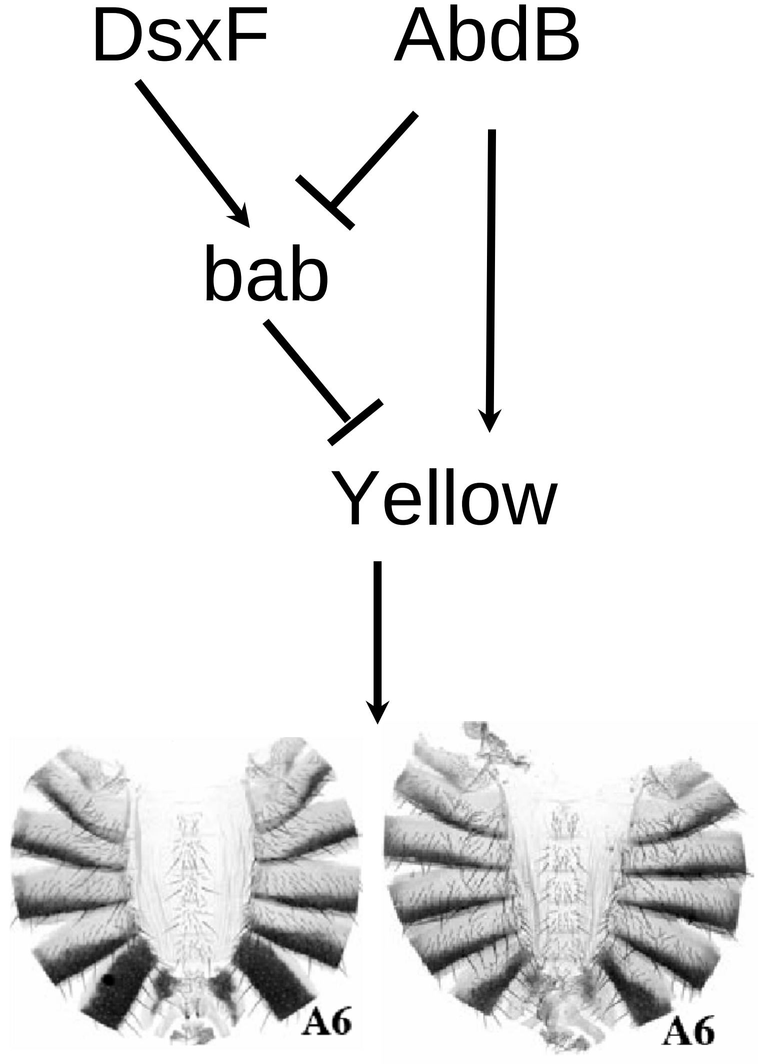





HLD modeling of the bab region
Association Score
0.0
0.2
0.4
0.6
0.8
0
10000
20000
30000
40000
50000
60000
70000
80000
90000
100000
110000
120000
130000
140000
150000
Position (bp)
bab1
bab2

Gene functioning might be assayed at the right time in the right tissues…





Intro

- Genetics of local adaptation of *A. lyrata*, QTL → model selection;
- Is outcrosser (*A. arenosa*) better than selfer (*A. thaliana*)?

- Gene-level association studies, *bric-a-brack* in *D. melanogaster*?
- How variation in pathways scales into phenotypic variation, InR/TOR pathway in *D. melanogaster*.

Quantitative genetics of transcript levels. Components of Variation.

- 9,863 genes ← no multiple splice products, probes higher than negative controls.

  1,609 *X* linked genes

  8,213 “autosomal” genes (*2, 3* only)

