

# Causes and consequences of genetic background effects in *Drosophila melanogaster*.



Ian Dworkin

Program in Ecology, Evolutionary Biology and Behavior

Program in Genetics

Department of Zoology

*The Integrative Study of Animal Biology*

MICHIGAN STATE  
UNIVERSITY

# Thanks to...

**Dworkin lab:**

**David Tack**

**Alycia Kowalski**

**Mike DeNieu**

**Sudarshan Chari**

**Lin Choi**

**Chris Chandler**

**Will Pitchers**

**Andrew Victory**

**MSU:**

**Titus Brown**

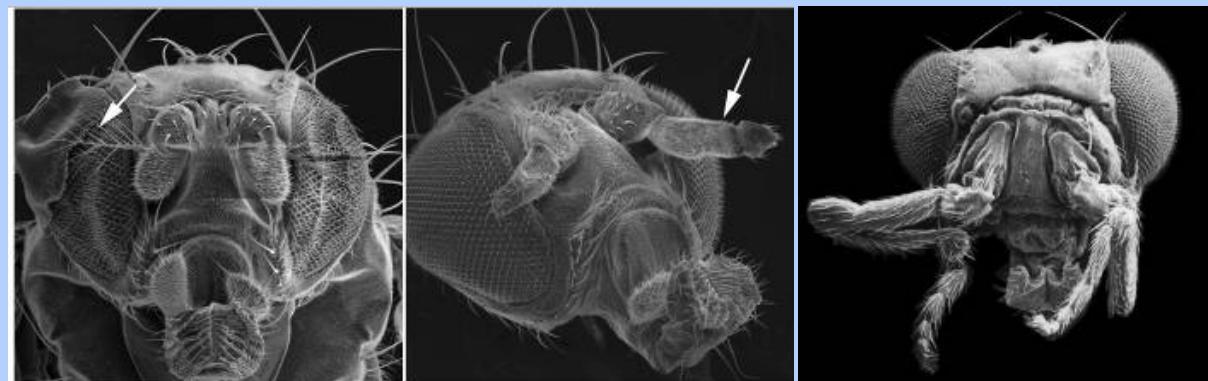
**Elsewhere**

**Greg Gibson (Georgia Tech)**

**Jim Mahaffey (NCSU)**



# The analysis of phenotypic variation



# Talk outline

- Introduction to conditional effects of mutations.
- Using induced mutations in *Drosophila* to study genetic background effects.
- Integrating genetic and genomic approaches to study the causes and consequences of genetic background effects.
- Mapping the loci that contribute to genetic background effects.

# The genetics of trait *expression* VS. *variation.*

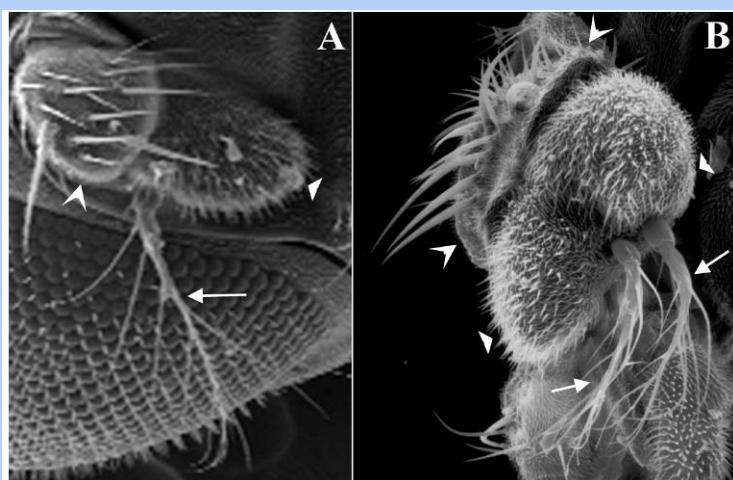
- Genetic analysis of trait expression:
  - Functional genetic analysis.
  - Utilize mutations of large effect (generated via mutagenesis).
  - Are genes “necessary” and “sufficient” for trait expression?
- Genetic analysis of phenotypic variation:
  - Domain of Quantitative and Population genetic analysis.
  - Studying standing genetic variation in natural populations.
  - Most often allelic variation of small effect.
  - How do alleles contribute to overall phenotypic variation for the trait?

Individual mutant alleles can have variable effects depending upon:

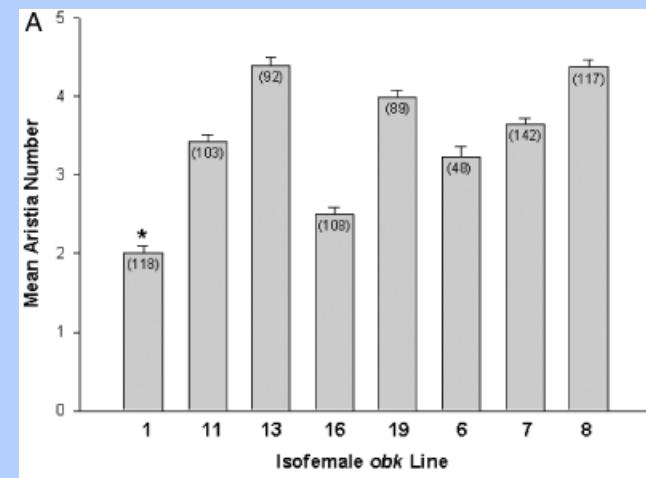
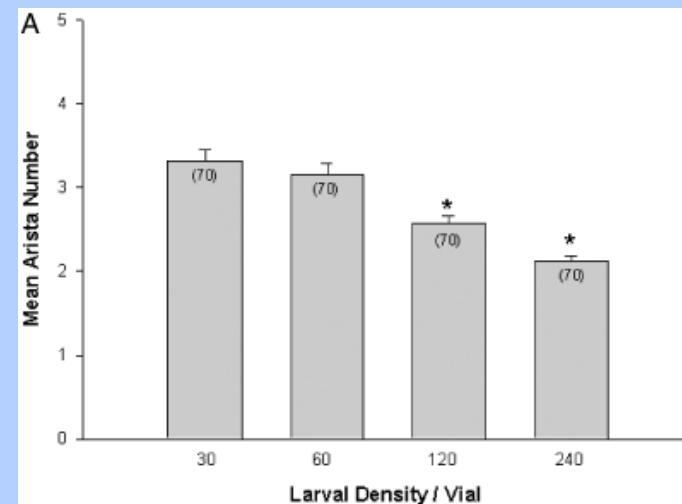
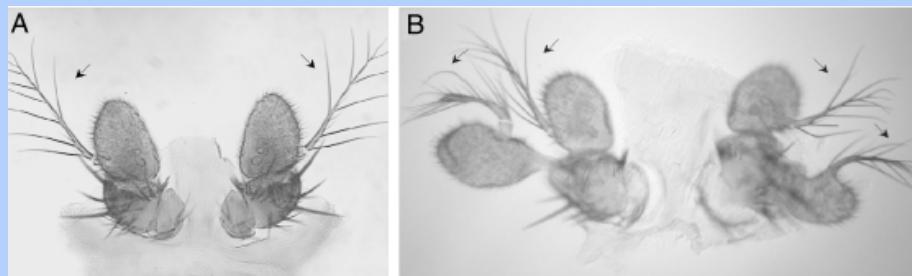
- Environmental effects such as rearing temperature, nutrition, density:  
Genotype by Environment (GxE) interactions.
- The genetic BACKGROUND in which they occur: Genotype by Genotype (GxG) interactions, or epistasis.
- The expressivity of an allele can itself be considered a quantitative trait.

# GxE and GxG for *obake*

wt



*obake* (*obk*)



Dworkin et. al. 2001; Atallah et. al. 2004

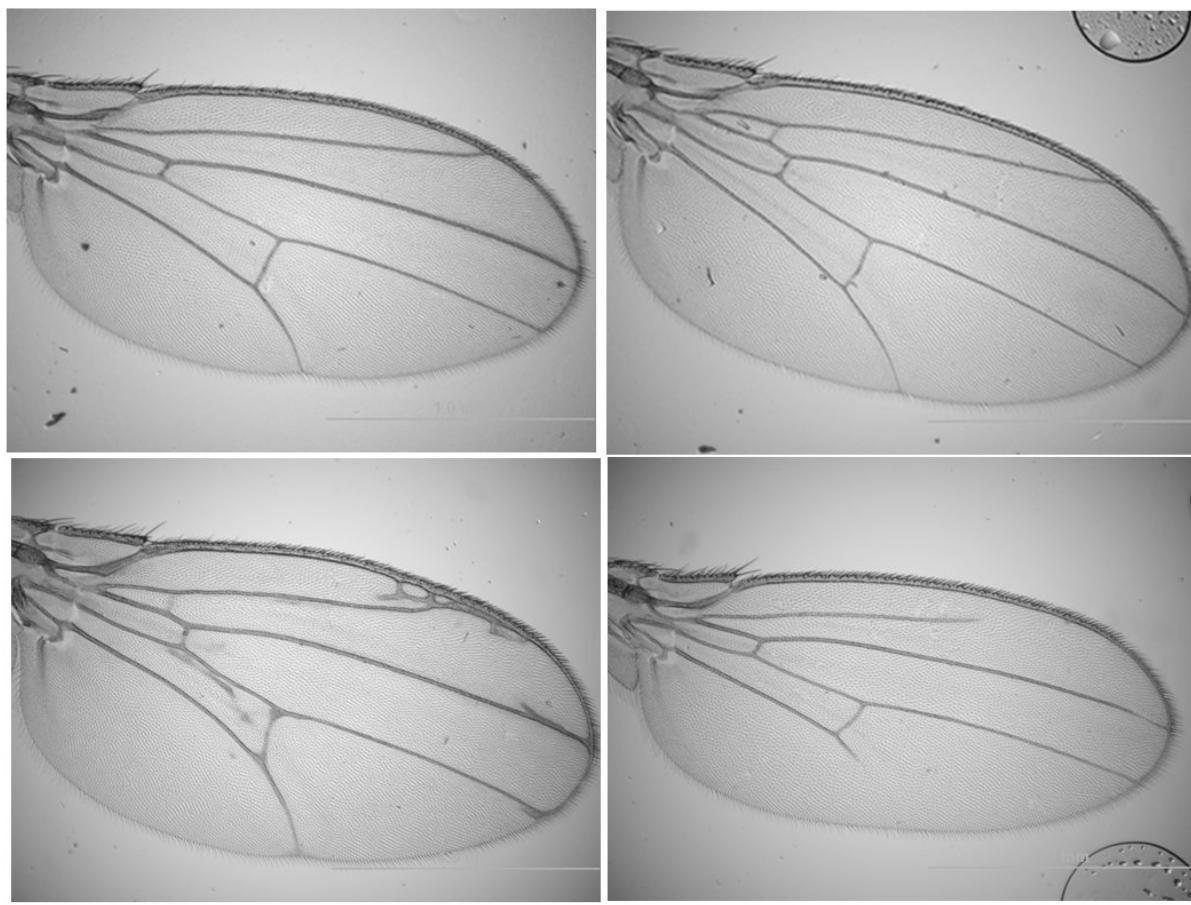
# The ubiquity of genetic background effects

- Genetic background effects have been observed in many organisms spanning bacteria, yeast, plants & animals.
- It is commonly observed in mouse, *C.elegans* and *Drosophila*, across the range of traits and allelic severity, including nulls/amorphic alleles!!

# For your consideration...

- Studying mutations in only one isogenic “wild-type” background may miss important components of the biology.
- From a functional perspective: it provides insight into how epistasis ( $G \times G$ ) occurs.
- From an evolutionary perspective: alter our predictions on the fate of new mutations, introgressed alleles & maintenance of genetic variation.

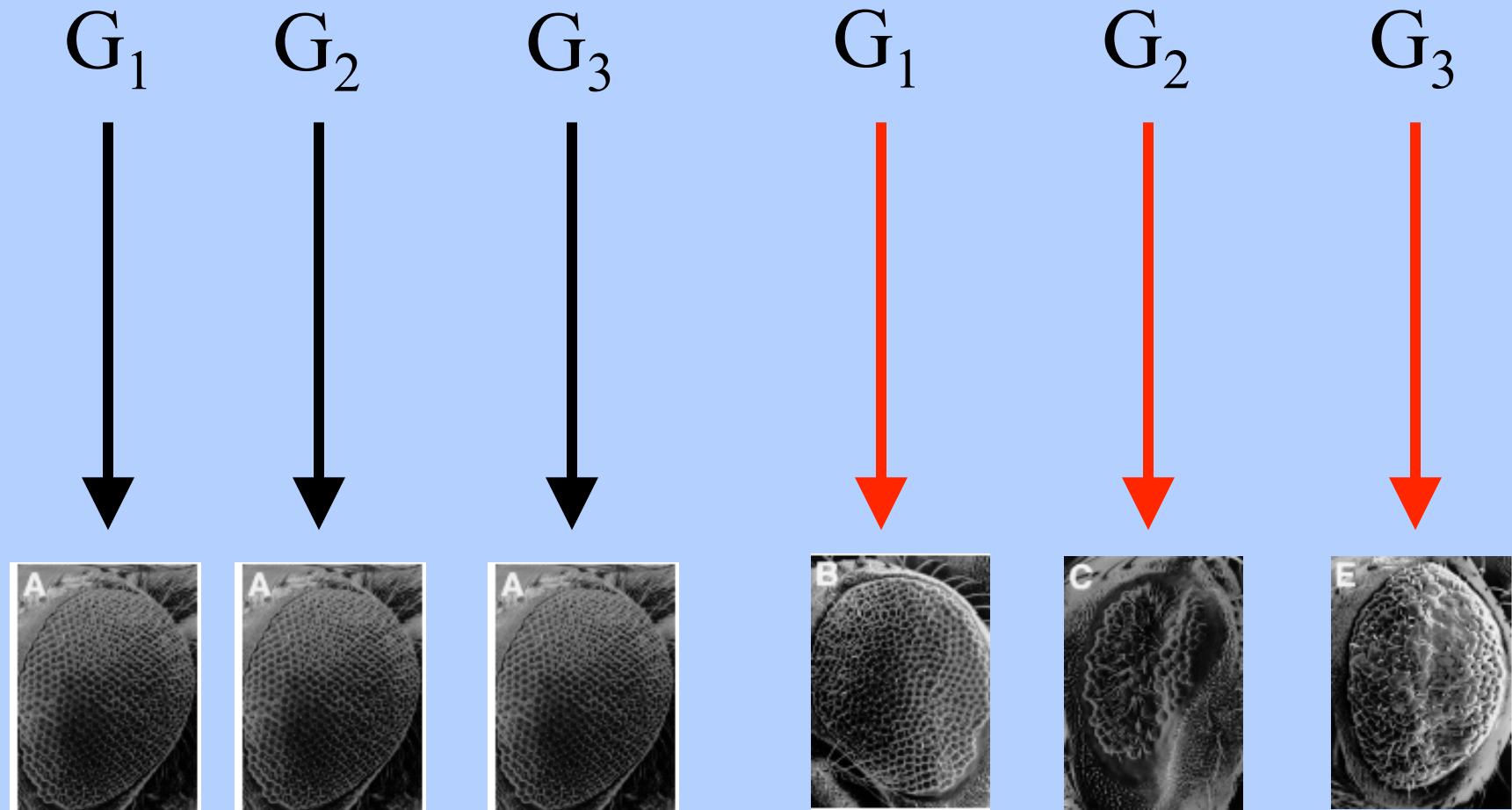
# High prevalence of “mutant” wing phenotypes in natural populations of *Drosophila melanogaster*.



~40% of inbred lines generated from natural populations of *Drosophila melanogaster* appear to have qualitative defects in wing morphology.

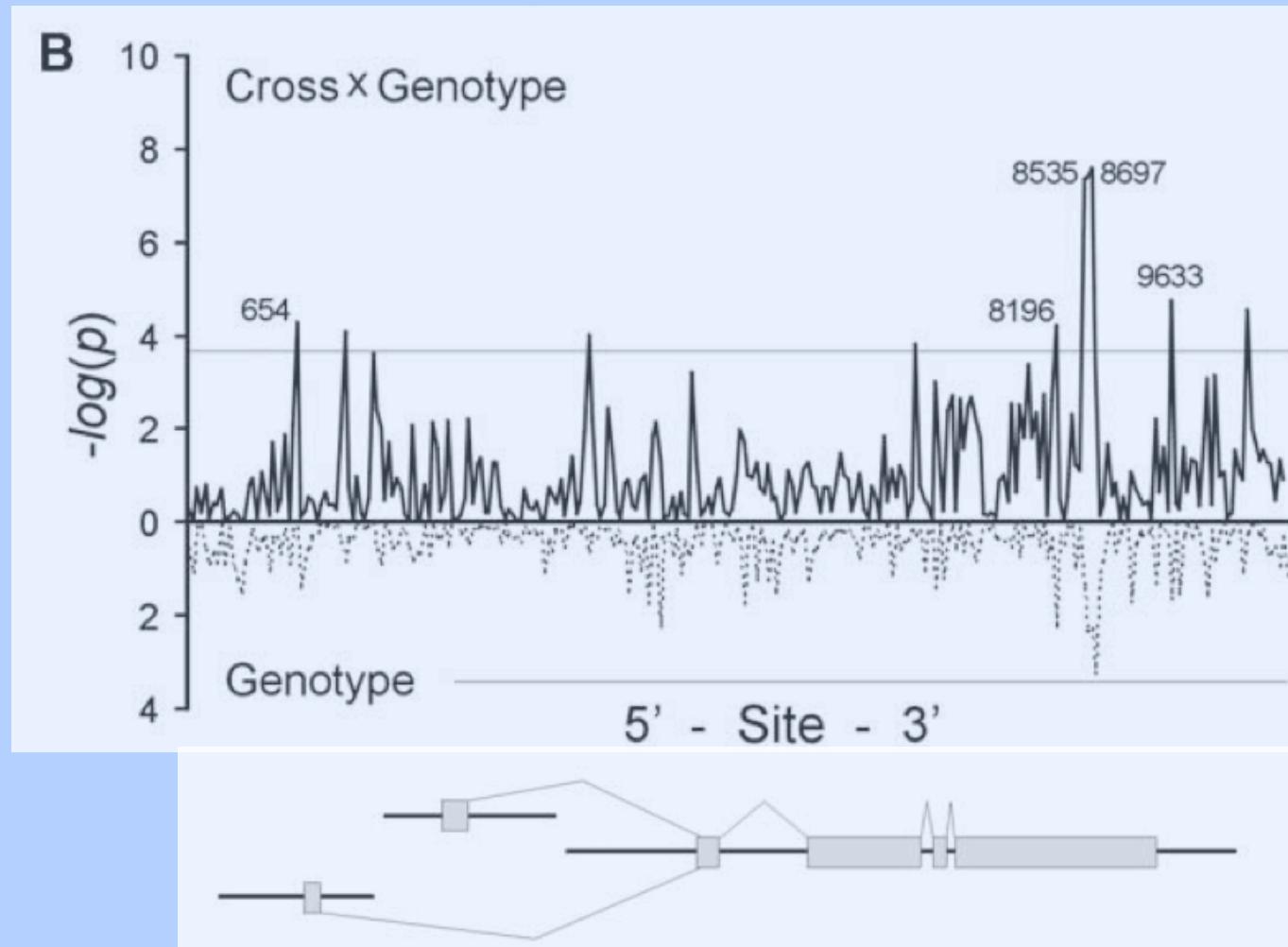
Often due to rare segregating alleles in single genes.

