

# Paradoxical evolution of a large segmental duplication in mouse

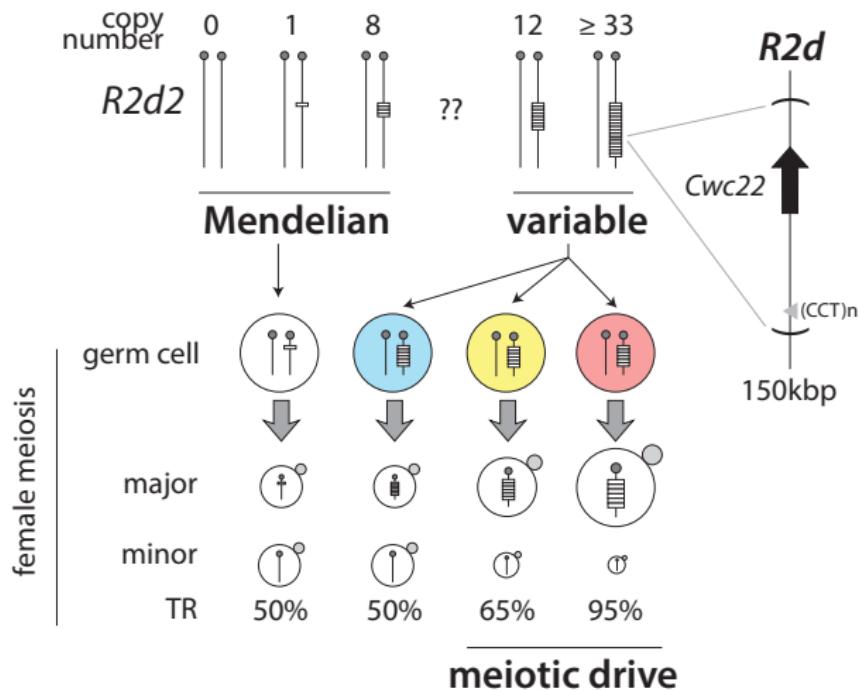
Andrew P Morgan

29<sup>th</sup> IMGC  
Yokohama, Japan

11 November 2015

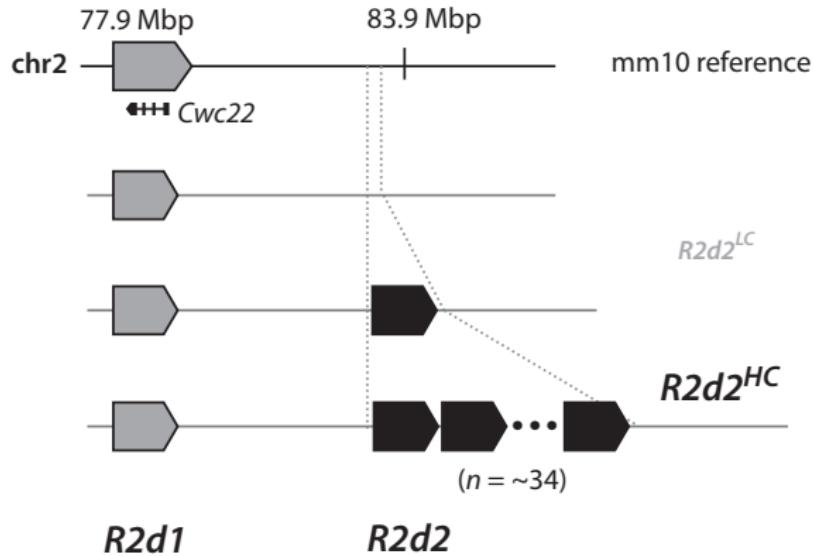


# The *R2d2* meiotic drive locus



Didion JP et al (2015) PLoS Genet

# Structure of the *R2d1-R2d2* locus



# Mutational processes at *R2d2*

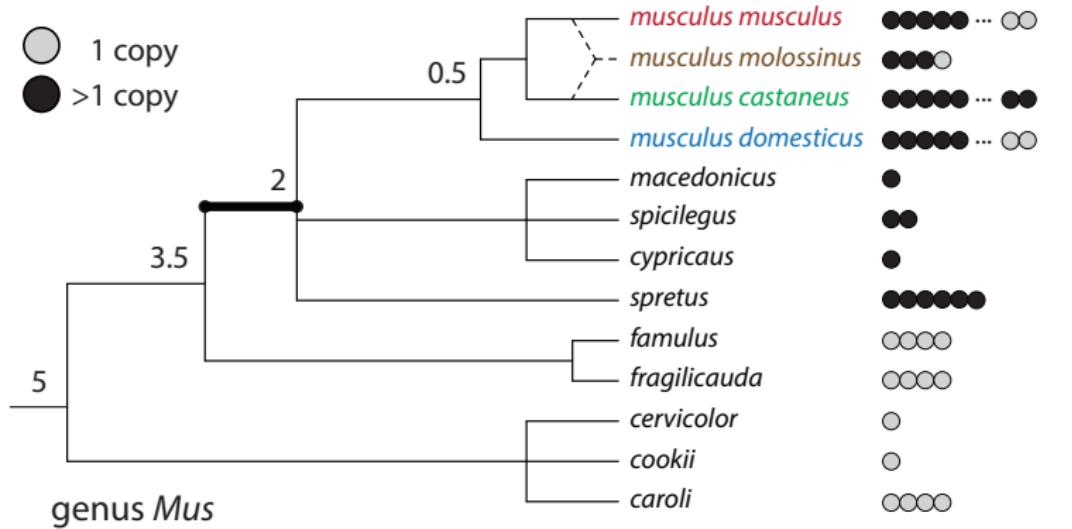