  35 *4th* chromosome, 6 *Y* chromosome

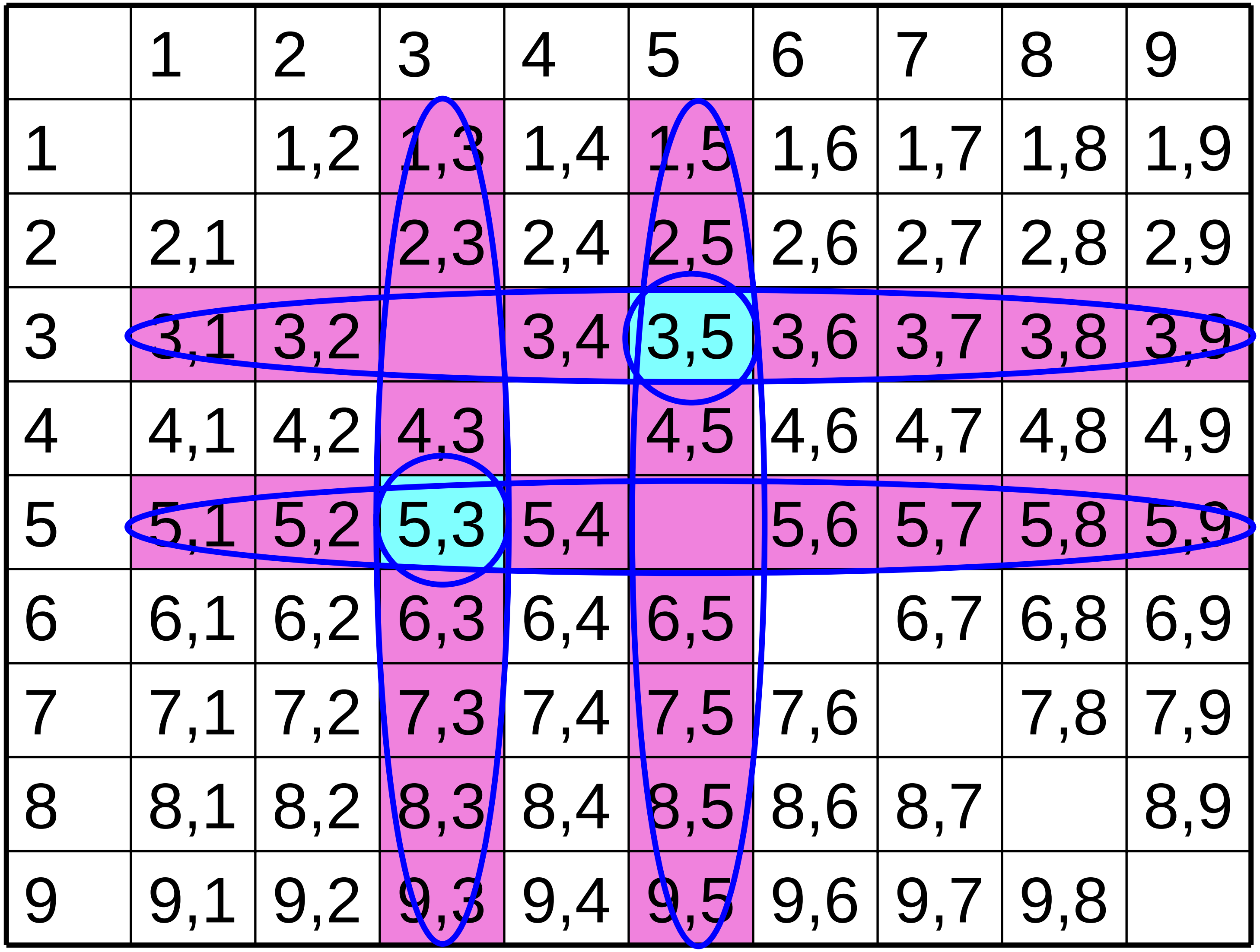
- Models run for sexes separately; evaluated significance at FDR = 0.20

$$\sigma^2_P = 2\sigma^2_{GCA} + \sigma^2_{SCA} + 2\sigma^2_{RGCA} + \sigma^2_{RSCA} + \sigma^2_{\varepsilon}$$

Quantitative genetics of transcript levels.