# Using mutations to unmask variation



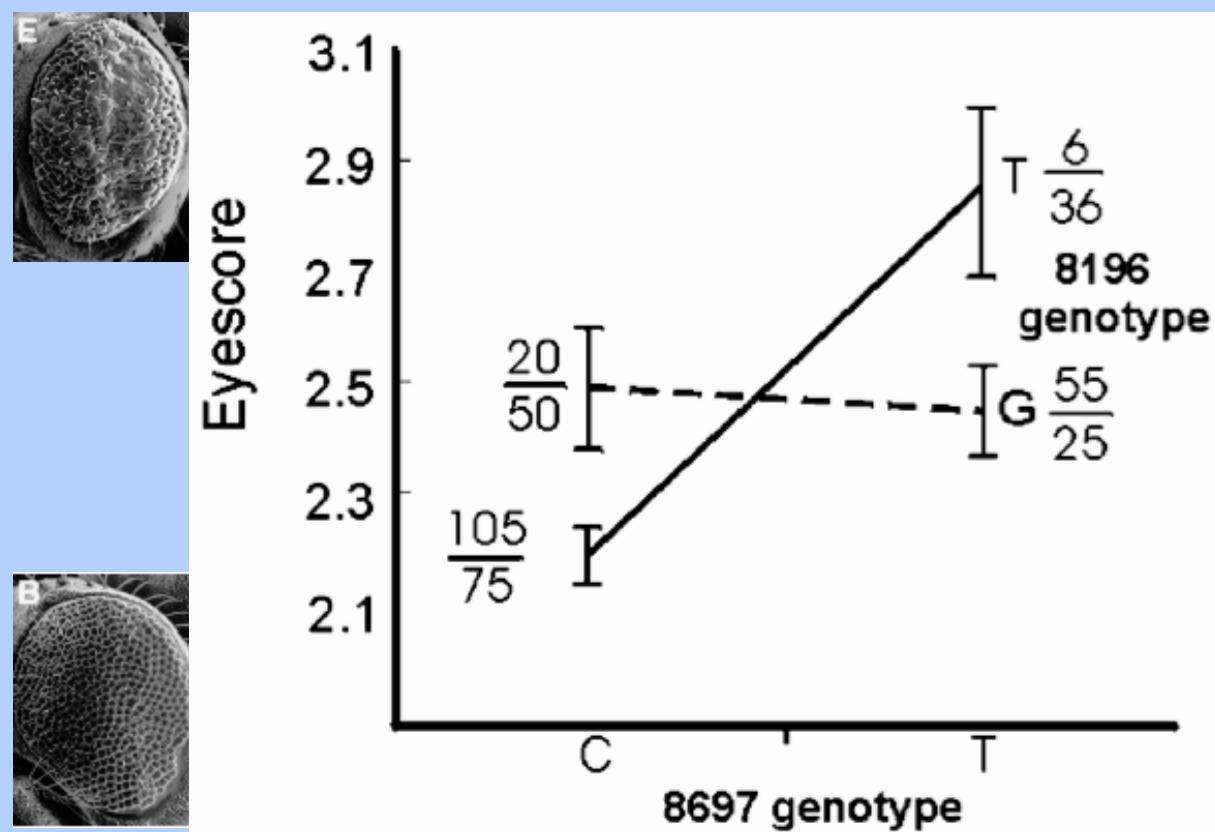
Sensitized by crossing to  
*Ellipse* allele

# SNP's within *Egfr* are associated with variation for photoreceptor determination.

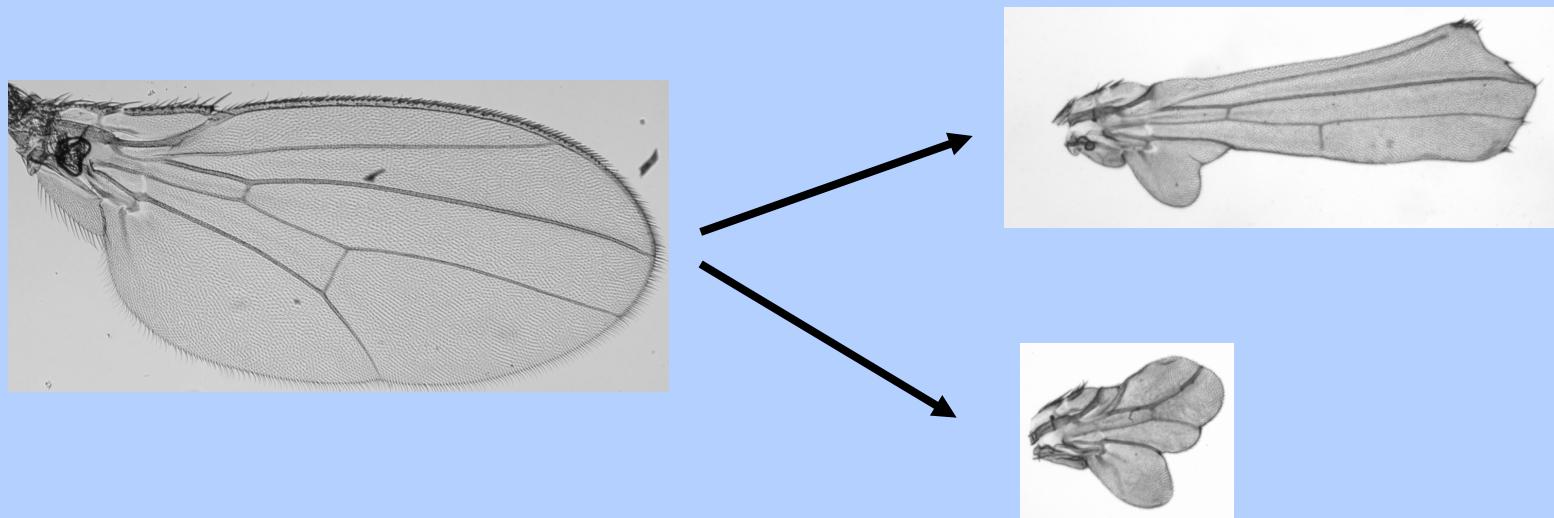


Dworkin et al. 2003

# Is this variation visible to natural selection?

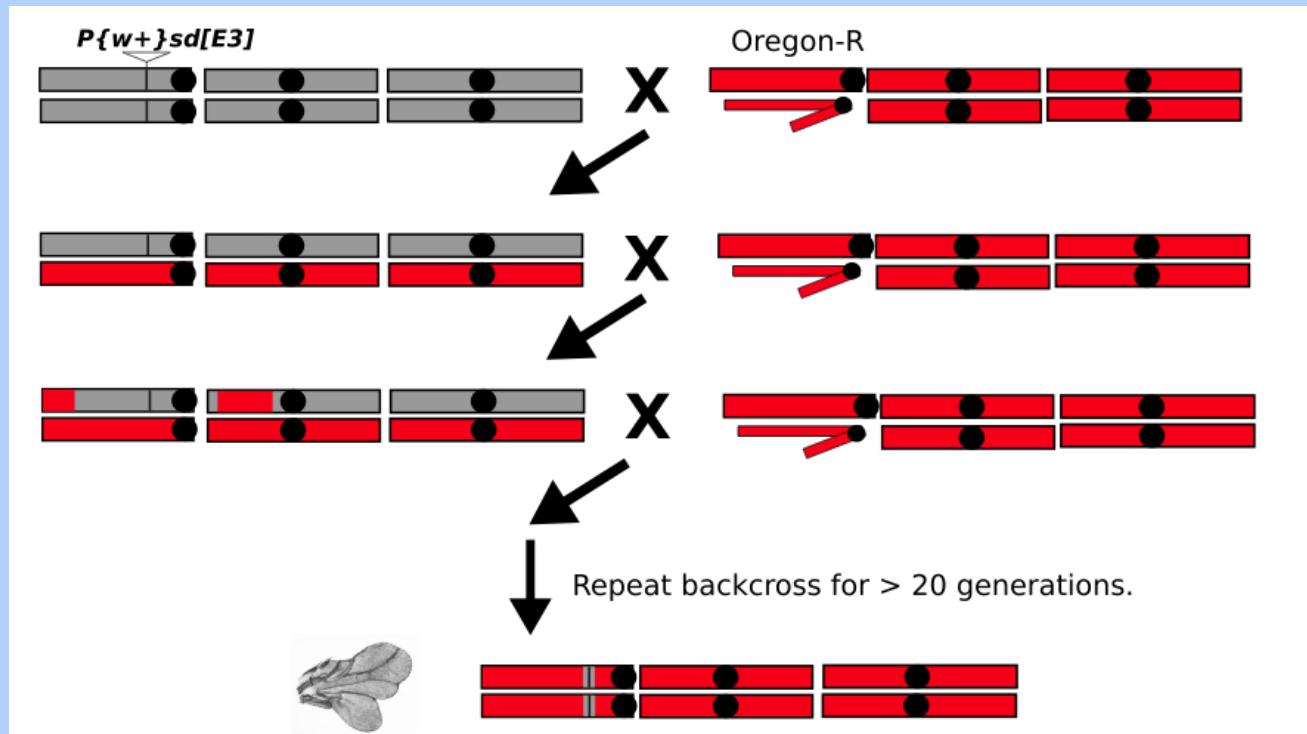


# Functional genetic analysis of background dependent effects



A case study with a *scalloped* mutant in the wing of  
*Drosophila*.

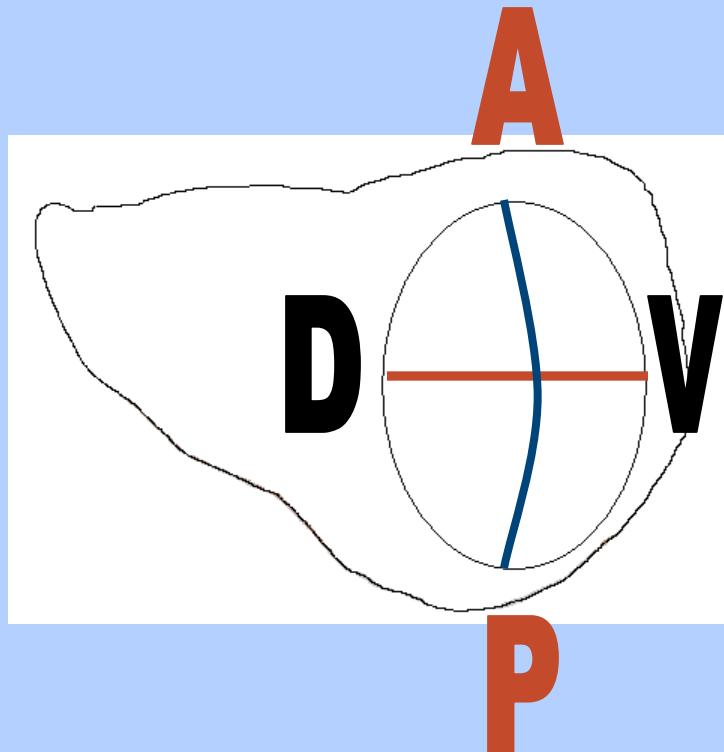
# Introgression of a *scalloped* allele



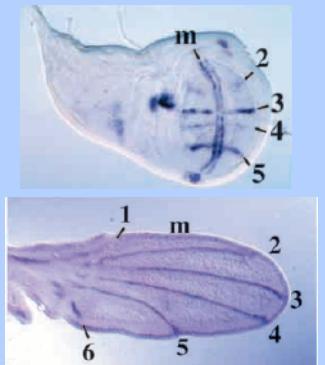
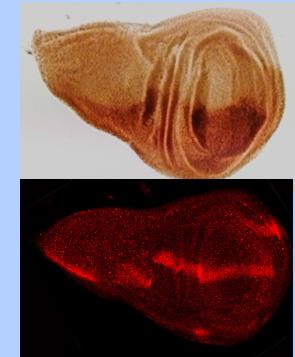
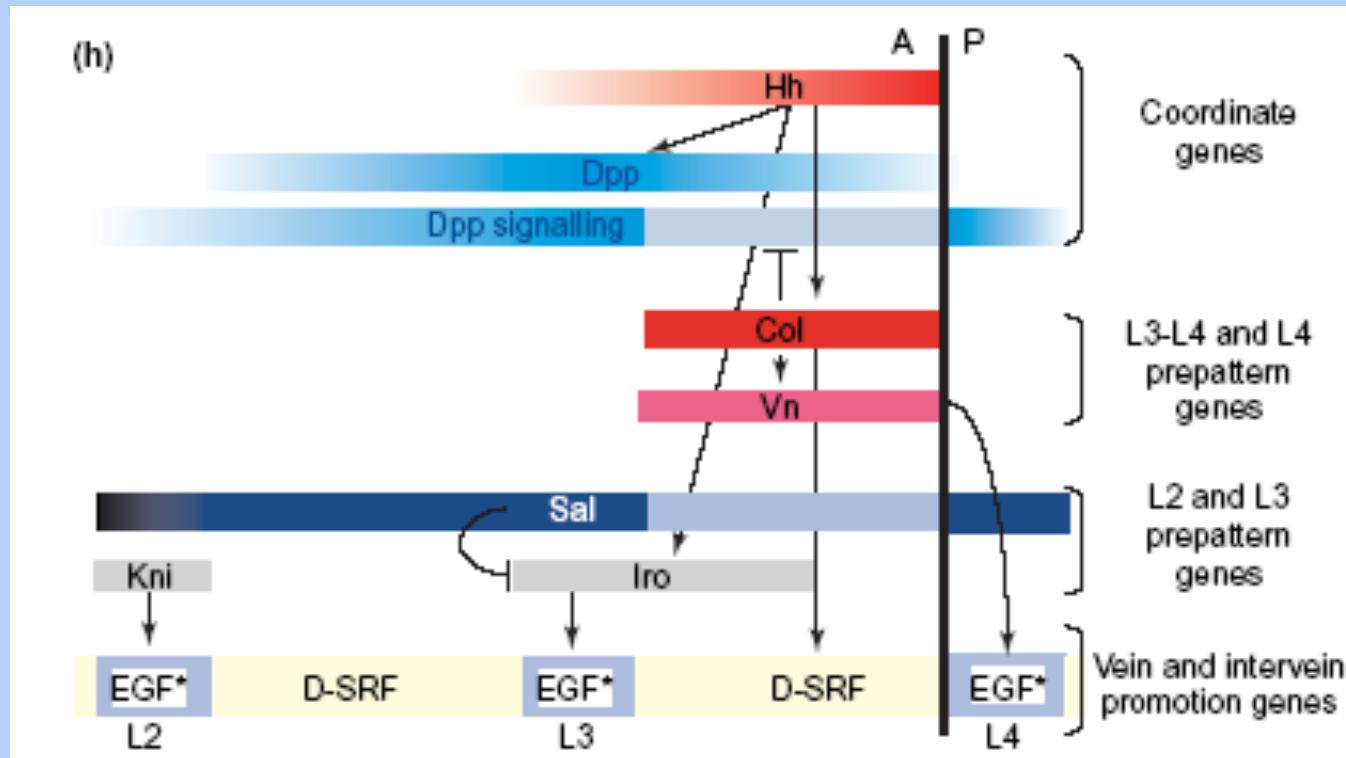
# Why focus on *scalloped*?

- Phenomenologically striking phenotypic variation between the effects observed in the two wild-type backgrounds (Sam vs. Ore).
- We understand a great deal about the role and function of *sd* during wing development, and to a lesser extent other phenotypes (olfaction and gustation).
- Potential pleiotropic effects on wing shape.

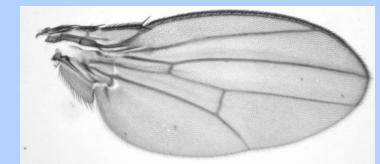
# Wing Development



# Patterning of the wing blade



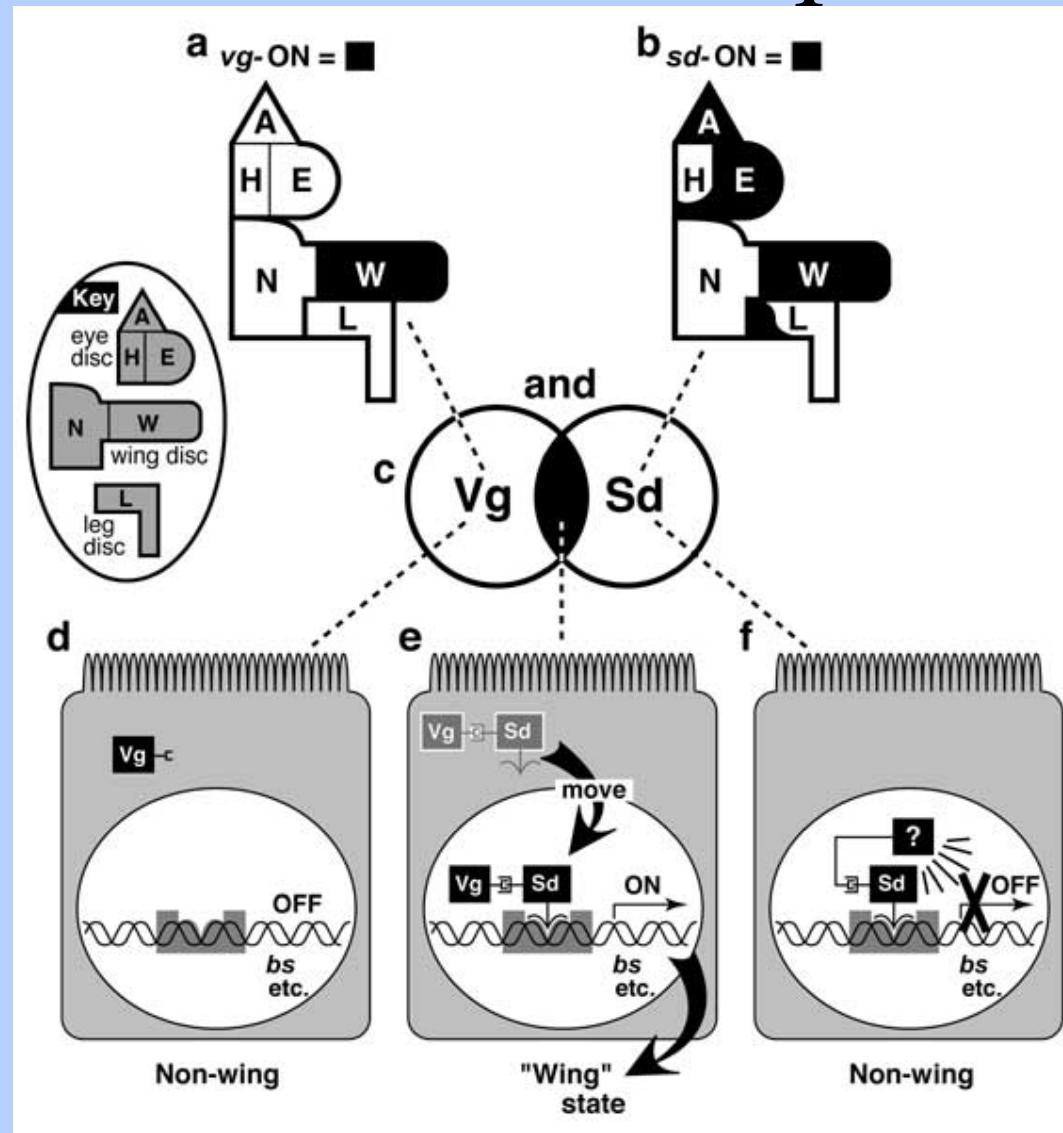
Crozatier et. al. 2004 TIG



# *sd* function during wing development

- Sd is a TEA class transcription factor that forms a heterodimer with Vestigial in the wing imaginal disc to promote wing cell fates.
- The Sd-Vg heterodimer regulates transcription at a number of known loci (*blistered, cut, wingless, vestigial, spalt,..*)