Three surprising observations:

- ① High rate of non-allelic gene conversion
- ② Copy-number dependent suppression of crossovers
- ③ Very high rate of *de novo* copy-number changes

Implications for:

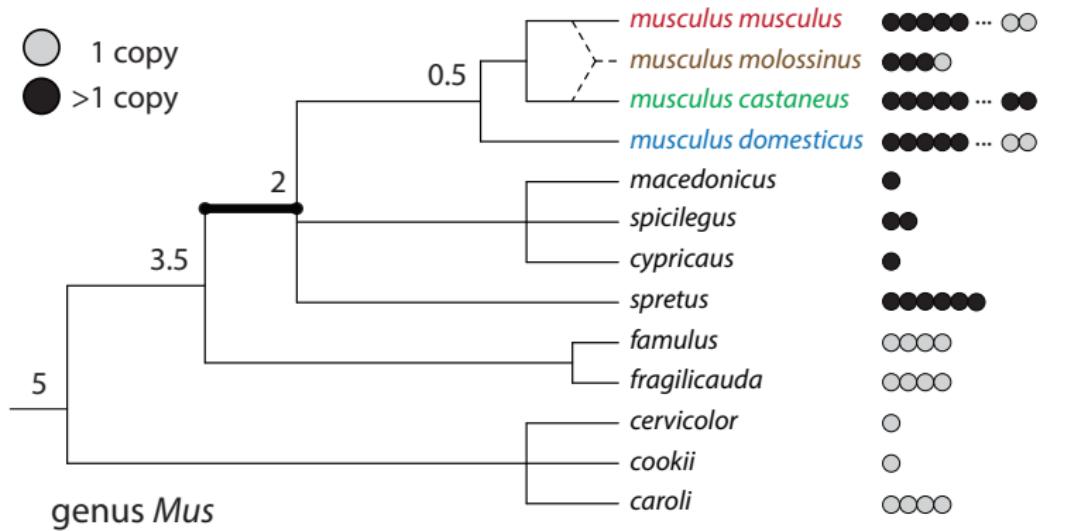
- evolution of a conserved gene involved in mRNA splicing, *Cwc22*
- genetic diversity at linked sites