Sires

Dams



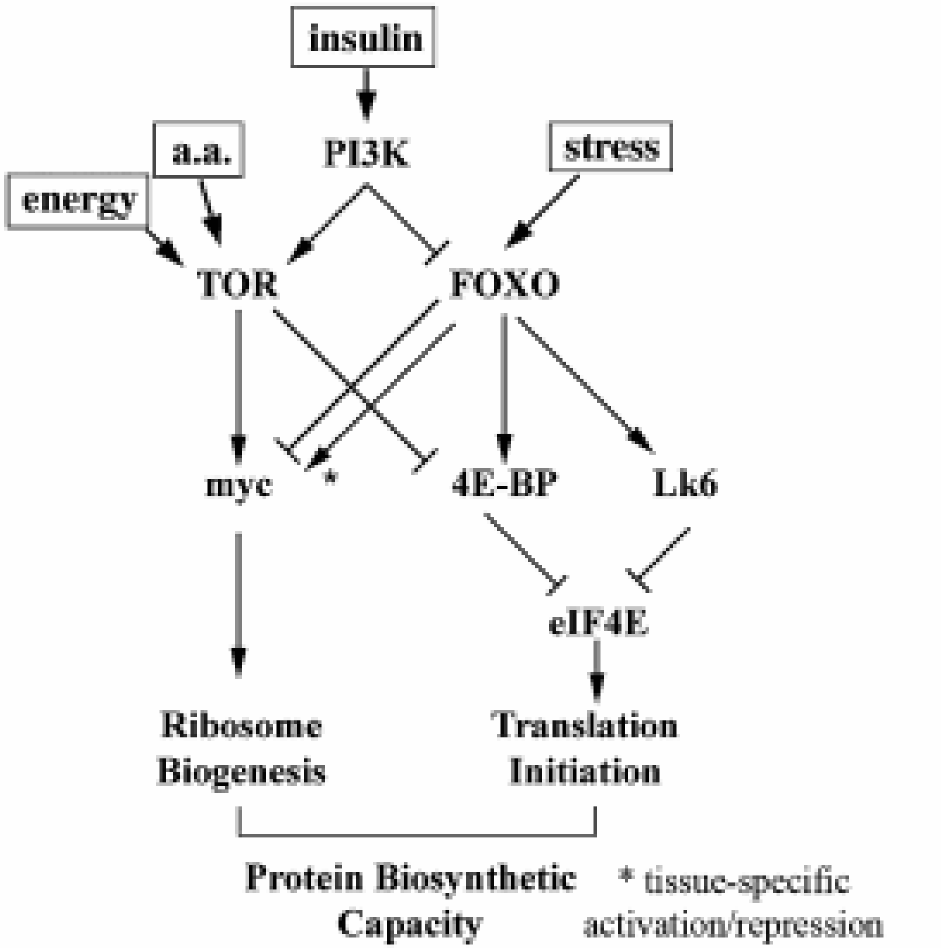
	1	2	3	4	5	6	7	8	9
1		1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9
2	2,1		2,3	2,4	2,5	2,6	2,7	2,8	2,9
3	3,1	3,2		3,4	3,5	3,6	3,7	3,8	3,9
4	4,1	4,2	4,3		4,5	4,6	4,7	4,8	4,9
5	5,1	5,2	5,3	5,4		5,6	5,7	5,8	5,9
6	6,1	6,2	6,3	6,4	6,5		6,7	6,8	6,9
7	7,1	7,2	7,3	7,4	7,5	7,6		7,8	7,9
8	8,1	8,2	8,3	8,4	8,5	8,6	8,7		8,9
9	9,1	9,2	9,3	9,4	9,5	9,6	9,7	9,8	

Conclusions: males possess ~no epistatic variation in transcript levels, while females are overwhelmed with it