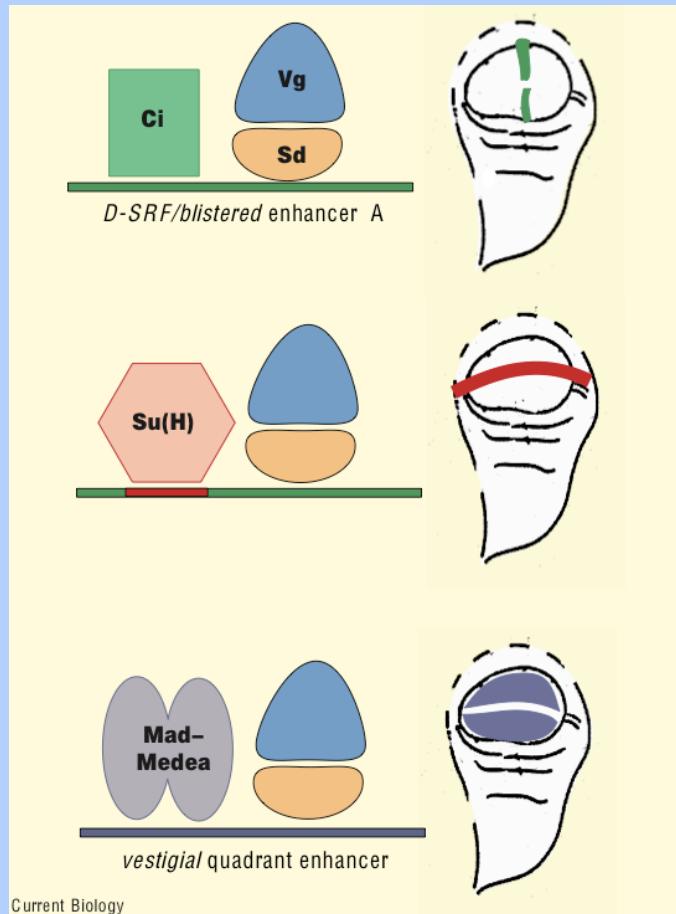
# The role of *scalloped*



Held 2003

# Necessary and Sufficient

Synthetic enhancers with  
Sd binding sites

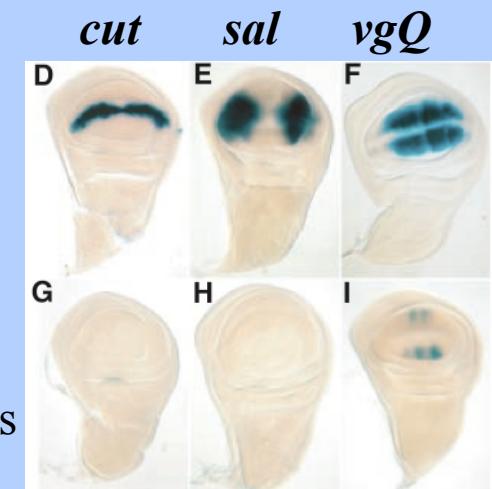


Current Biology

Bray 1999

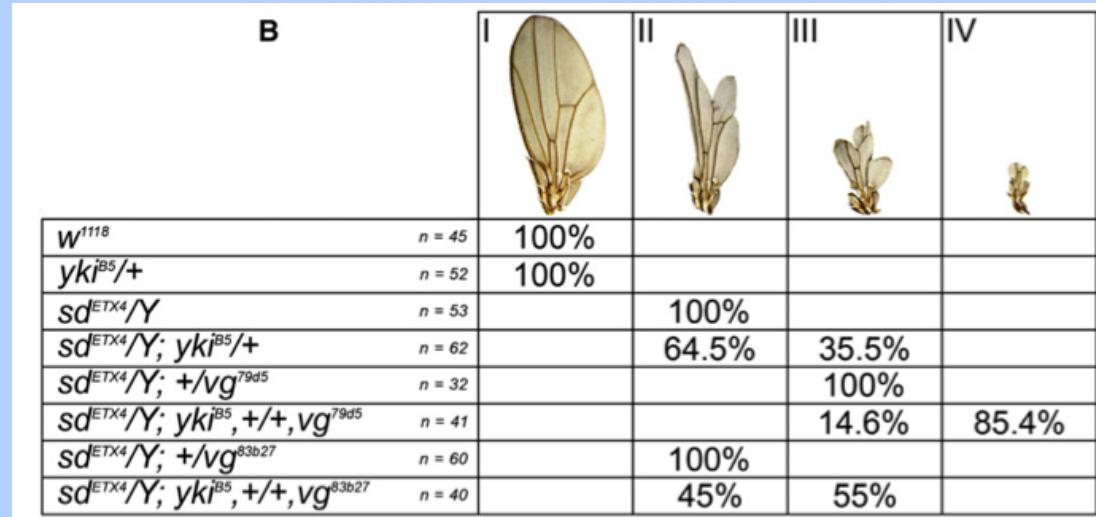
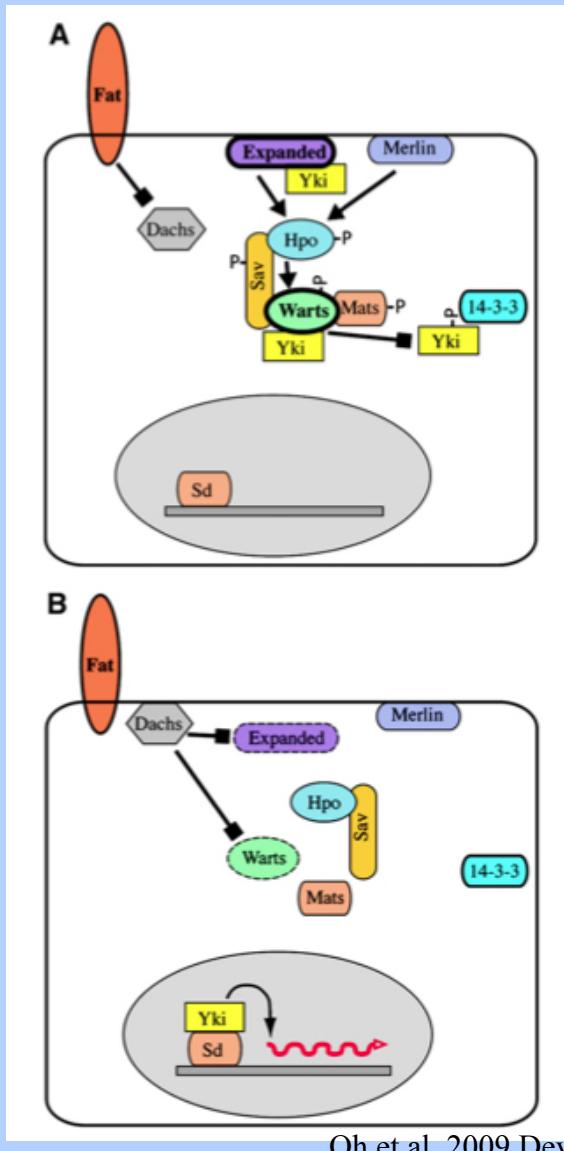
Wild-type  
reporter  
constructs

Mutated Sd  
binding sites



Guss et al. 2001

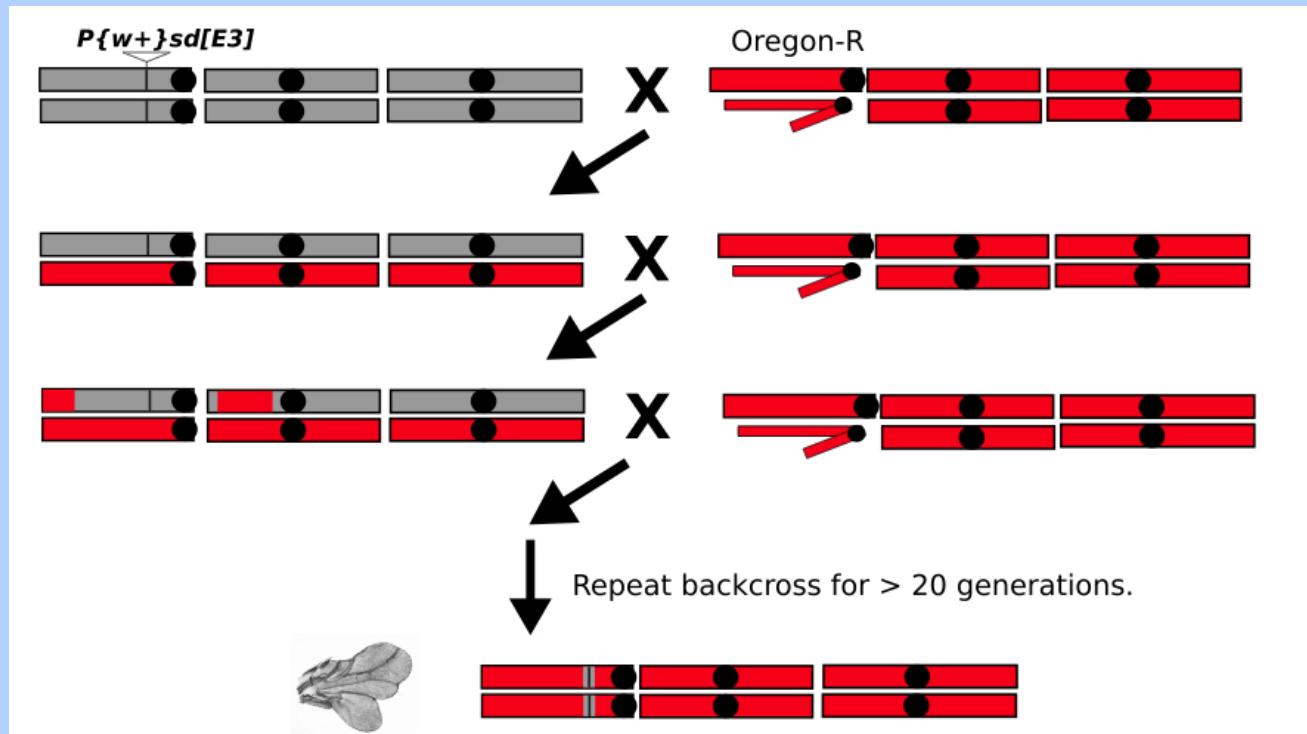
# Sd is also an important transcription factor influencing tissue growth and cell proliferation & tissue polarity



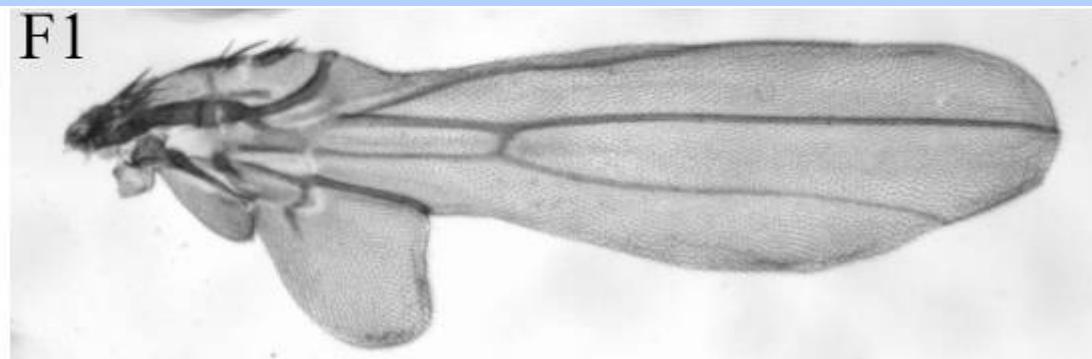
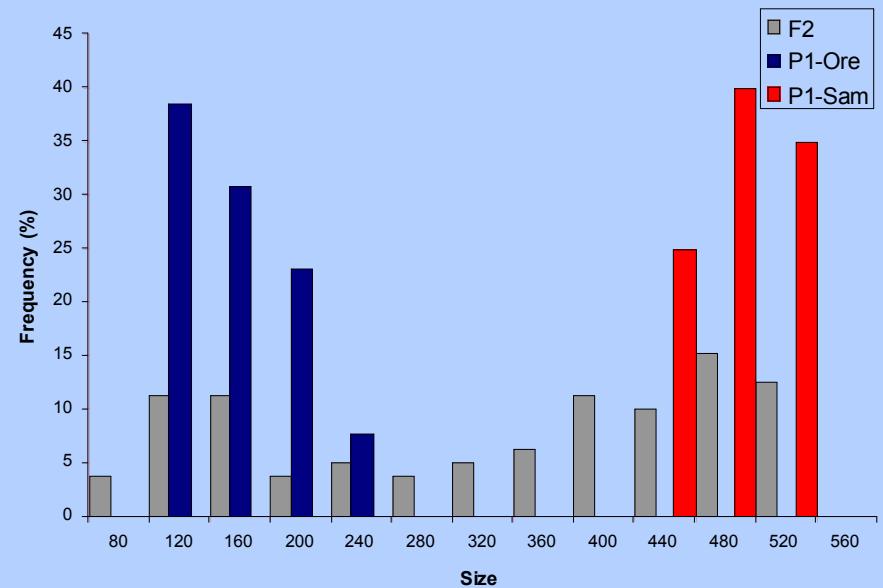
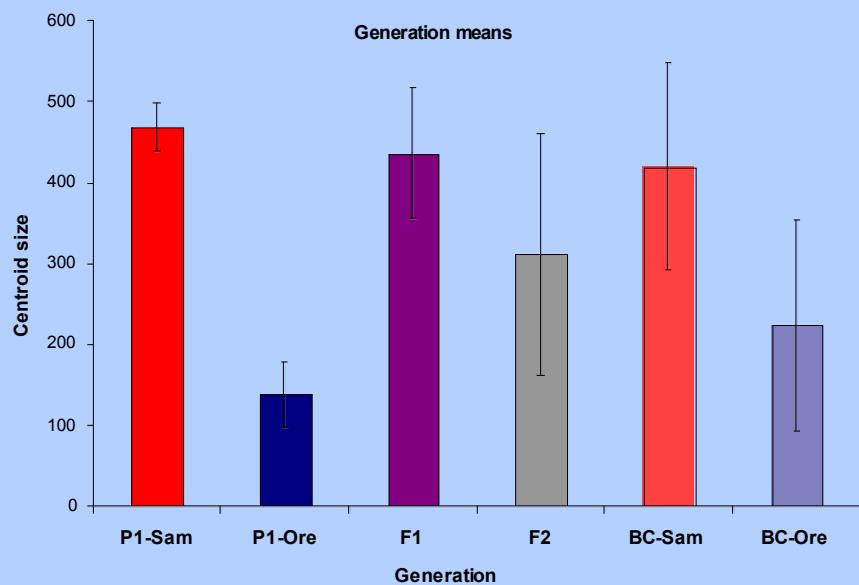
Goulev et al. 2008

Sd + Yrk regulate transcription of several genes, including: *diap1 (thread)*, *CyclinE*, *bantam*, *merlin*, *expanded*

# Introgression of a *scalloped* allele



# Generations means analysis for *sd* background effects

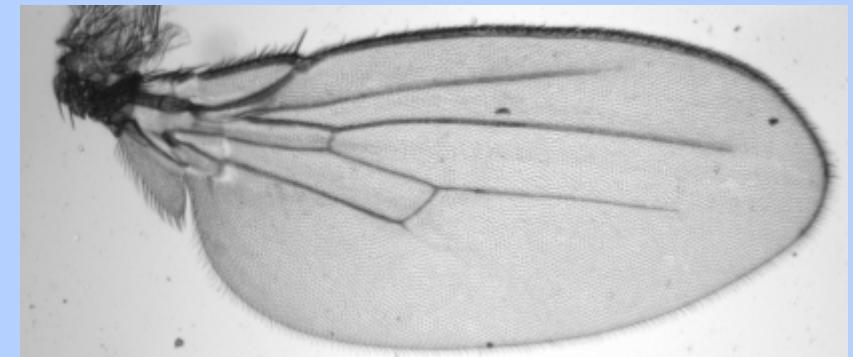
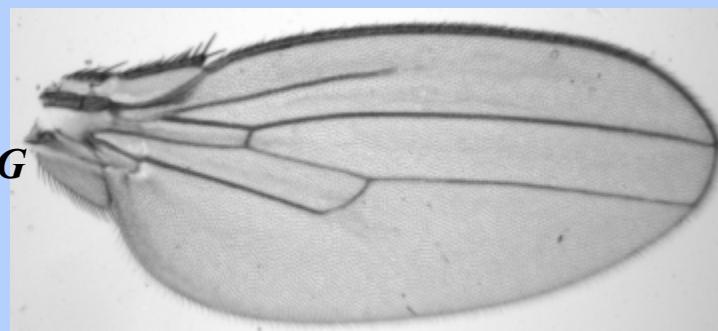


$sd^{E3}/Y$

# linked Genetic background effects?

At least one of the modifiers of the genetic background effect for *sd* is linked to the same marker for a background modifier of *rhomboid*, and for quantitative variation for wing shape? Are they the same polymorphisms?

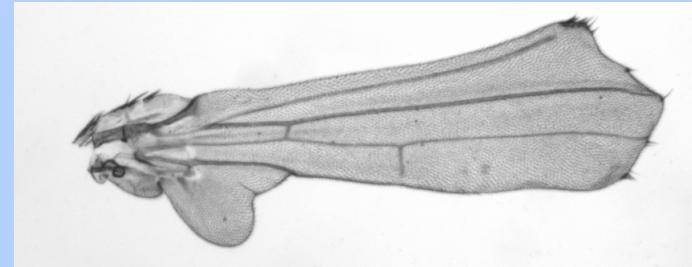
*rhomboid*<sup>KG</sup>



*scalloped*<sup>E3</sup>



Ore



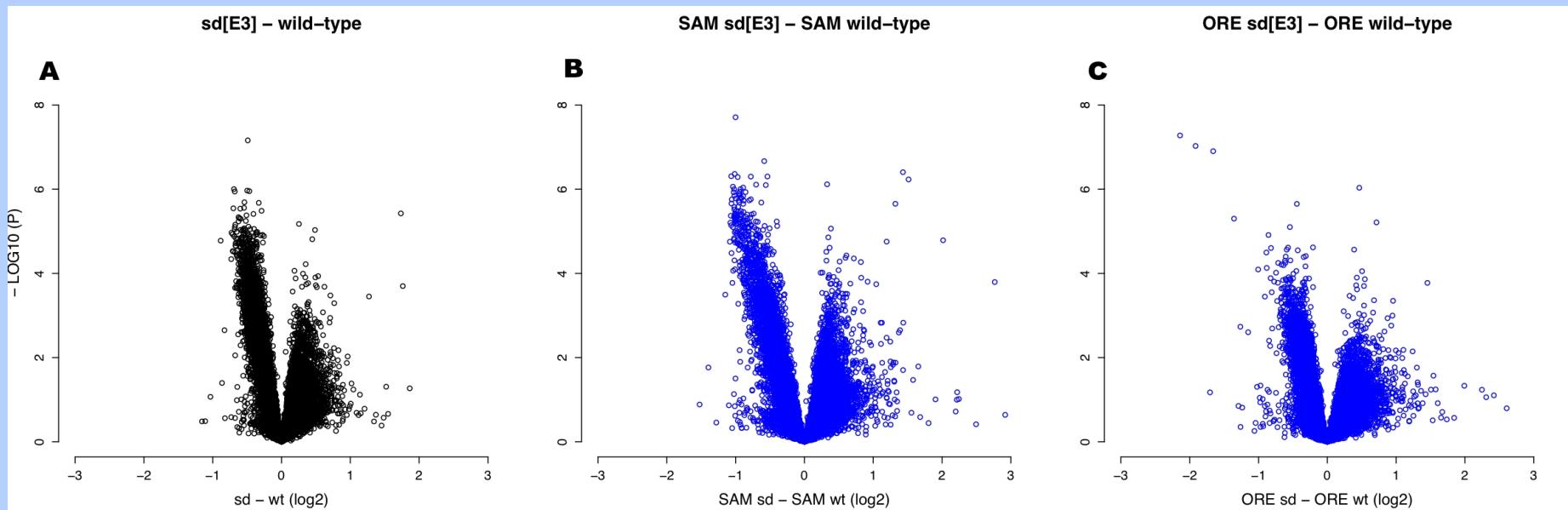
Sam

# Models to explain the phenotypic consequences of genetic background.

1. *sd* mutation enhances background specific differences in transcription, mediating the observed phenotypic differences.
2. Background effects involve a set of genes that partially overlap with differentially expressed genes between mutant and wild-type, with quantitative differences in transcription **that correlate with variation for the *sd* phenotype.**
3. Background effects involve a **different set of genes** than those that mediate the main effects of the mutant *sd* allele.