# R2d2 duplication is old



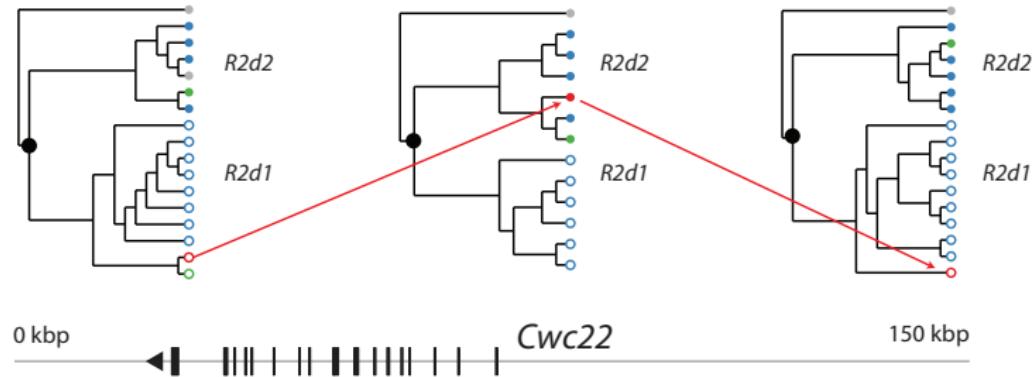
R2d was duplicated once  $\sim$  2 Mya, but one of the resulting copies has been lost multiple times.

## R2d2 duplication is old

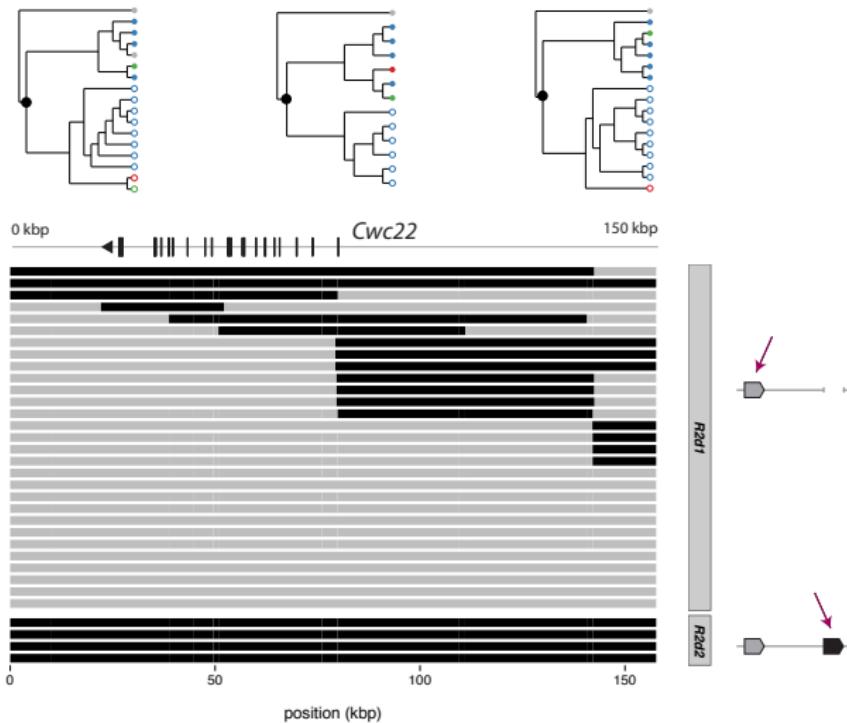


R2d was duplicated once  $\sim$  2 Mya, but one of the resulting copies has been lost multiple times.

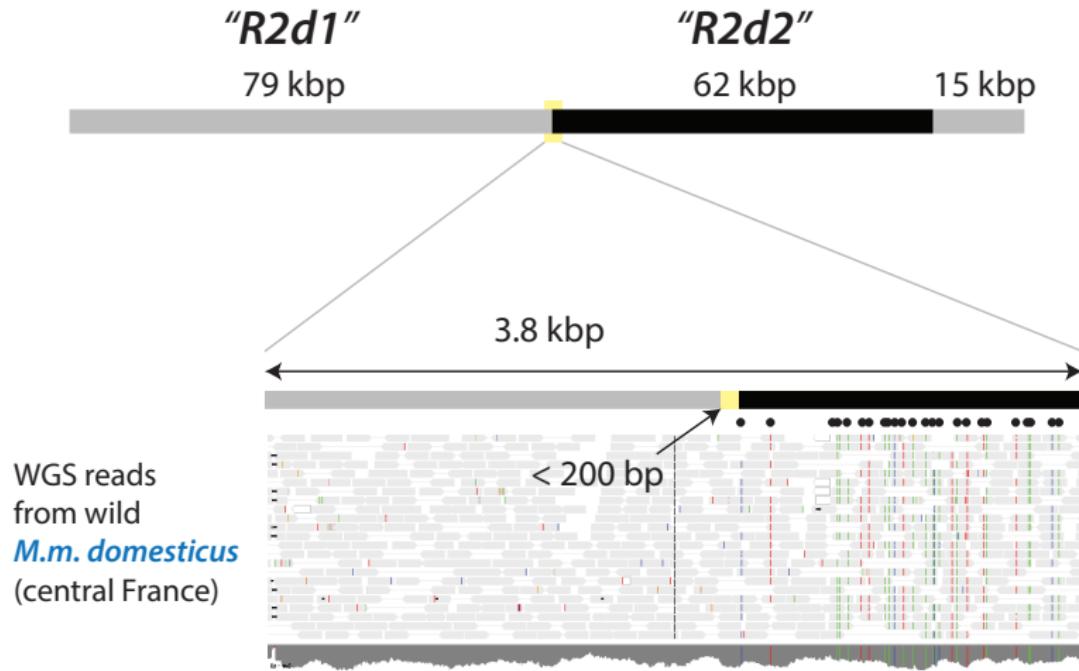
# *R2d2* is a focus for non-allelic gene conversion