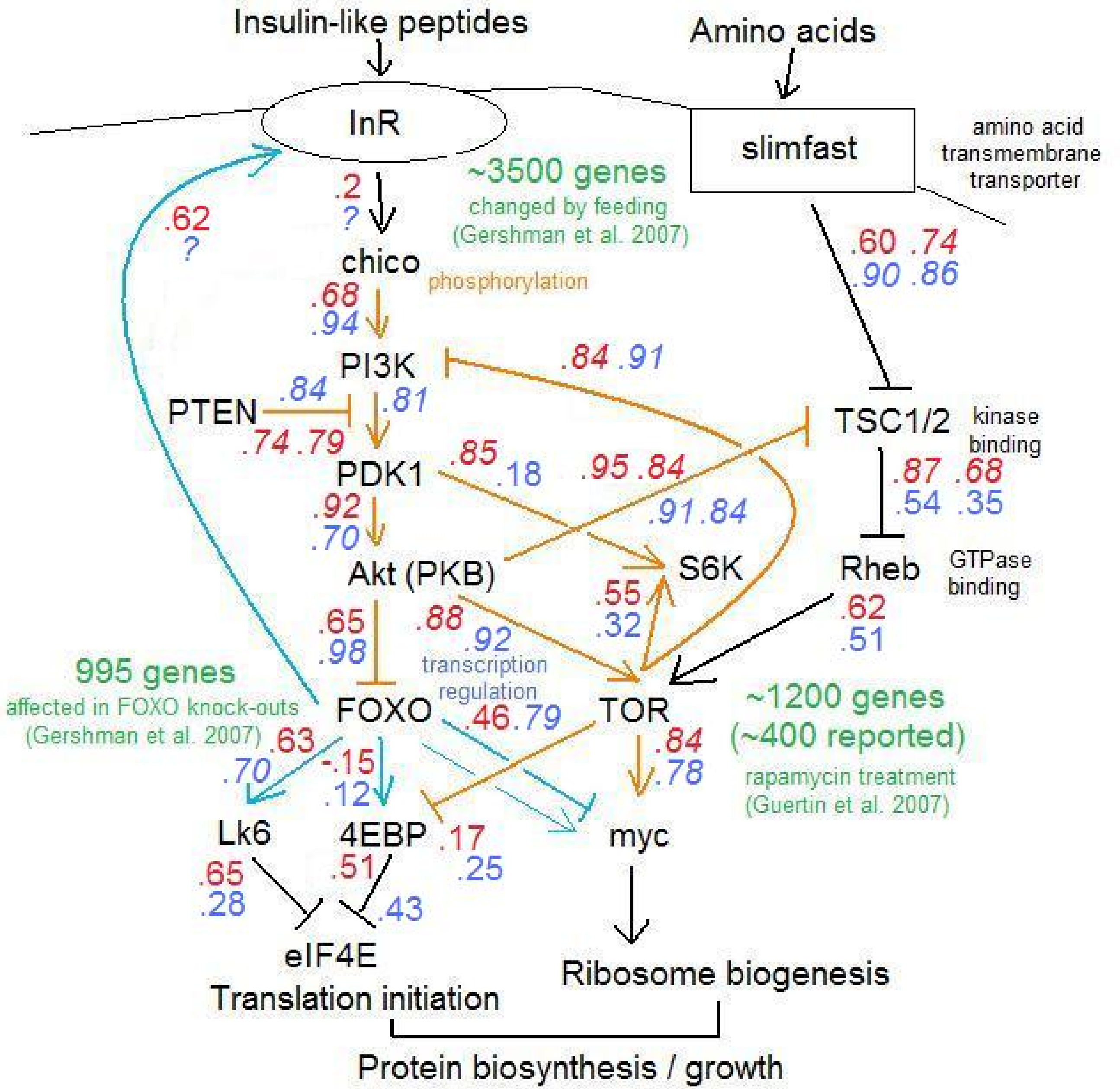
	Male		Female	
	X	Autosome	X	Autosome
Genetic	192	1337	604	2720
GCA	**185**	**1336**	**88**	**630**
SCA	**1**	**2**	**361**	**1284**
rGCA	44	23	1	1
rSCA	2	1	341	1380

How variation in pathways scales into phenotypic variation, InR/TOR pathway in *D. melanogaster*.