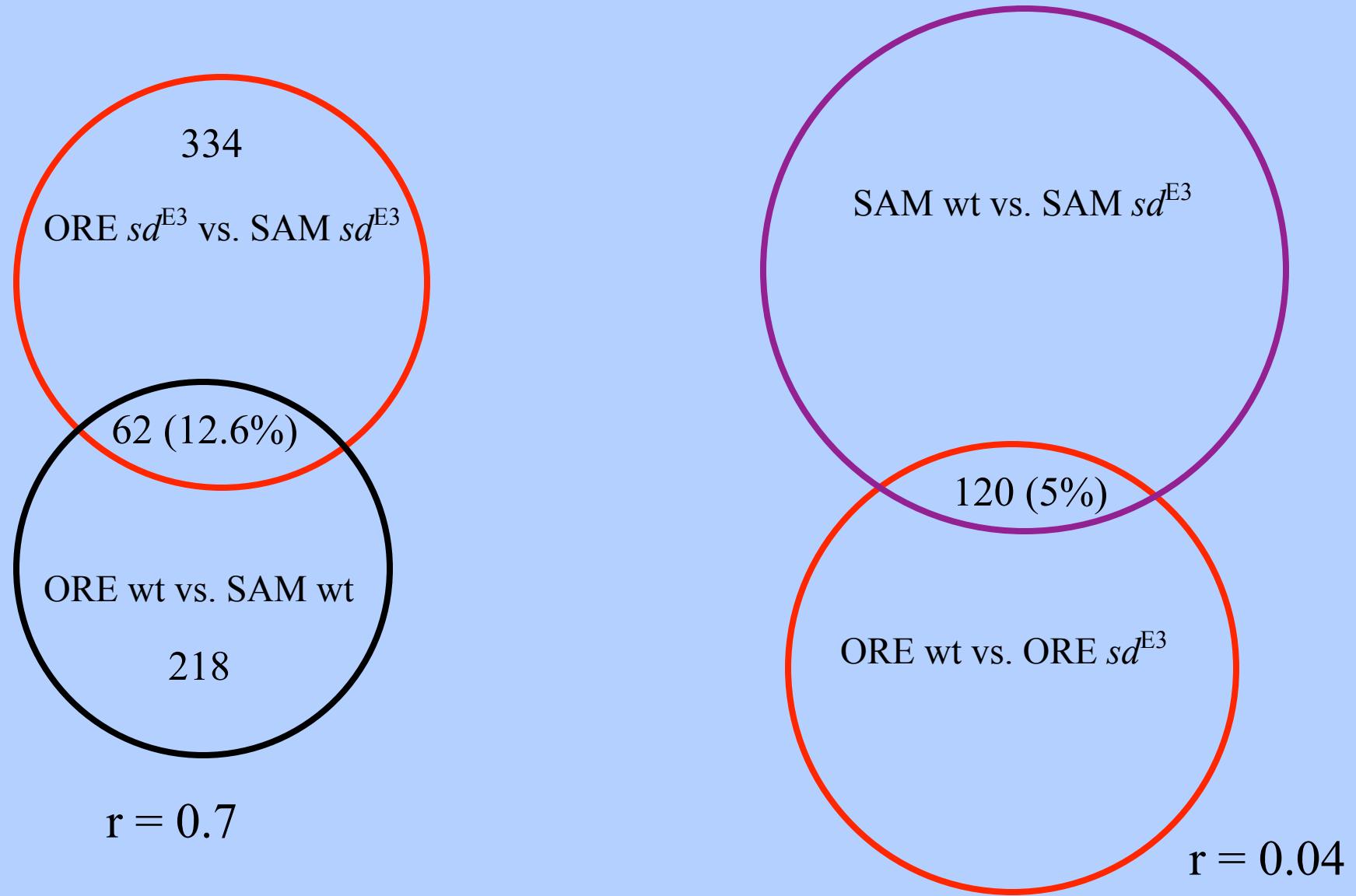
# $sd^{E3}$ vs.wild-type

overall

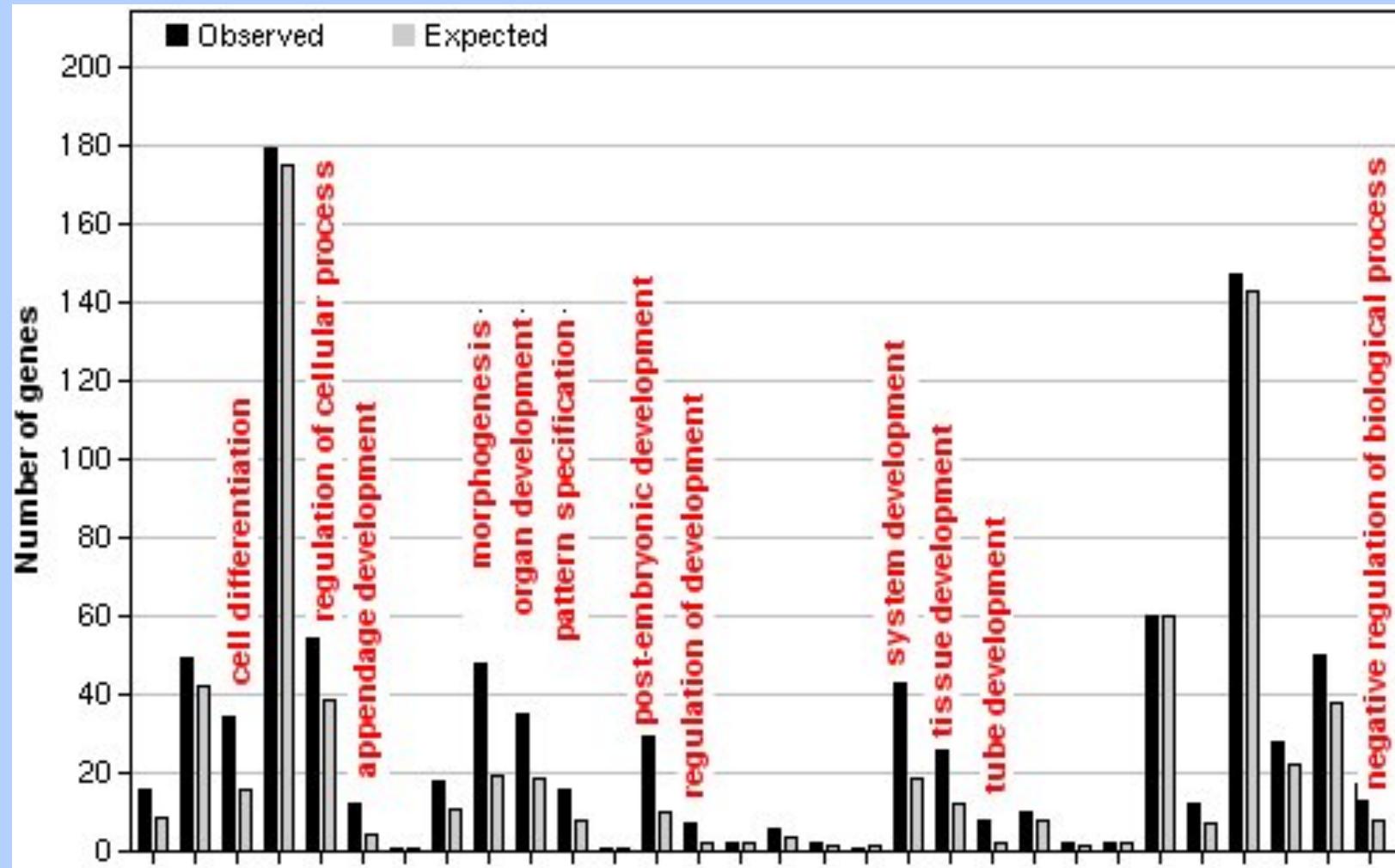


These “volcano plots” demonstrate that the genes (~1000) that are differentially expressed between  $sd^{E3}$  & wild-type, are down-regulated in the  $sd^{E3}$  genotype, but in a background dependent manner.

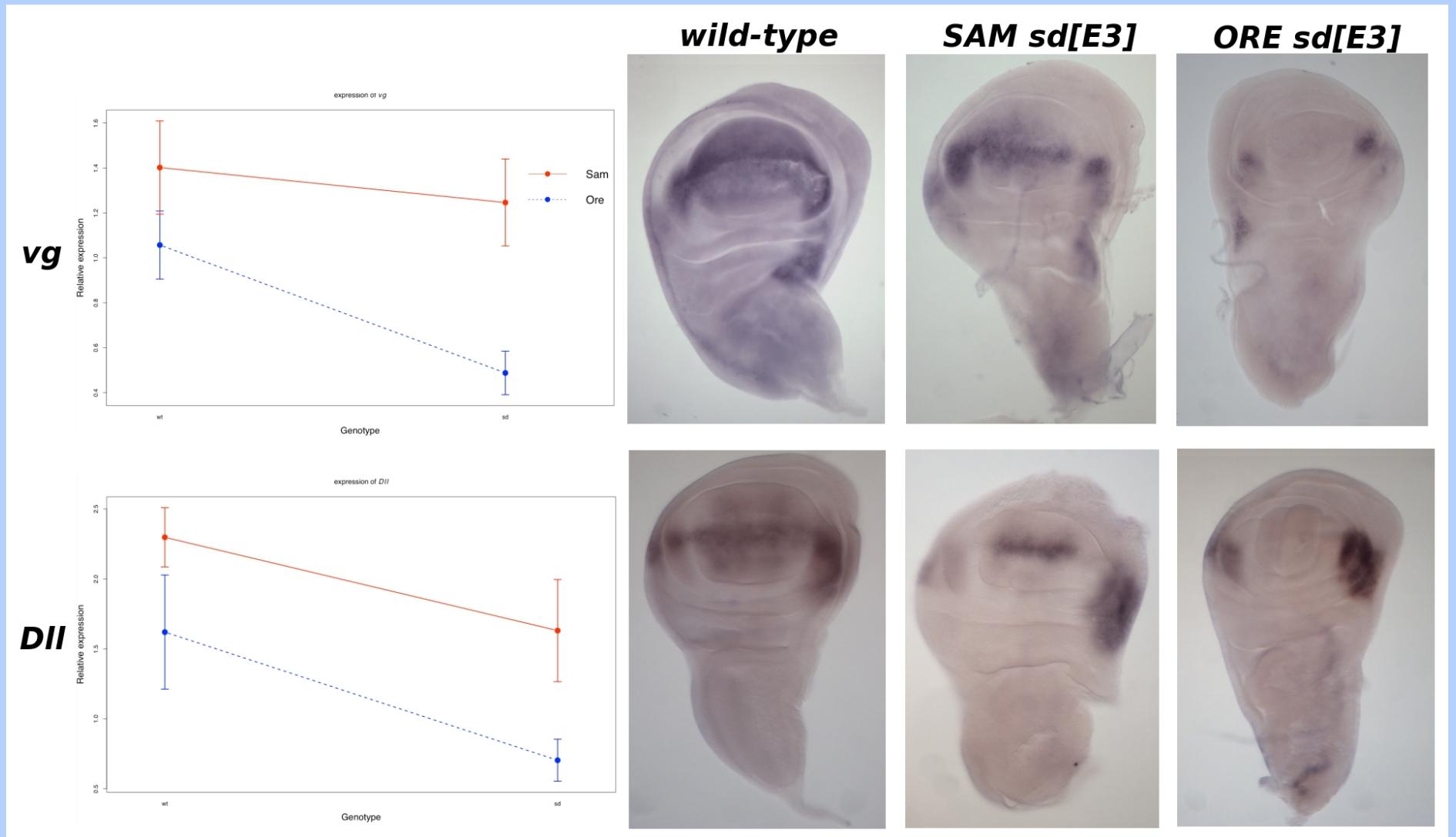
# Relationships in expression profiles are consistent with models 2&3.



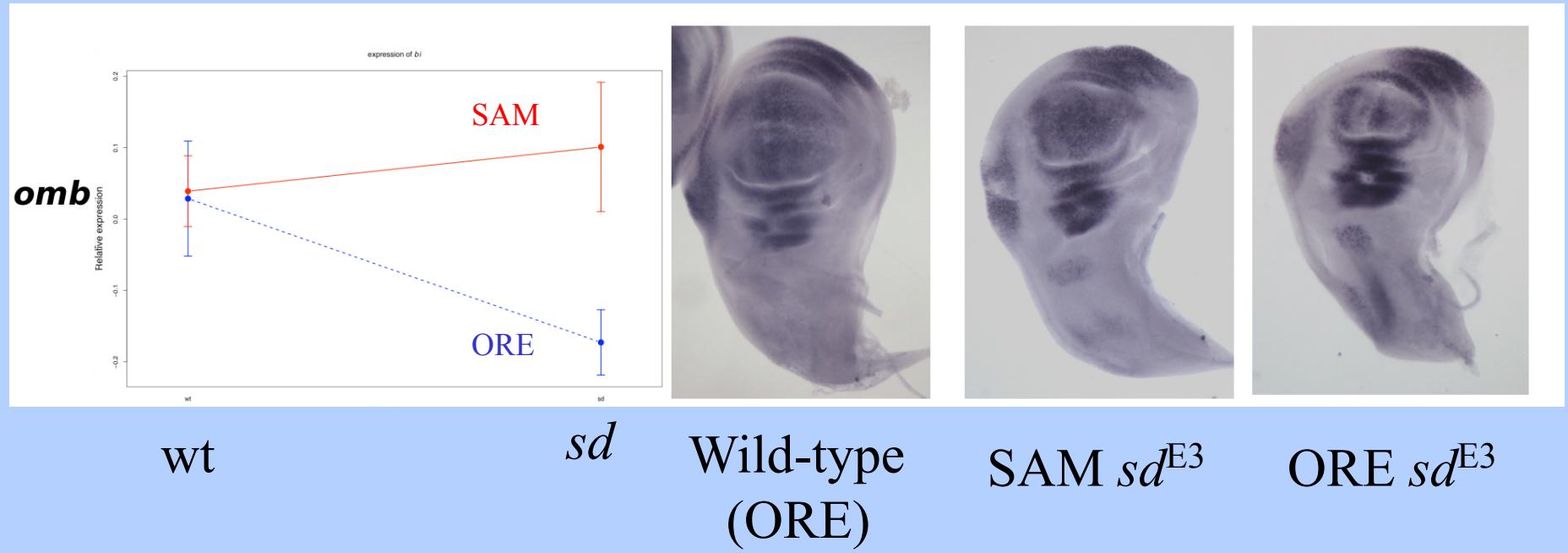
# Background effects of *sd* appear to act on developmental regulatory genes



# Examining the expression consequences of Sd target genes



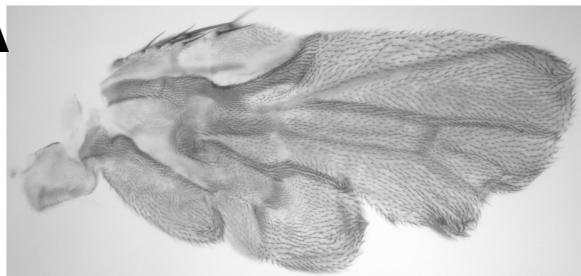
# Genetic background effects: higher order genetic effects- an interesting observation



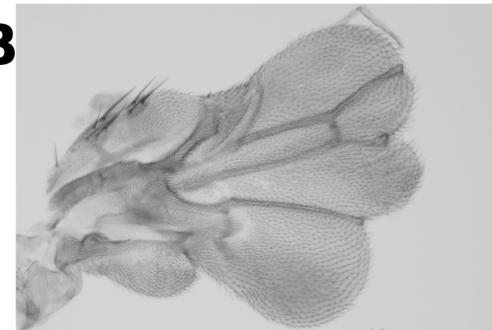
# The effect of background on “higher order” genetic effects (epistasis)

***SAM w omb[md] sd[E3]***    ***ORE w omb[md] sd[E3]***

**A**



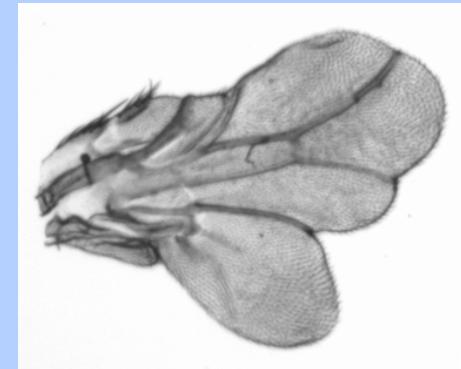
**B**



**SAM *sd*<sup>E3</sup>**



**ORE *sd*<sup>E3</sup>**

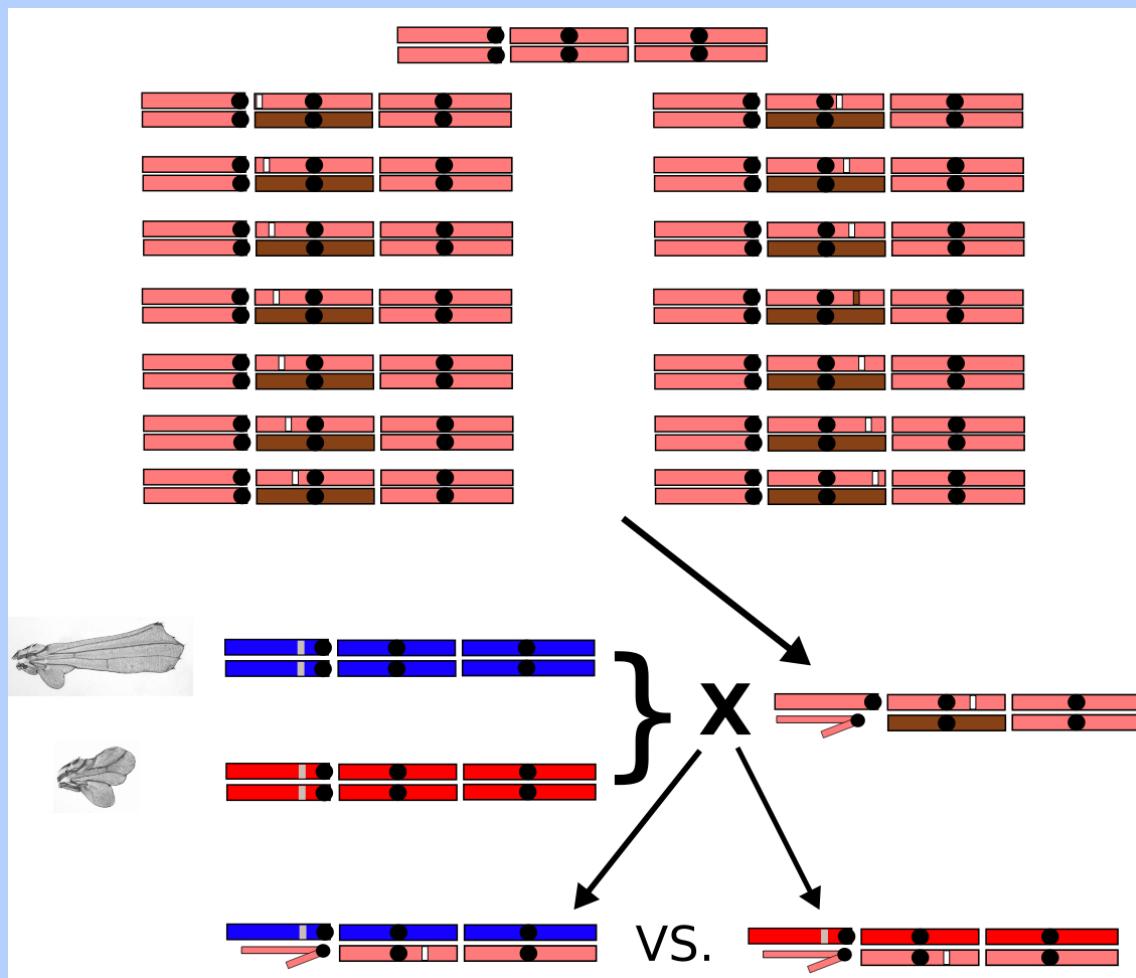


Dworkin et al. 2009

# Genetic background and inferring genetic networks.

- Many genetic networks are inferred from genetic interaction data such as these.
- If genetic interactions are often background dependent, what are the implications for our estimates of the networks?
- **How common are background specific genetic interactions?**