# *R2d2* is a focus for non-allelic gene conversion

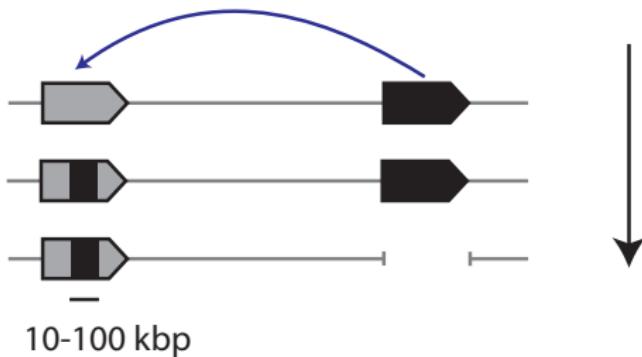


# Defining the boundary of a conversion event



## Properties of conversion events

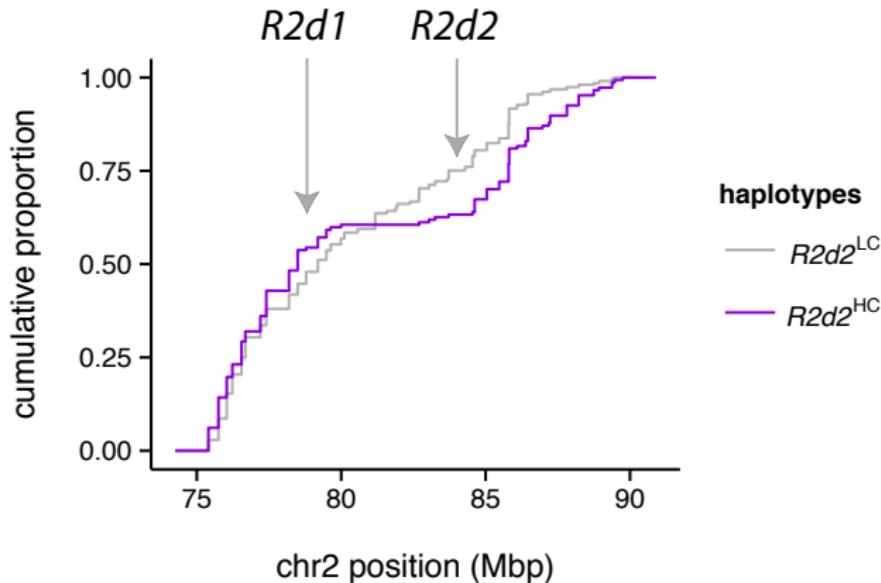
- Apparently frequent: 9 of 36 chromosomes examined from the wild (though difficult to estimate precise rates)
- Long conversion tracts (10 – 100 kbp) comprising thousands of paralogous variants
- Create mosaic *Cwc22* isoforms
- Breakpoints are clustered, but not identical



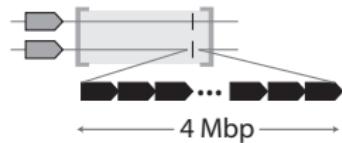
# Crossovers are locally suppressed around *R2d2*