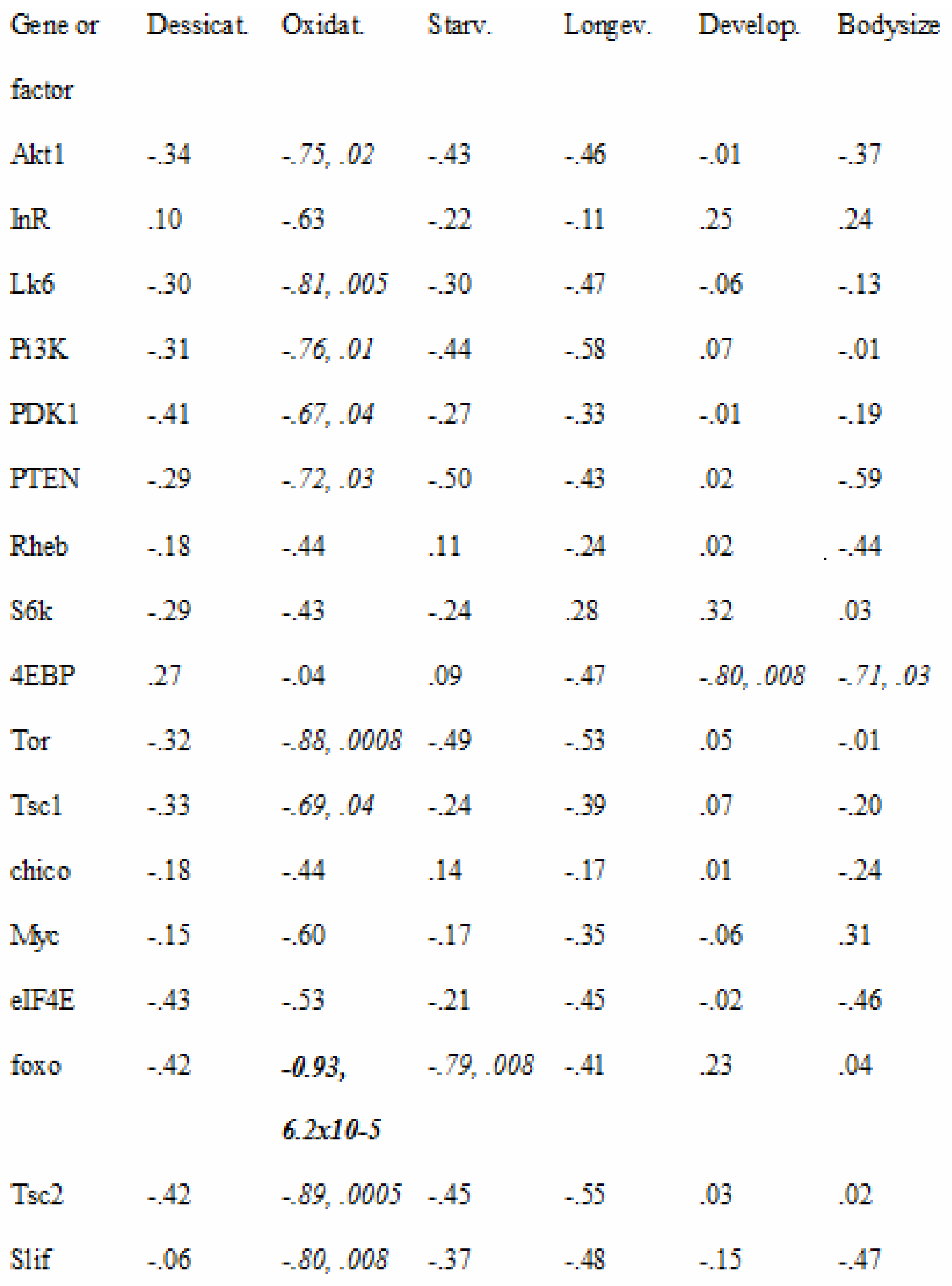
- Does transcriptome variation shape with the networks inferred with major-effect mutations?
- Does genetic variation in networks map into phenotypic variation?



Does transcriptome variation shape with the networks inferred with major-effect mutations?