# The consequence of background on higher order genetic effects

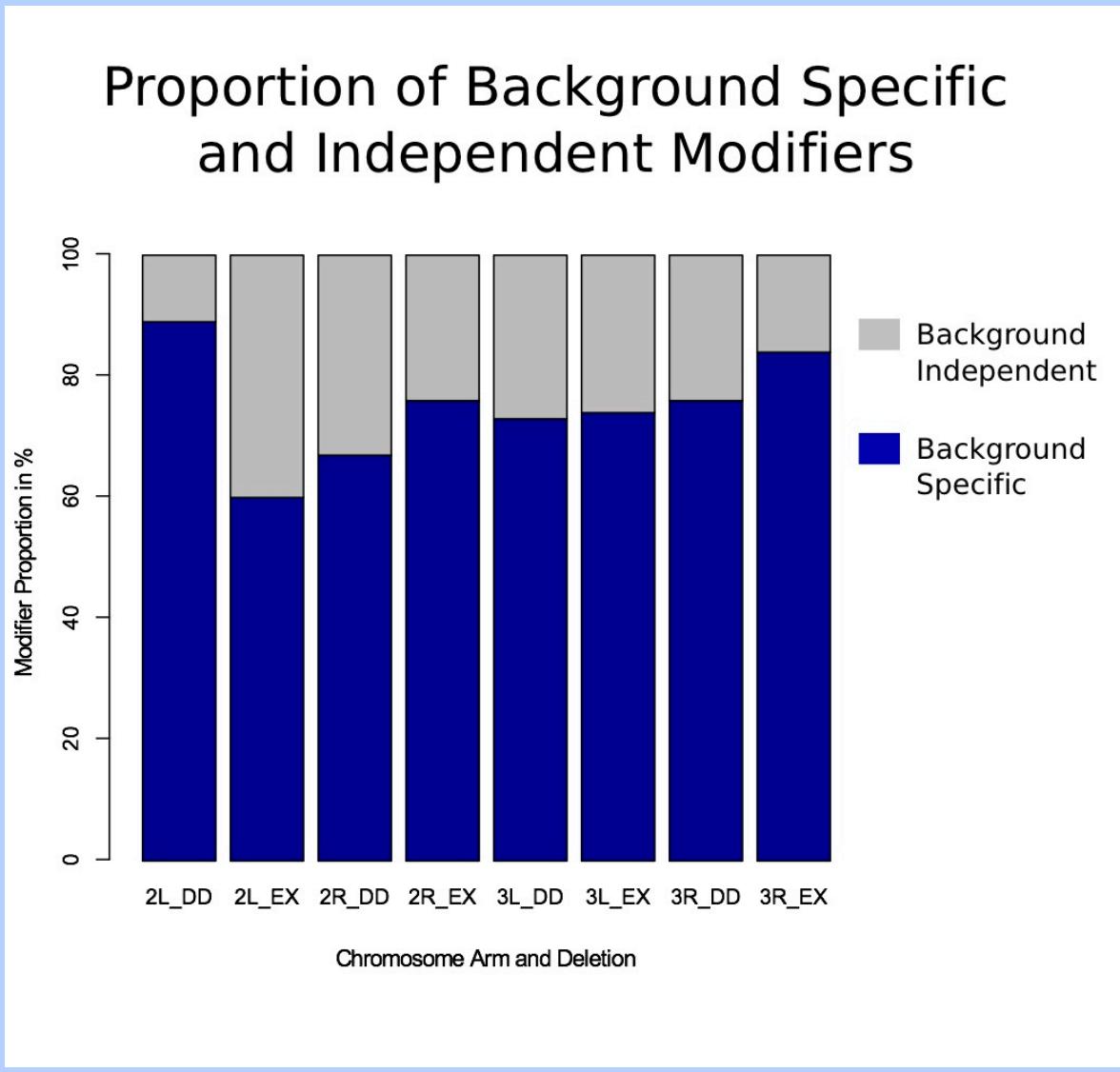


Used a panel of small genomic deletions in an otherwise isogenic background.

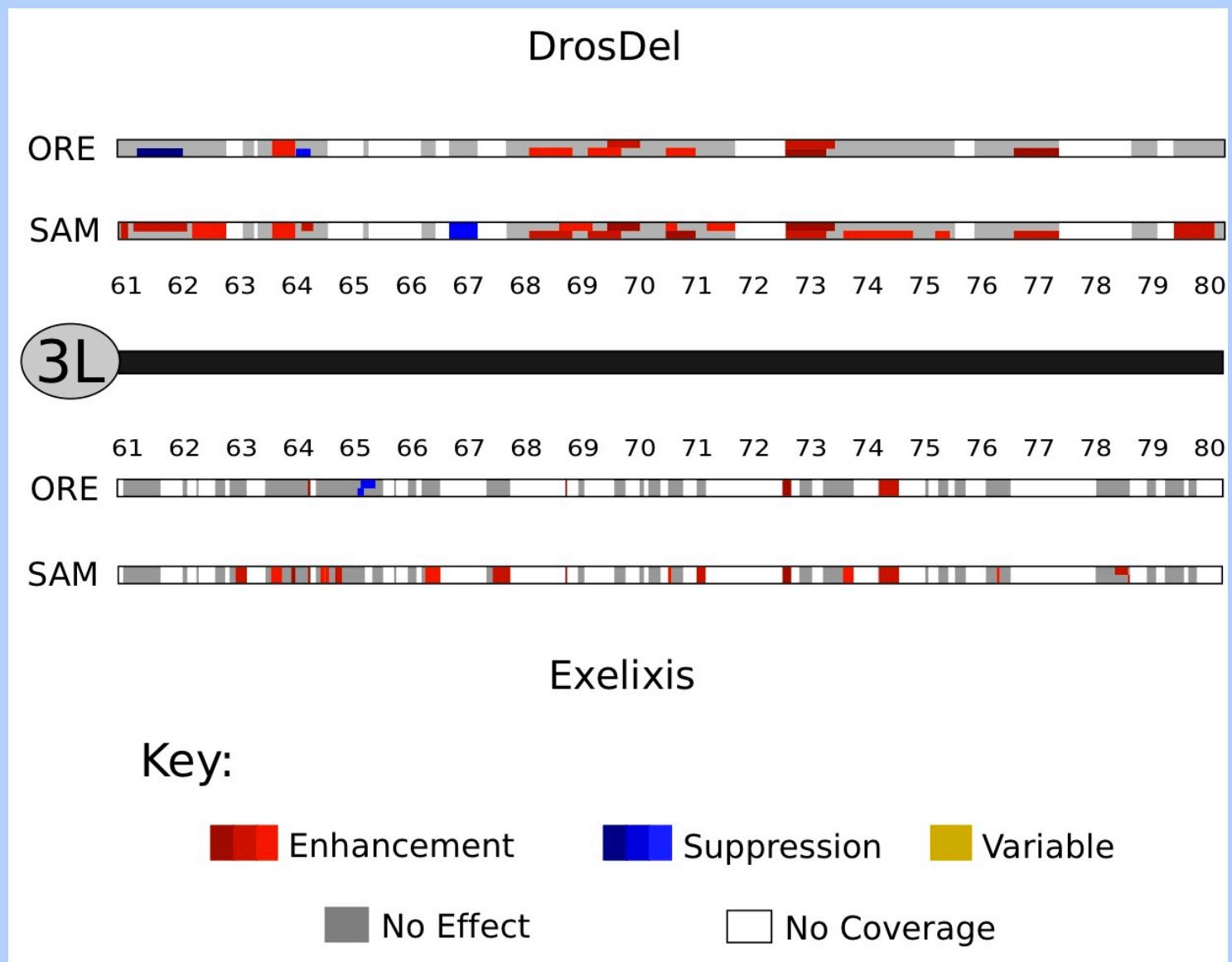
How many dominant modifiers will be background dependent?



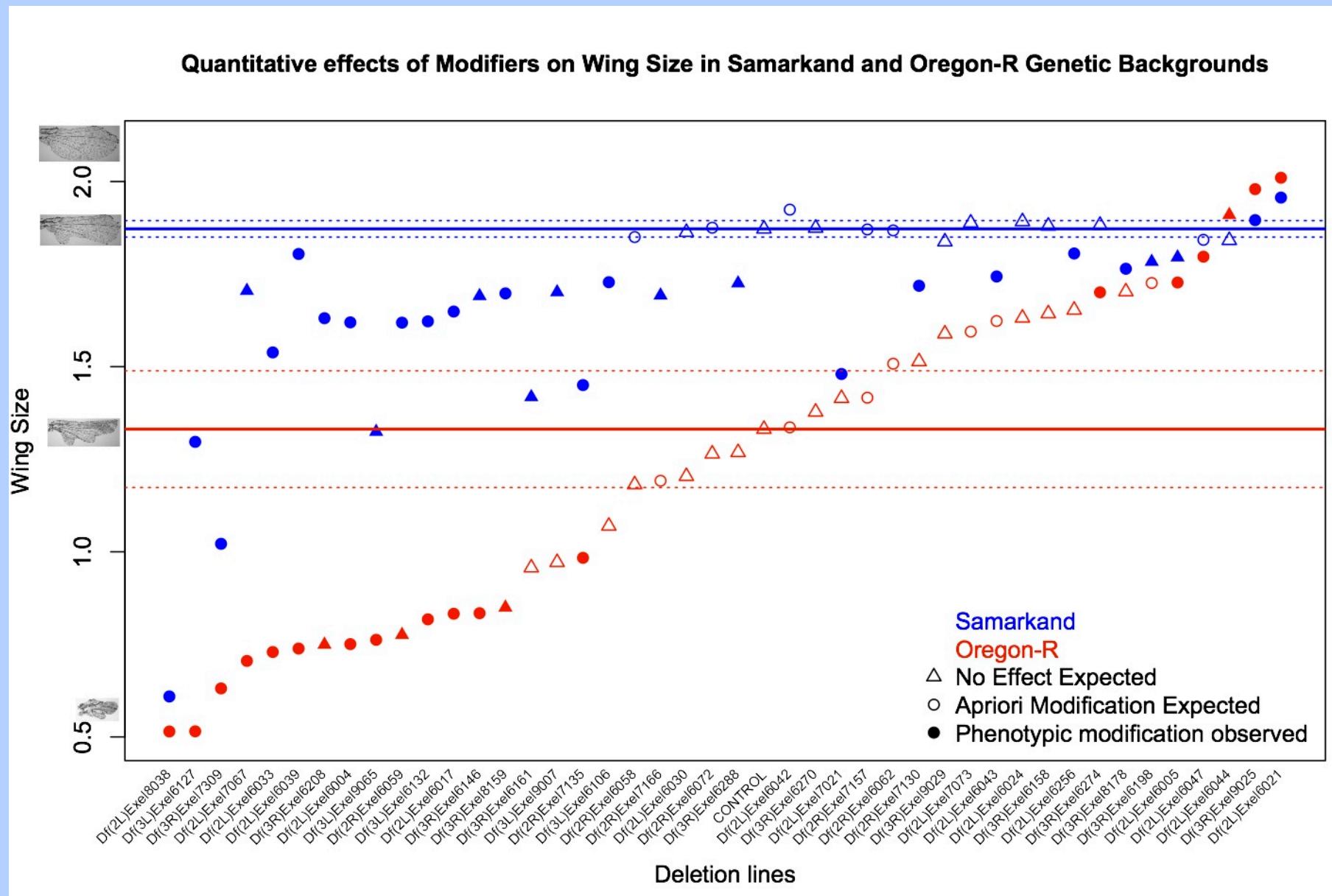
*The majority of dominant modifiers show background specific genetic interactions!!!*



# Modifier Map for Sam $sd^{E3}$ & Ore $sd^{E3}$ in DrosDel & Exelixis for Chromosome 3L



# Differential sensitivity to mutational perturbation between wild-type strains.



# The influence of genetic background on the outcome of genetic interactions is substantial.

- While genetic interaction studies, and in particular modifier screen will (and should) remain an important approach to deconstructing genetic networks, there is an additional “dimension” that should be considered.
- It is likely that there is considerable flexibility in such networks.

# Genetic mapping of background modifiers for *sd* expressivity,

- We generated a panel of markers between the progenitor strain and generated an initial map.
- Two regions of the genome were linked to the background effects.
- How best to fine map these loci?

The conditional effects of mutations: How does genetic background effect our understanding of “higher-order” genetics, such as the ordering of allelic series?

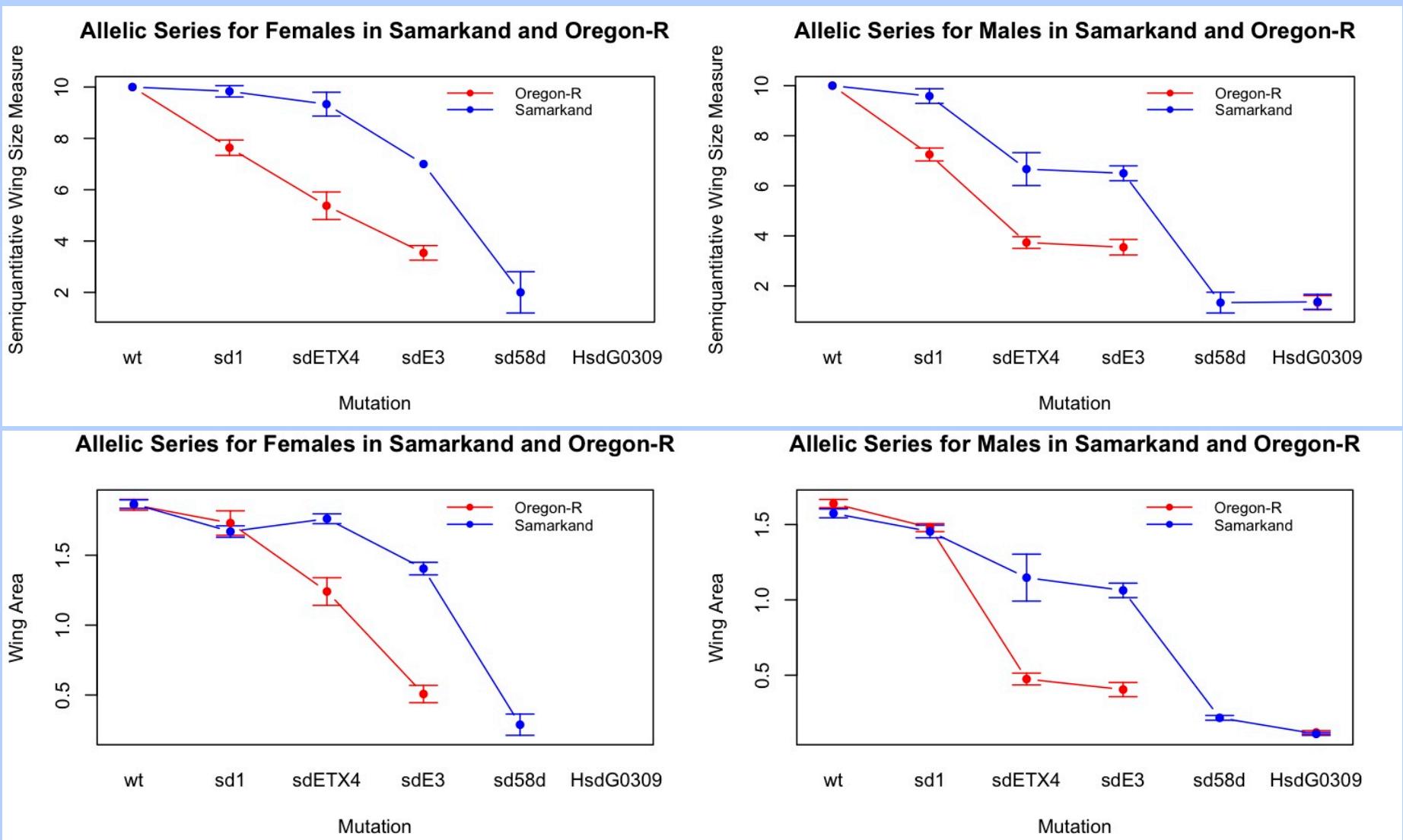
# The ordering of allelic series

- In traditional genetic analysis multiple alleles within genes are “ordered” with respect to the severity of their phenotypic effects.
- Wild-type > allele1 > allele2 > null allele.
- Are such orderings conditional?

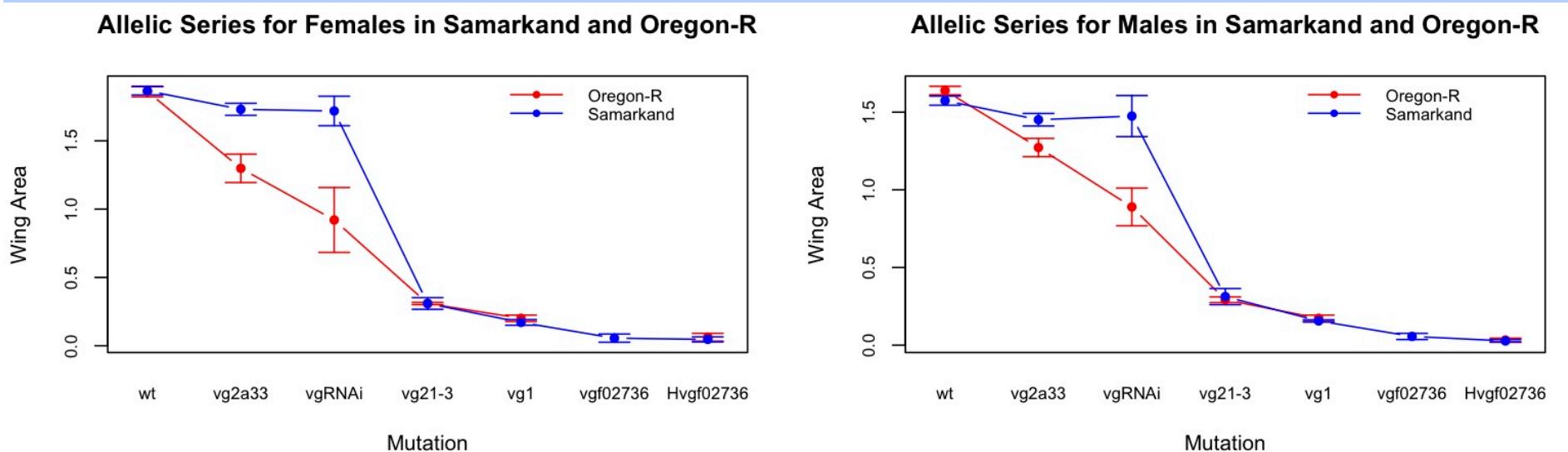
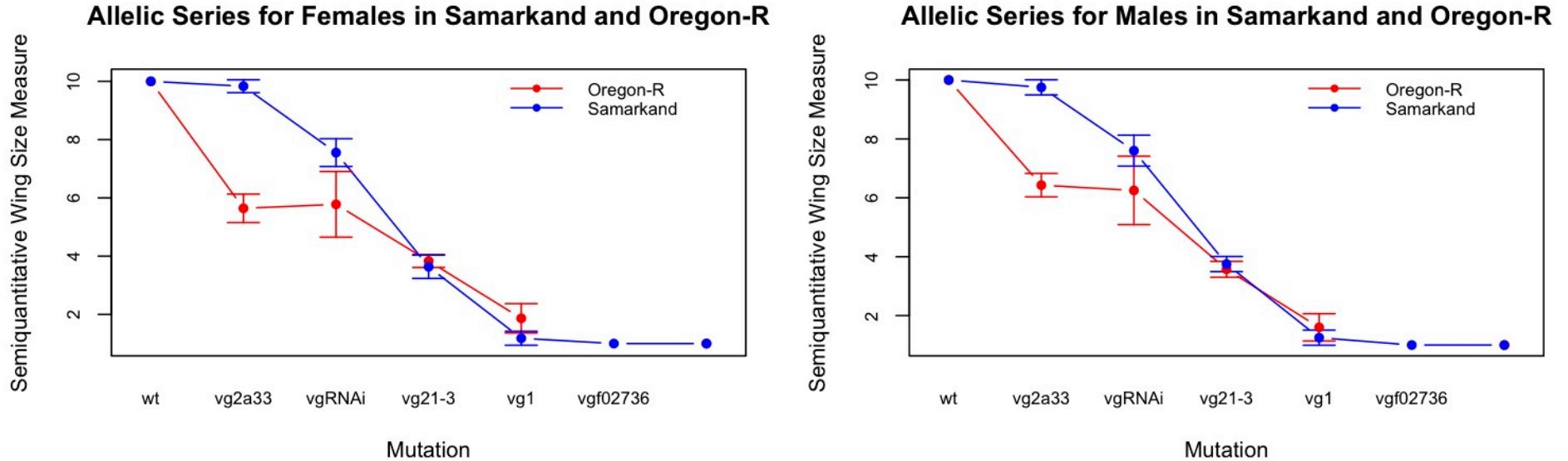
# The “BIG CROSS”

- Our lab has just completed a large study investigating the joint consequences of genetic background, larval nutrition and rearing temperature on the ordering of allelic series, patterns of intra-genic complementation and epistasis (inter-genic interactions).
- Here is a taste....

# The influence of genetic background on the ordering of alleles of *scalloped*.

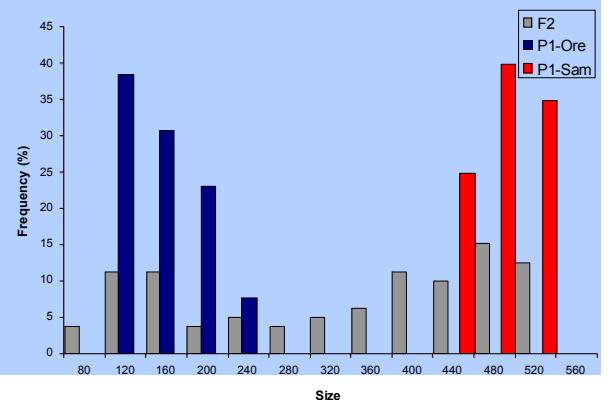


# The influence of genetic background on the ordering of alleles of *vestigial*.

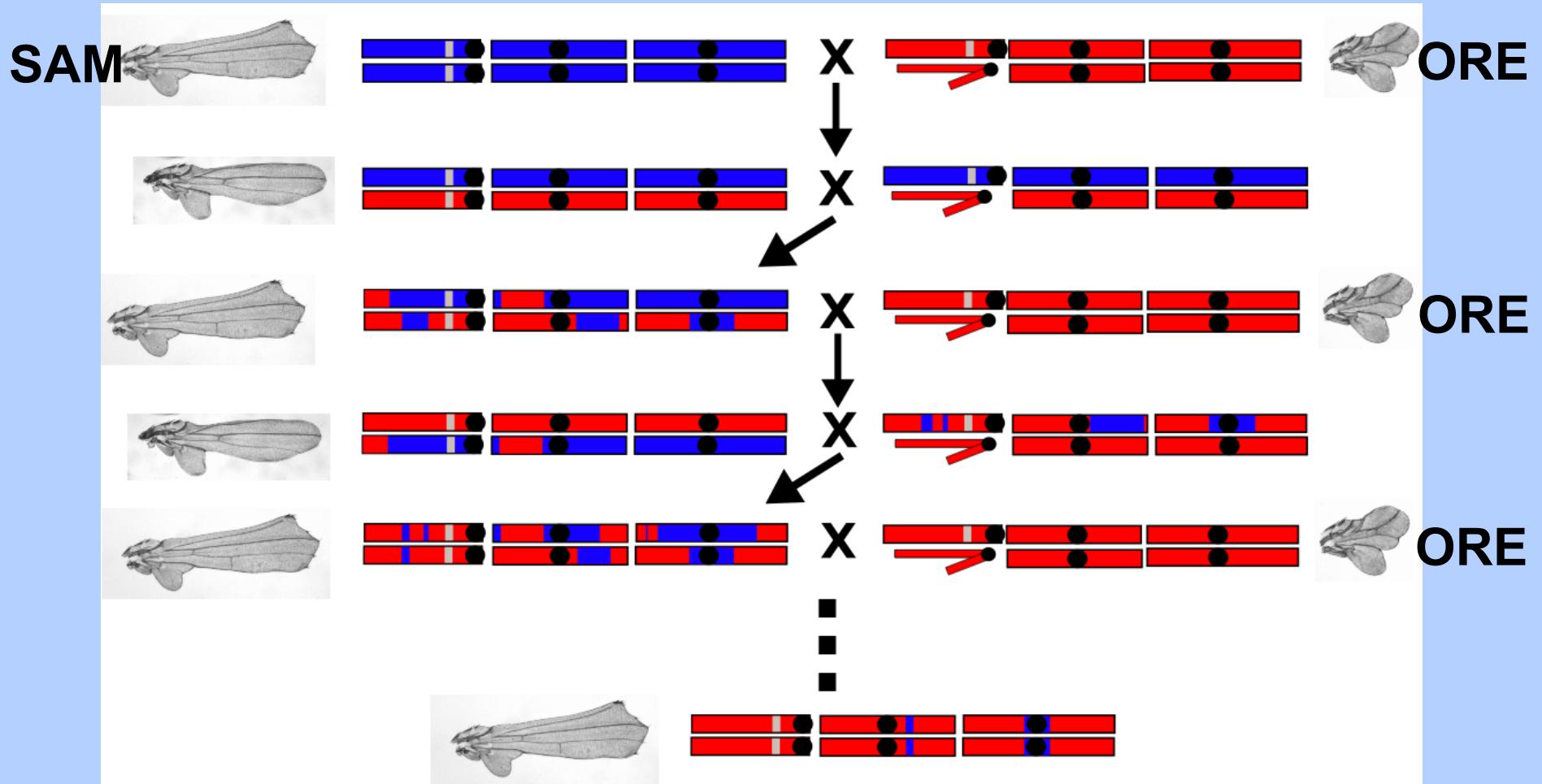


# Mapping of genetic background modifiers using “Next-Gen” Sequencing

- Generated > 50,000 F2 individuals for mapping population.
- Using extremes of the distribution we pooling individuals to perform a Bulk segregant analysis (BSA).
- Sequenced the pools of the long and short winged extremes.

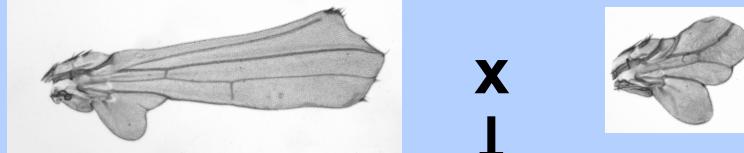


# Backcross Design

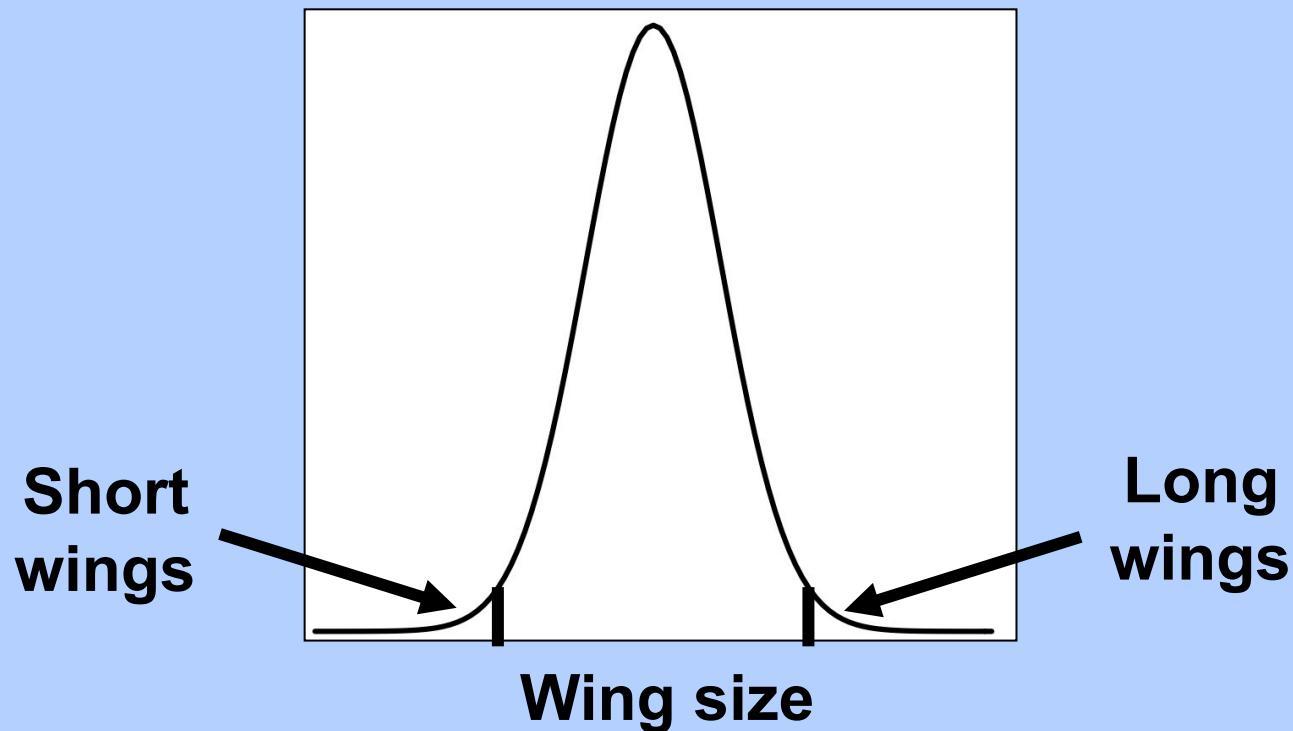


Long wings but mostly ORE (short wing) genome  
Next-gen sequencing: which loci still have SAM alleles?

# F2 Mapping Design

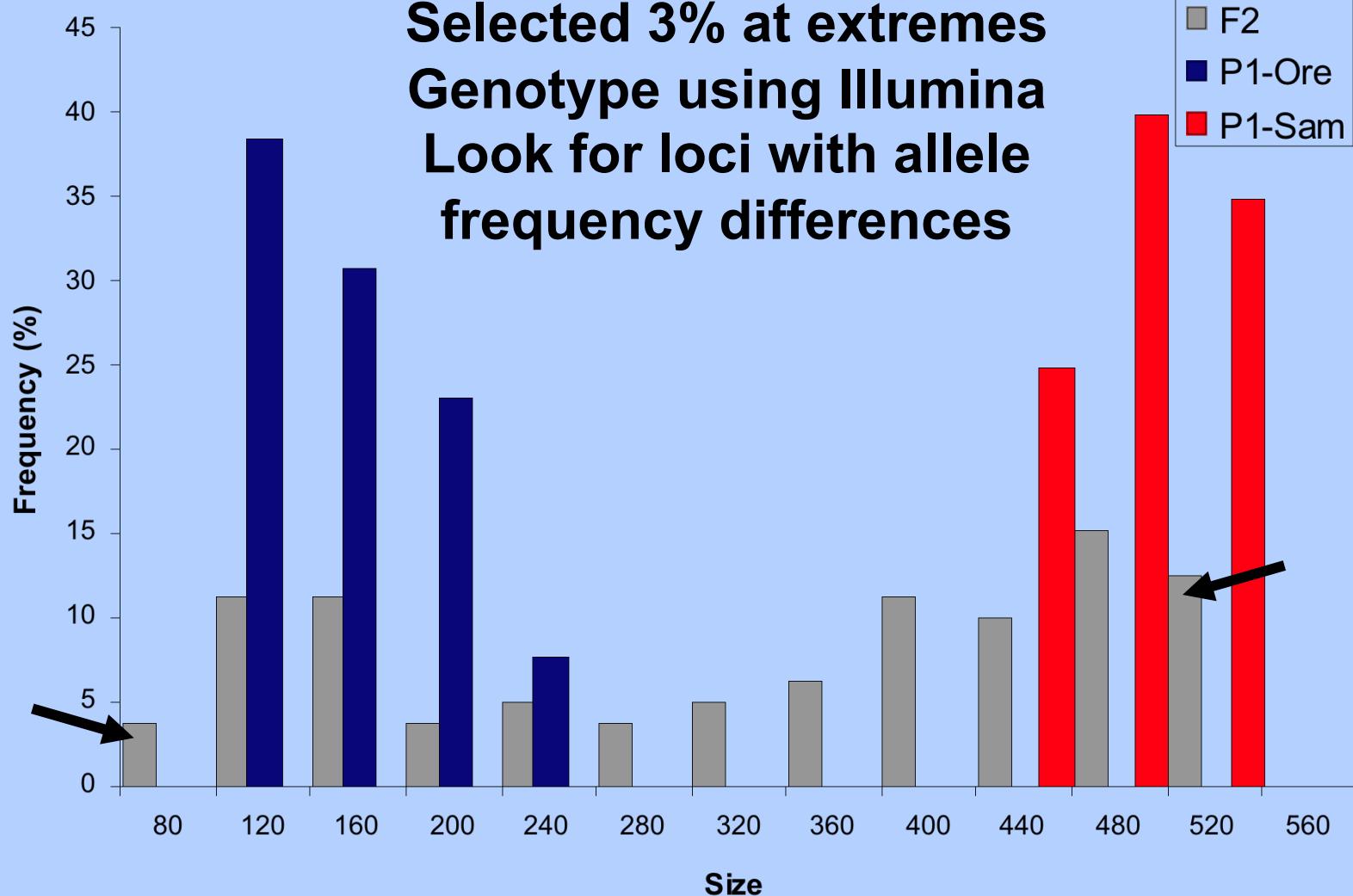


X  
↓  
F1  
↓  
F2

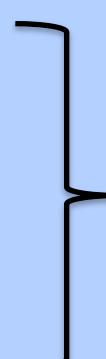


# F2 Mapping Design

~50,000 F2's screened  
Selected 3% at extremes  
Genotype using Illumina  
Look for loci with allele frequency differences



# Sequencing

- ORE & SAM
  - 1 lane each:
    - 1 long backcross population (paired end)
    - 3 short backcross populations (paired end)
    - 2 long F2 bulk segregant populations (single end)
    - 2 short F2 bulk segregant populations (single end)
    - ORE & SAM (paired end)
- 
- Multi-plexed

# Sequence Analysis

- What do we want our data to look like?

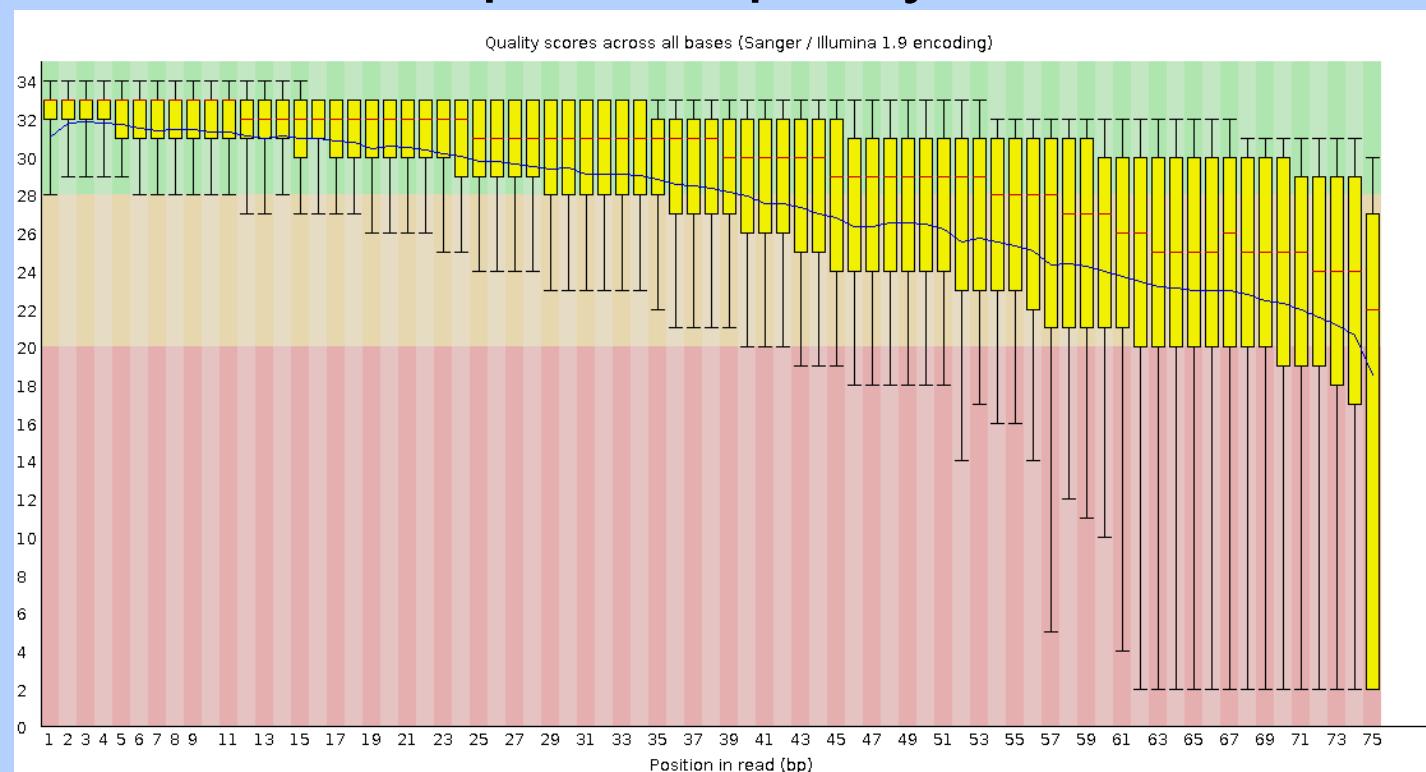
Chrom.	Pos.	Long-1 #ORE	Long-1 #SAM	Short-1 #ORE	Short-1 #SAM
X	150	8	0	0	6
X	227	12	1	0	17
X	283	0	22	14	1

# Sequence Analysis

- Clean/trim & align all reads to reference genome
- Call SNPs
- At each SNP, for each sample, count number of reads with ORE allele and number of reads with SAM allele