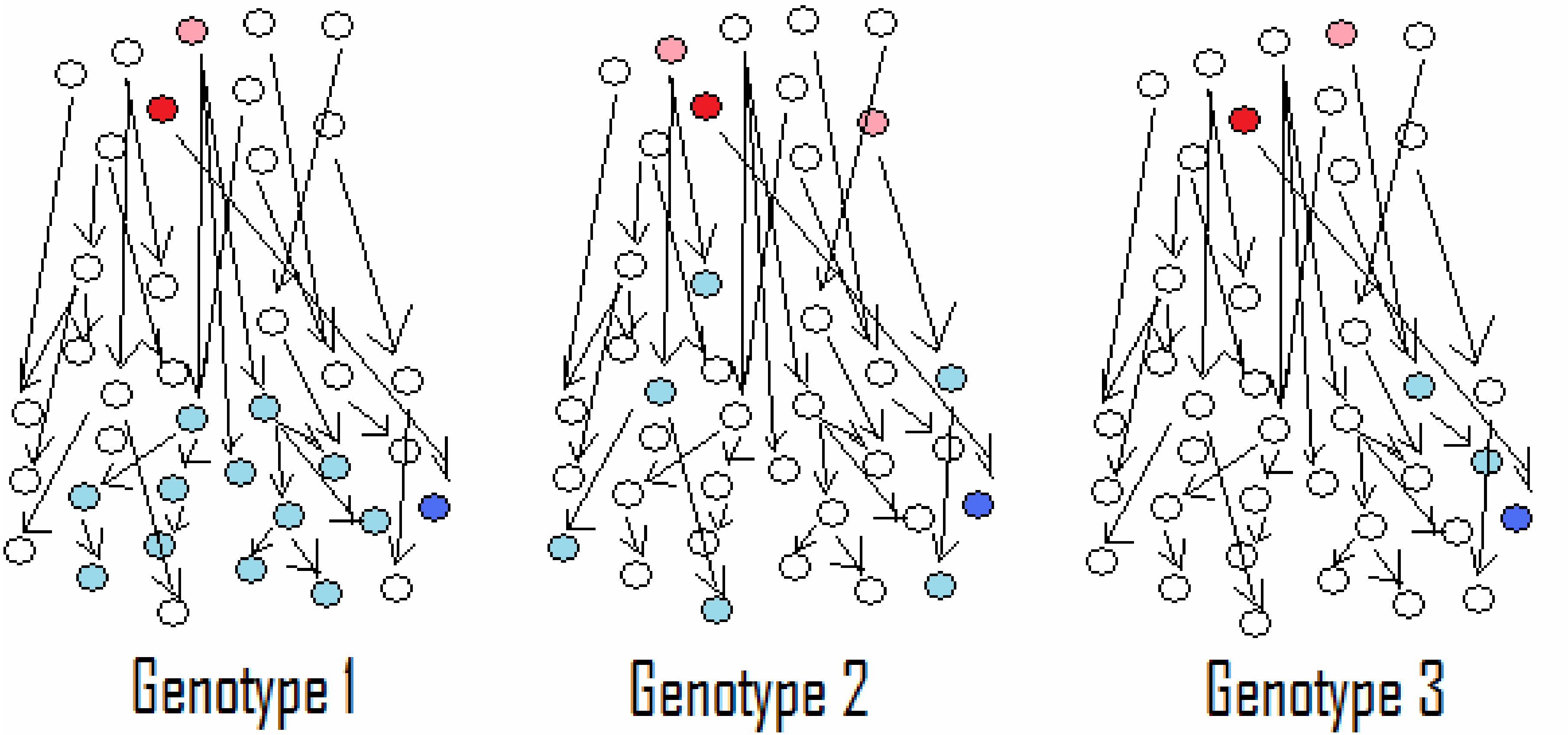


Tons of correlations, naturally



Gene or factor	Dessicat.	Oxidat.	Starv.	Longev.	Develop.	Bodysize
Akt1	-.34	*-.75, .02*	-.43	-.46	-.01	-.37
InR	.10	-.63	-.22	-.11	.25	.24
Lk6	-.30	*-.81, .005*	-.30	-.47	-.06	-.13
Pi3K	-.31	*-.76, .01*	-.44	-.58	.07	-.01
PDK1	-.41	*-.67, .04*	-.27	-.33	-.01	-.19
PTEN	-.29	*-.72, .03*	-.50	-.43	.02	-.59
Rheb	-.18	-.44	.11	-.24	.02	-.44
S6k	-.29	-.43	-.24	.28	.32	.03
4EBP	.27	-.04	.09	-.47	*-.80, .008*	*-.71, .03*
Tor	-.32	*-.88, .0008*	-.49	-.53	.05	-.01
Tsc1	-.33	*-.69, .04*	-.24	-.39	.07	-.20
chico	-.18	-.44	.14	-.17	.01	-.24
Myc	-.15	-.60	-.17	-.35	-.06	.31
eIF4E	-.43	-.53	-.21	-.45	-.02	-.46
foxo	-.42	***-0.93, 6.2x10-5***	*-.79, .008*	-.41	.23	.04
Tsc2	-.42	*-.89, .0005*	-.45	-.55	.03	.02
Slif	-.06	*-.80, .008*	-.37	-.48	-.15	-.47

What’s causing what?



- Allele specific expression → *cis*;
- The rest are *trans*;
- 200 genotypes?

Should work…., might be?