# Sequence Analysis

- Clean/trim:
  - FastQC
  - Script to truncate reads from F2 populations due to low sequence quality



# Sequence Analysis

- Align
  - bwa
  - Up to 4 mismatches for 76-bp reads

# Sequence Analysis

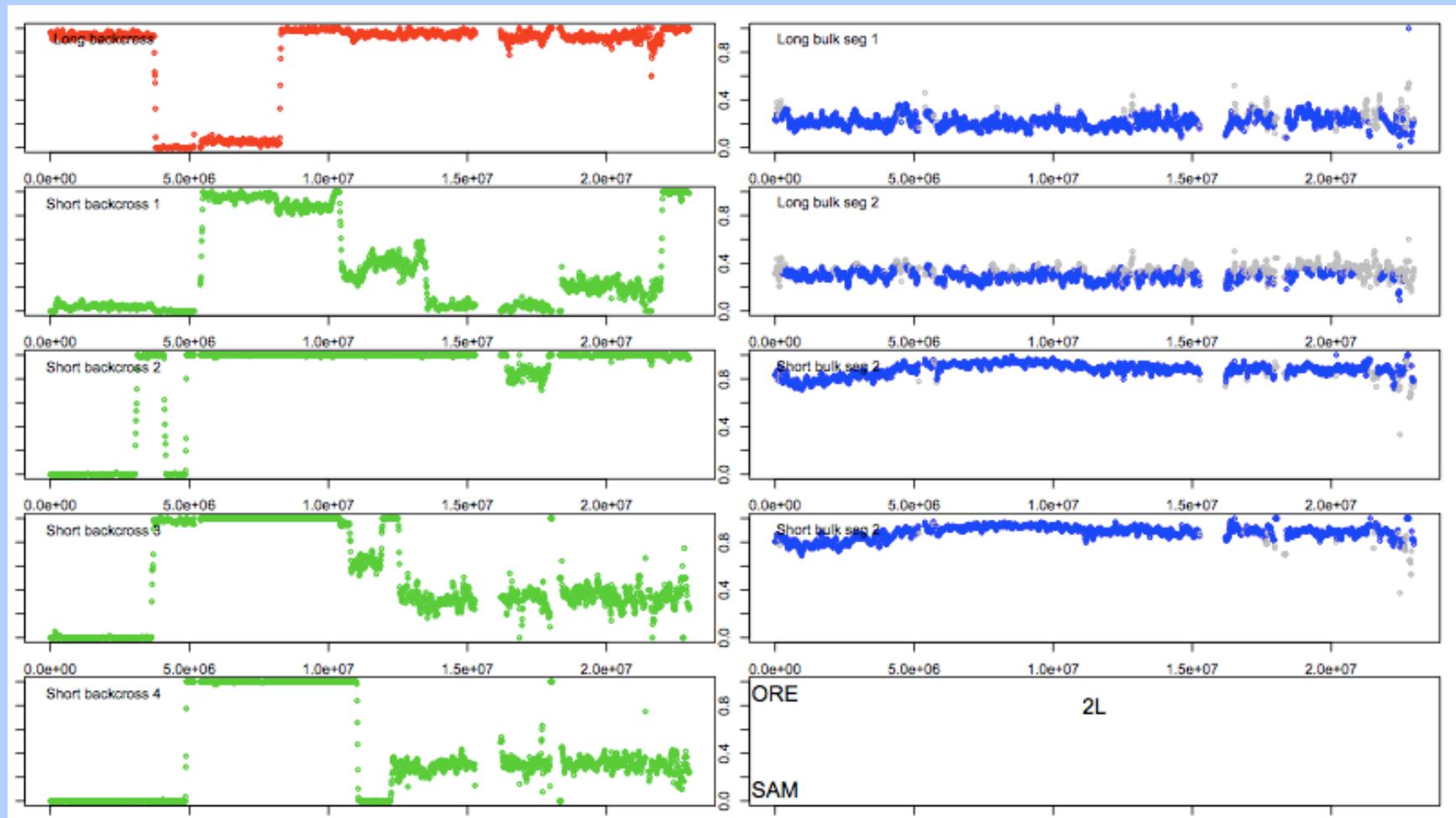
- Call SNPs
  - Samtools
  - Then generated pileup at SNP loci

# Sequence Analysis

- Counting alleles/reads
  - Python script
  - Keep only SNPs at which ORE & SAM are homozygous and different from each other
  - For each experimental line (i.e., backcross or F2 population), at each SNP, figure out whether each read represents ORE or SAM, and then count how many of each

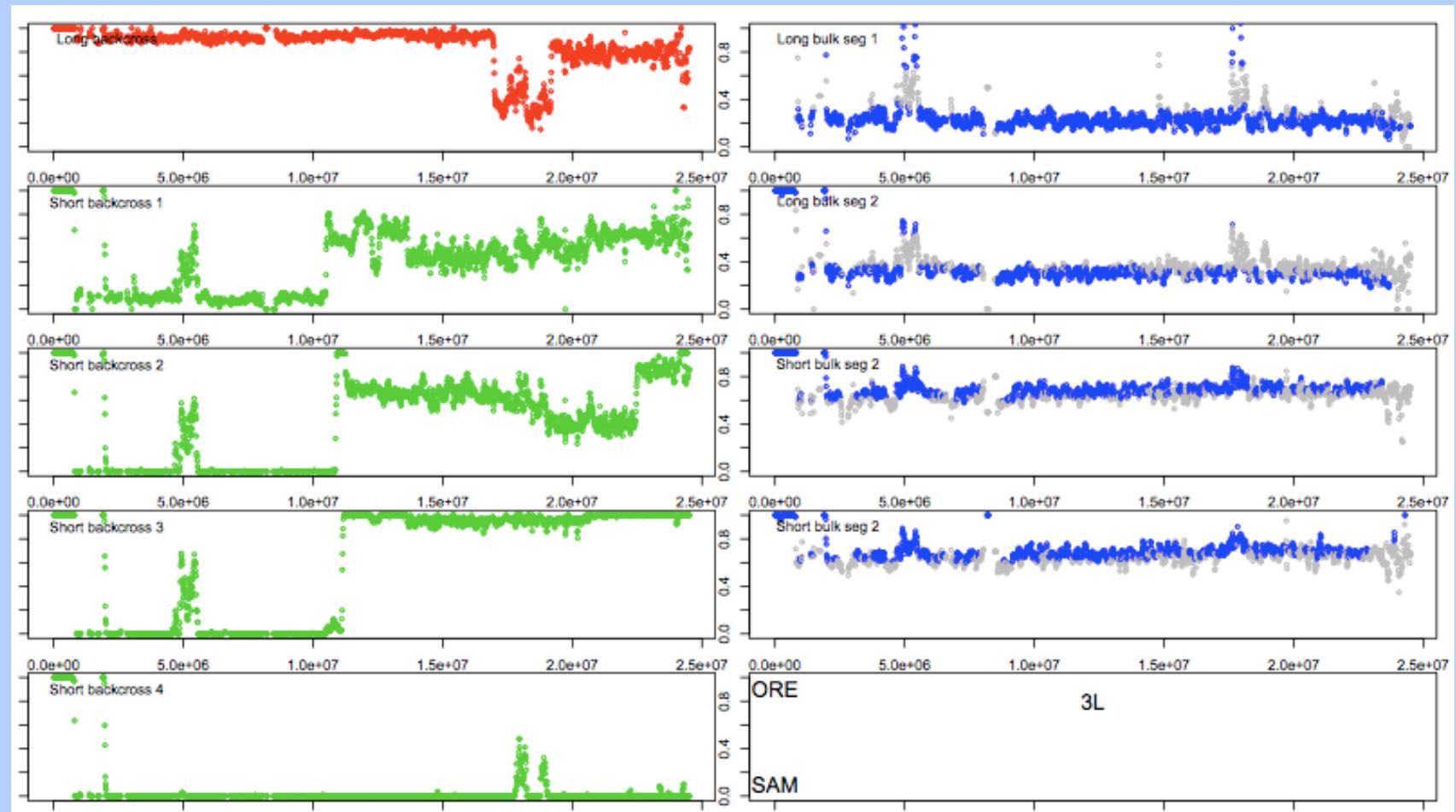
# Sequence Analysis

- “Analysis” (plotting)
  - R



# Sequence Analysis

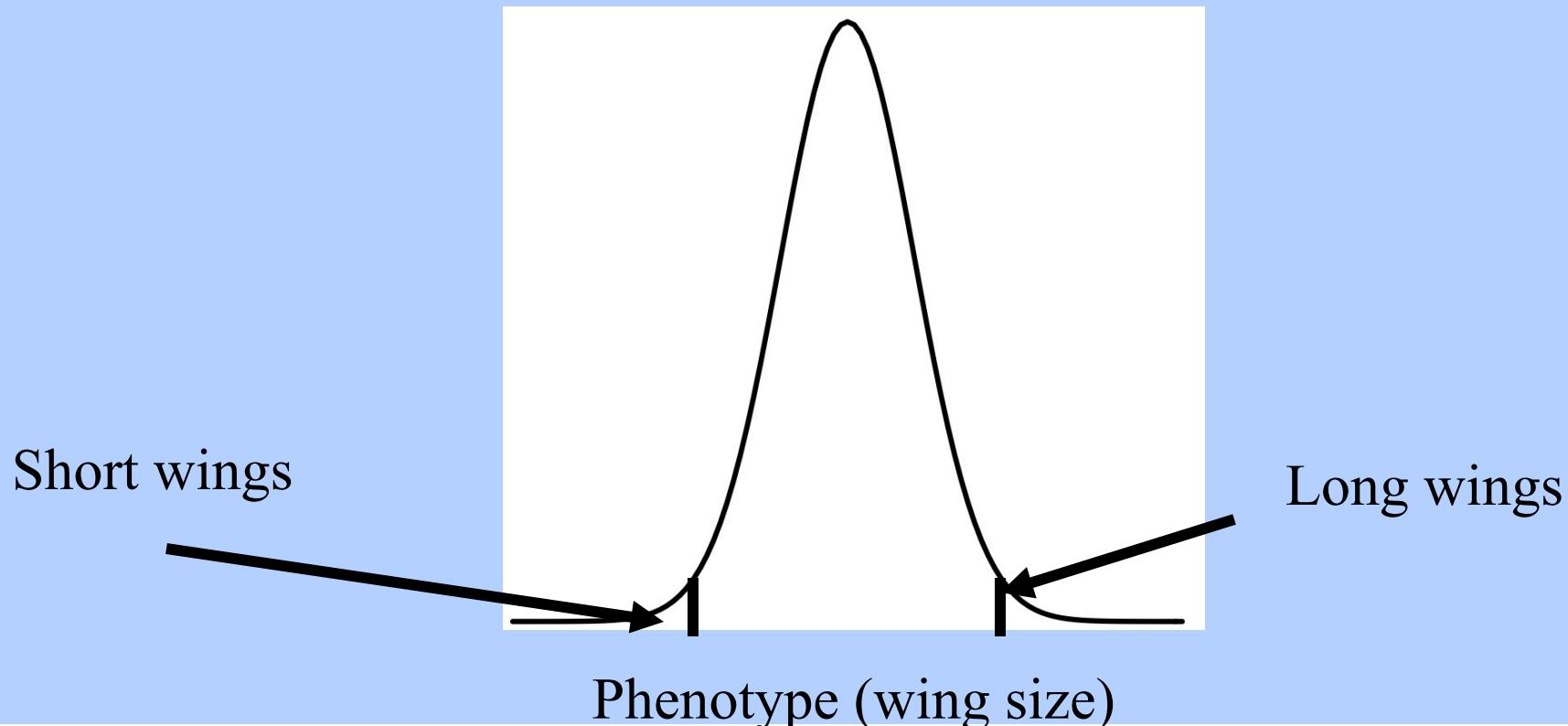
- “Analysis” (plotting)
  - R



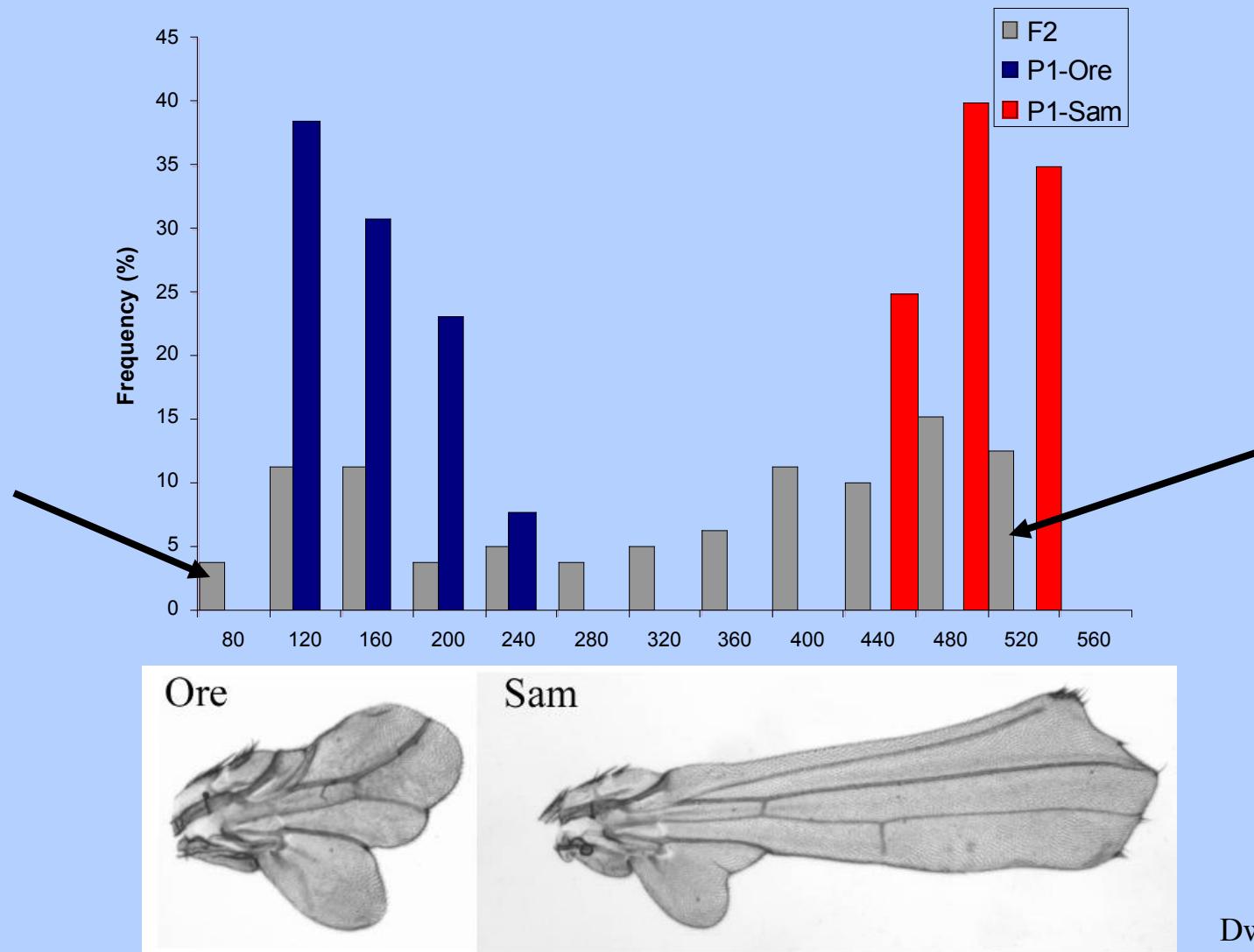
Onto the tutorial...

# Bulk segregant analysis

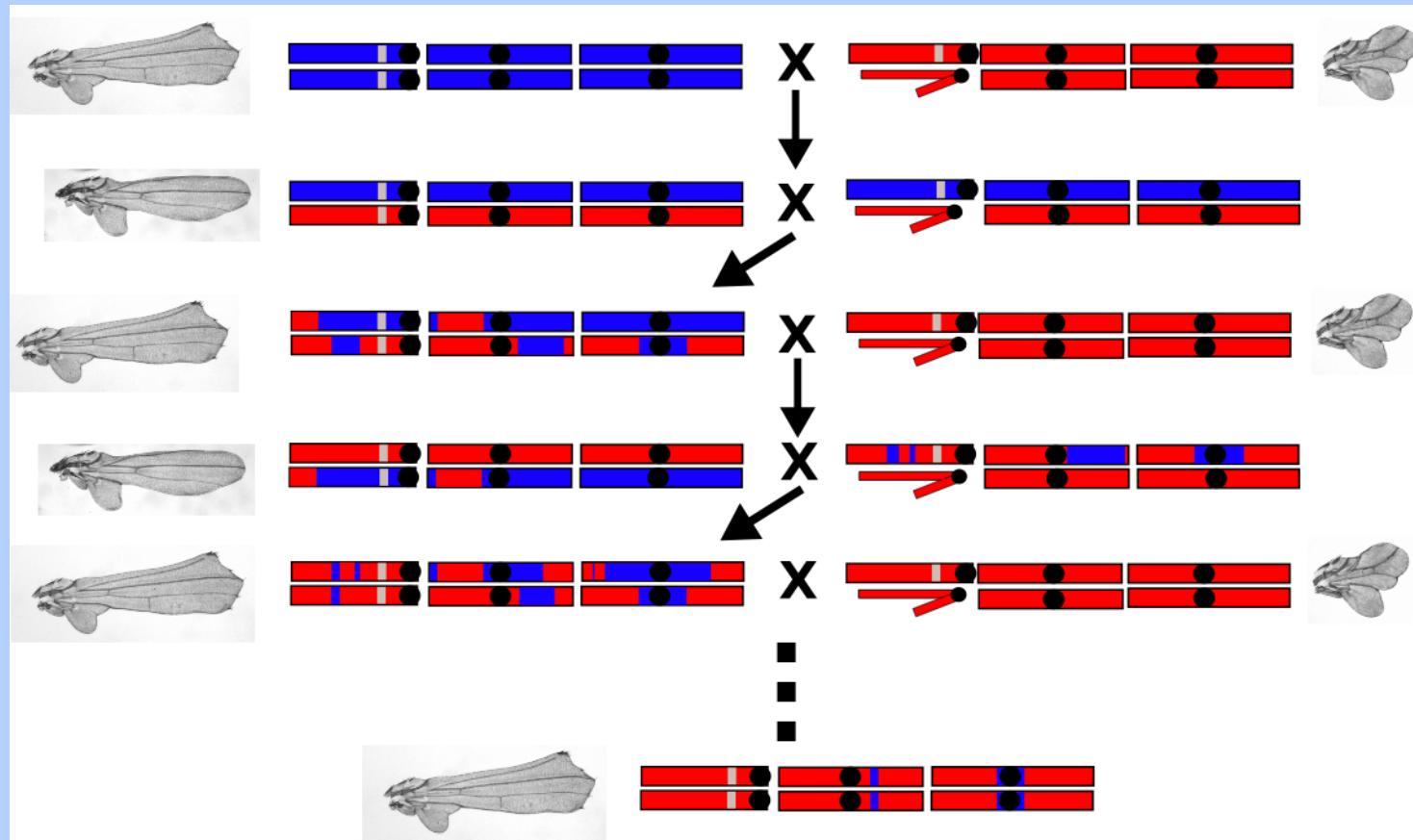
- 1: Generate a panel of F2 (or any intercross design) individuals that vary for the phenotype.
- 2: Select individual at the extreme of the distributions
- 3: Genotype select individuals only, in pools.
- 4: Look for regions of the genome with frequency differences for alleles that deviate from genome wide average.



# Genetic mapping using the distribution of the F2's



# Mapping using Backcross-Selection



The genetic background modifiers are linked to at least 2 regions (in blue). We are currently fine mapping these modifiers.

2R modifier: Linked to, but **unlikely vestigial**.

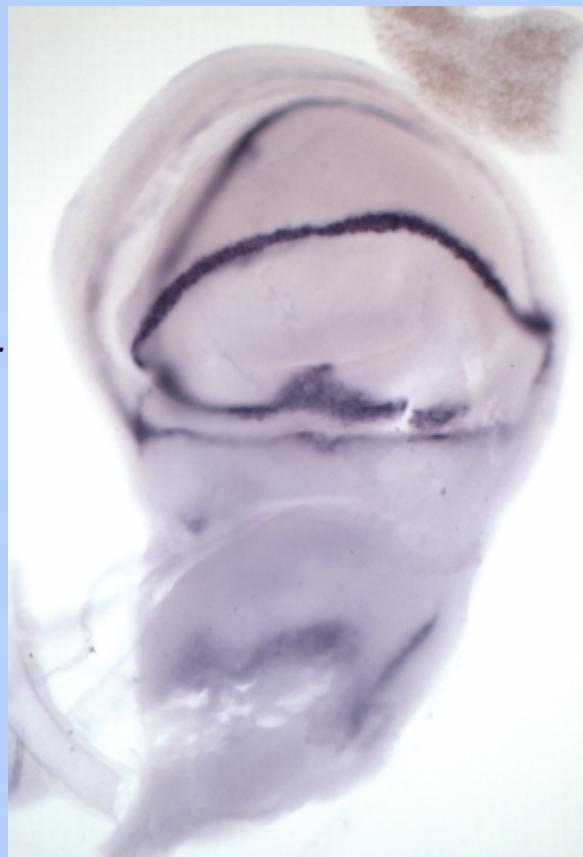
# BSA for background modifiers

- We generated ~ 50000 F2 individuals.
- Selected ~ 3% of flies that were longest or shortest for wing size.
  - Pooled “long” flies, Pooled “short” flies.
  - Illumina re-sequencing of DNA for pools
  - 2 Lanes for each pool, for 76bp reads. ~20 Million reads/lane
  - Drosophila has a 180Mb genome.
  - Total coverage ~ 33X
  - Once trimming and successful mapping consider coverage is ~ 25X

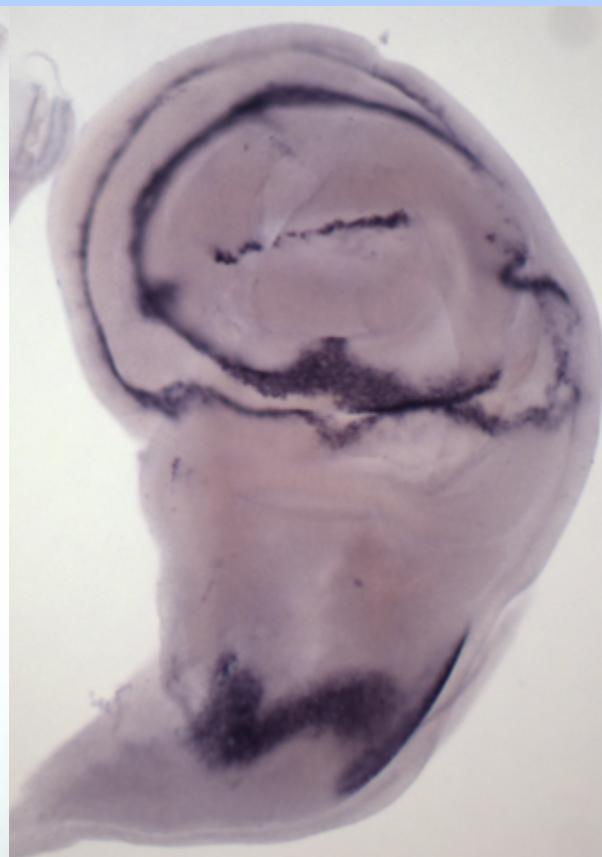
# wingless

*wg*

Sam wt



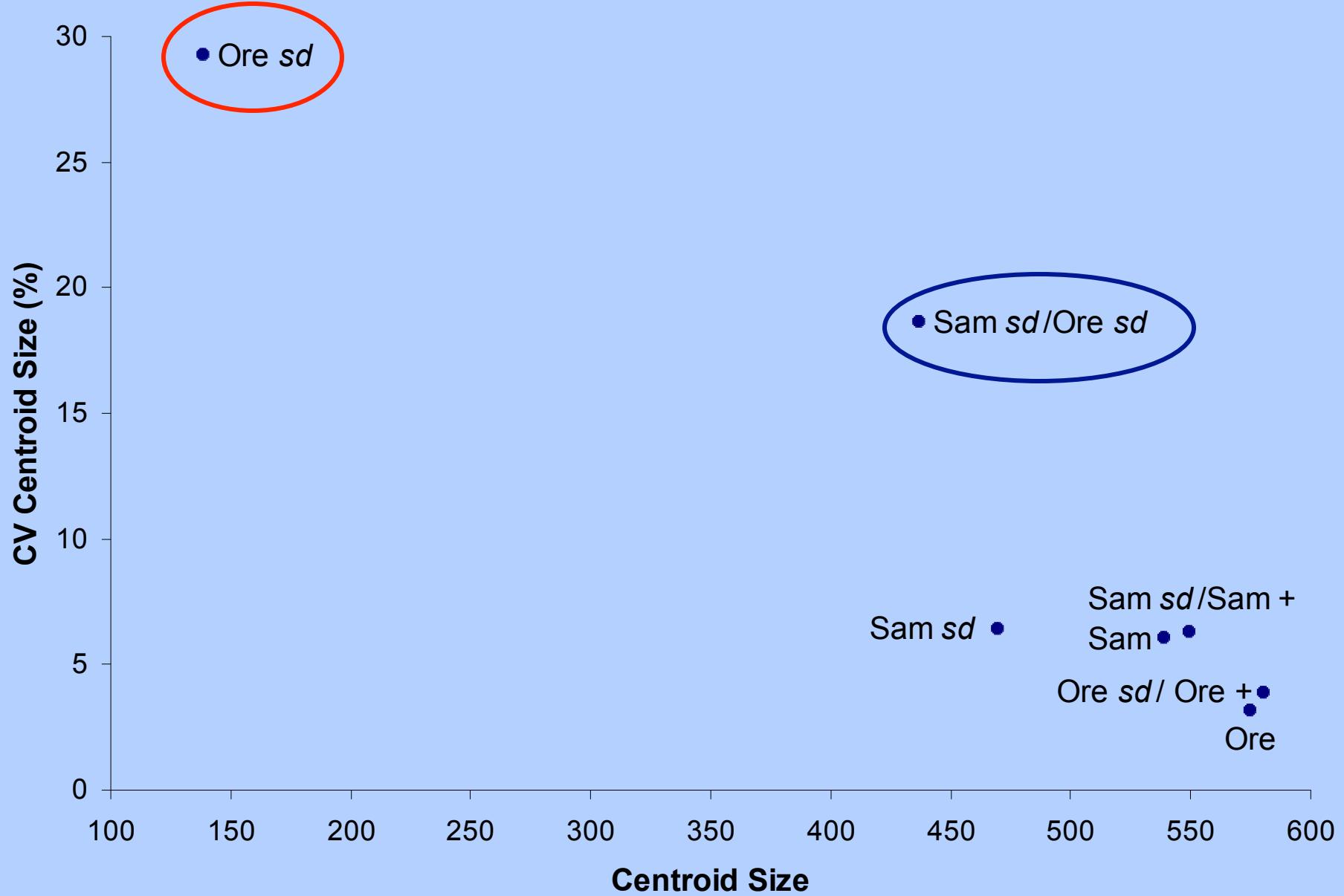
Sam *sd*



Ore *sd*



# Separation of mean and variance?



# Genomic analysis of Genetic background effects

- We have now generated allelic series of mutations of *sd* and *vg* (among other genes) in each genetic background.
- We will perform transcriptional profiling to test competing models for the mechanism generating background effects.

# Current and Future work

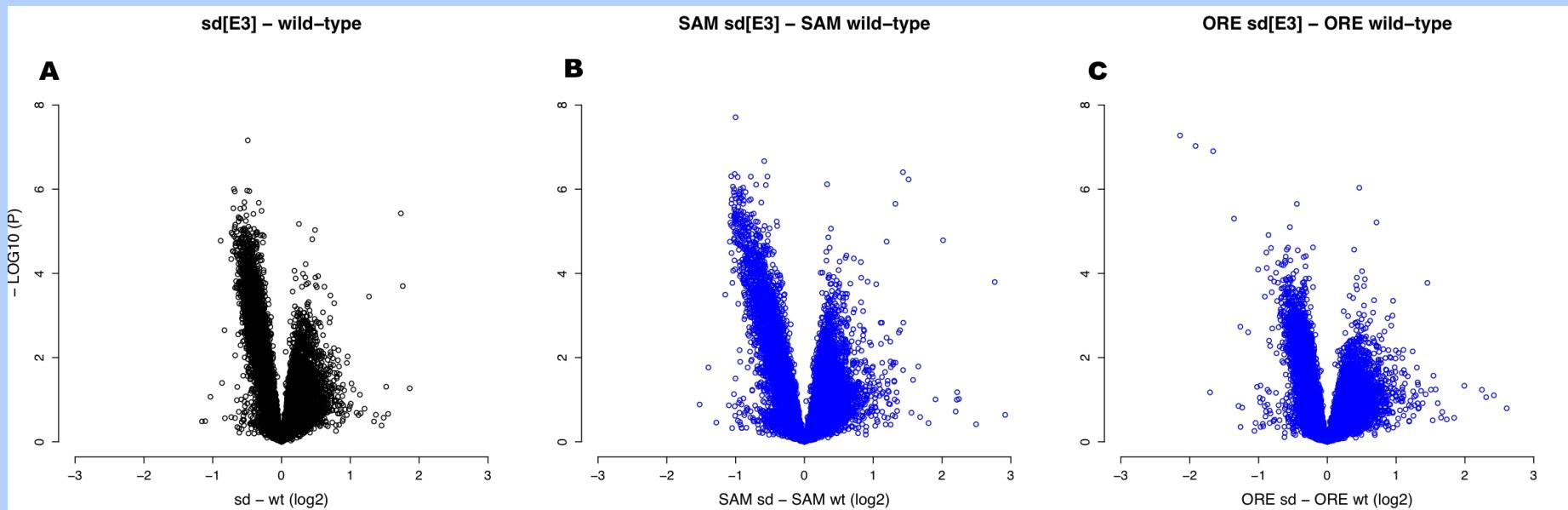
- Mapping the variants responsible for the genetic background effect. *What is their phenotypic effect in natural populations?*
- What are the consequences of genetic background effects on “higher order” genetic effects such as intra-genic complementation and epistasis?
- Introgression of allelic series for *sd*, *vg* and mutations in other genes involved with the wing to address these questions.

# Models to explain the phenotypic consequences of genetic background.

1. *sd* mutation enhances background specific differences in transcription, mediating the observed phenotypic differences.
2. Background effects involve a set of genes that partially overlap with differentially expressed genes between mutant and wild-type, with quantitative differences in transcription **that correlate with variation for the *sd* phenotype.**
3. Background effects involve a **different set of genes** than those that mediate the main effects of the mutant *sd* allele.

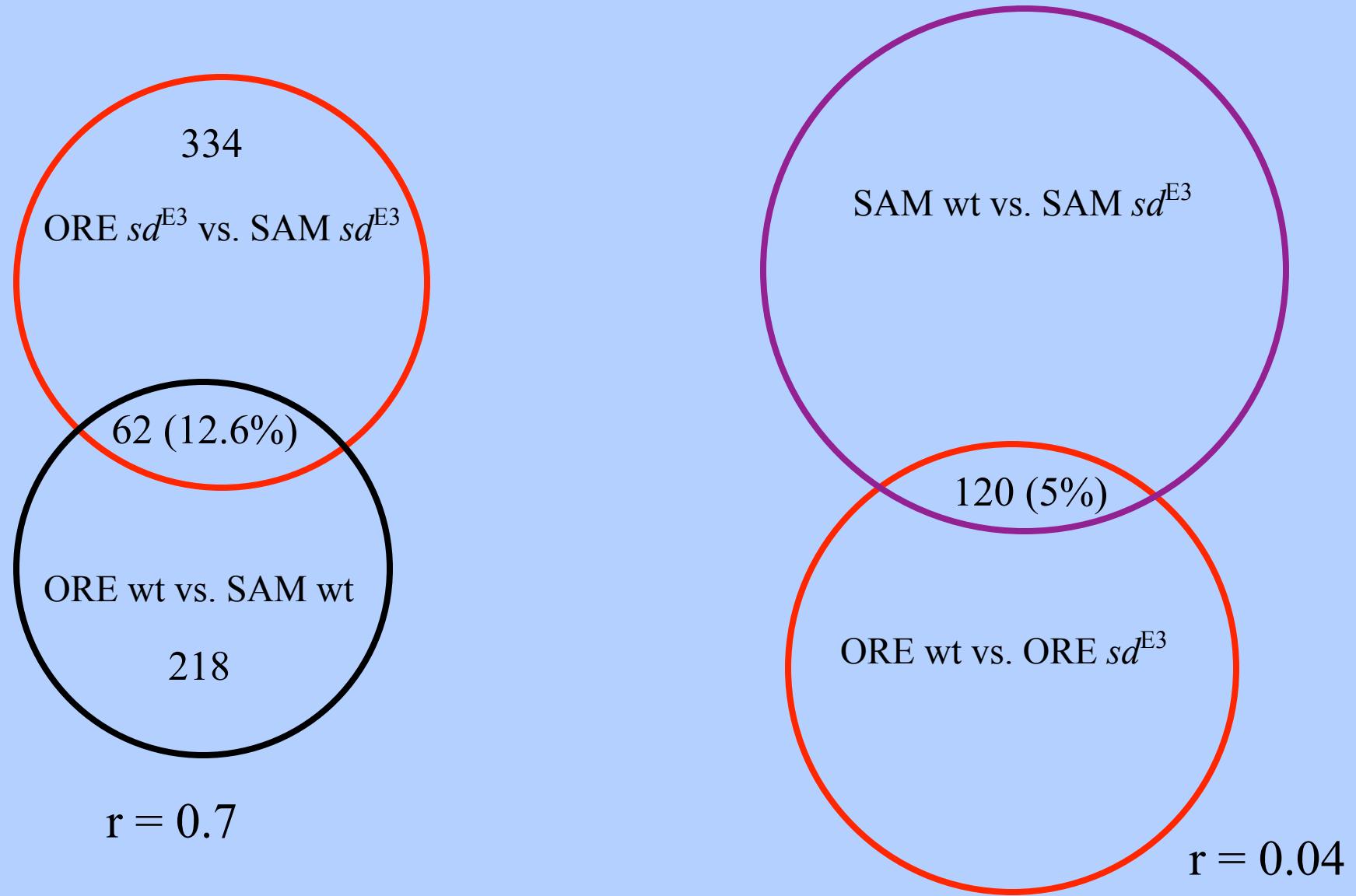
# $sd^{E3}$ vs.wild-type

overall

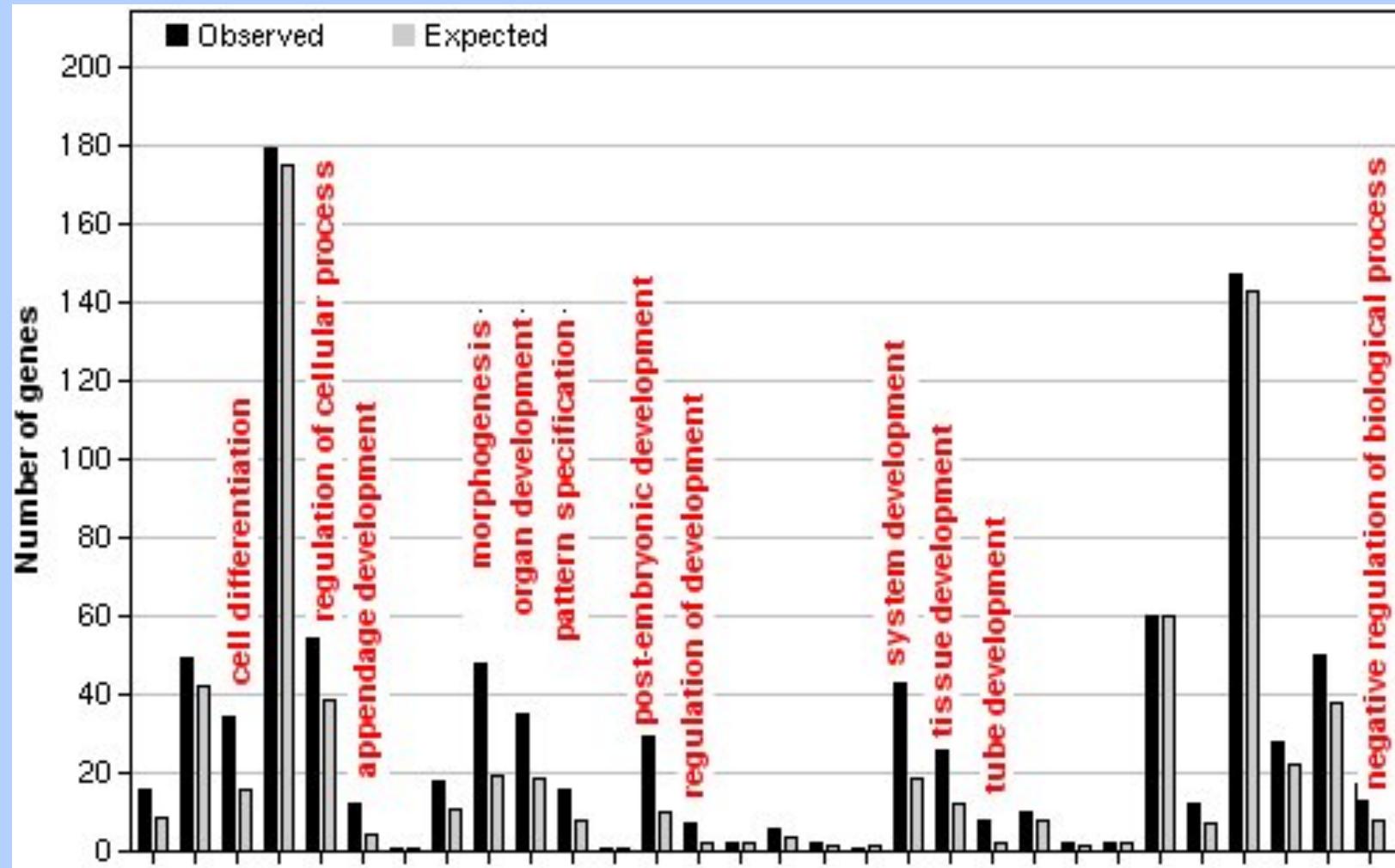


These “volcano plots” demonstrate that the genes (~1000) that are differentially expressed between  $sd^{E3}$  & wild-type, are down-regulated in the  $sd^{E3}$  genotype, but in a background dependent manner.

# Relationships in expression profiles are consistent with models 2&3.



# Background effects of *sd* appear to act on developmental regulatory genes



# Most genetic interactions are background dependent.

- We have scanned the majority of the autosomal regions for evidence of modifiers.
- Evidence for 98 dominant modifiers (enhancers & suppressors).
- **72/98 (73%) of these modifiers are background dependent!!!!**

# Summary for genetic background

- Studying mutations in only one isogenic “wild-type” background may miss important components of the biology.
- Genetic background contributes to “higher order genetic effects”, and not just the expressivity of individual mutations.
- We need to further explore how such effects can influence the evolutionary trajectories of alleles in natural populations.