In 4640 Diversity Outbred mice:

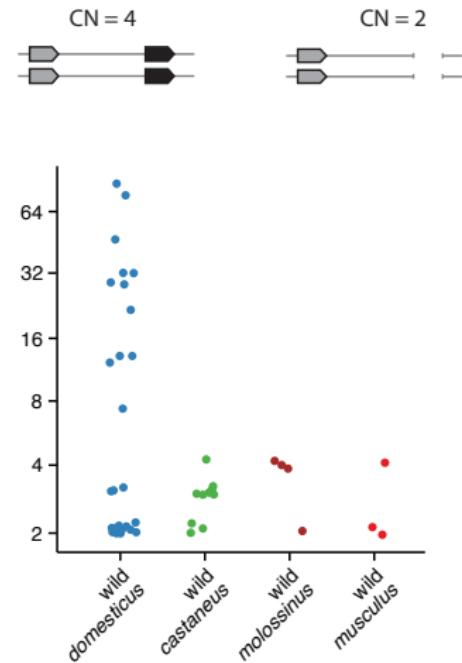
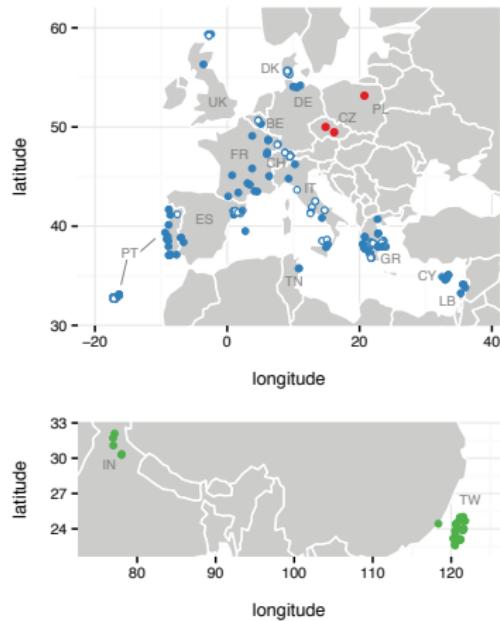
**7-fold** reduction in crossovers,  $R2d2^{HC}$  vs  $R2d2^{LC}$  ( $p < 10^{-8}$ )



# Crossovers are locally suppressed around *R2d2*

|                | configuration   | rate |
|----------------|---|------|
| $R2d2^{LC/LC}$ |  | +    |
| $R2d2^{HC/LC}$ |  | -    |
| $R2d2^{HC/HC}$ |  | ?    |

# *R2d2* is a CNV hotspot



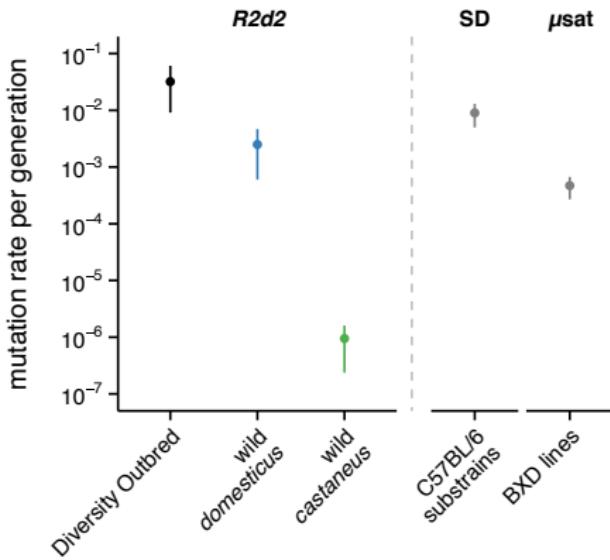
$n = 542$  wild or wild-derived mice

# *R2d2* is a CNV hotspot

- 9 new alleles in 183 chromosomes from the Diversity Outbred
- 0 new alleles in > 50 generations of an *R2d2*<sup>HC/HC</sup> inbred strain
- 0 new alleles in > 2000 meiosis through *R2d2*<sup>HC/LC</sup> females

Mutations increase with:

- higher copy number
- heterozygosity
- male sex?



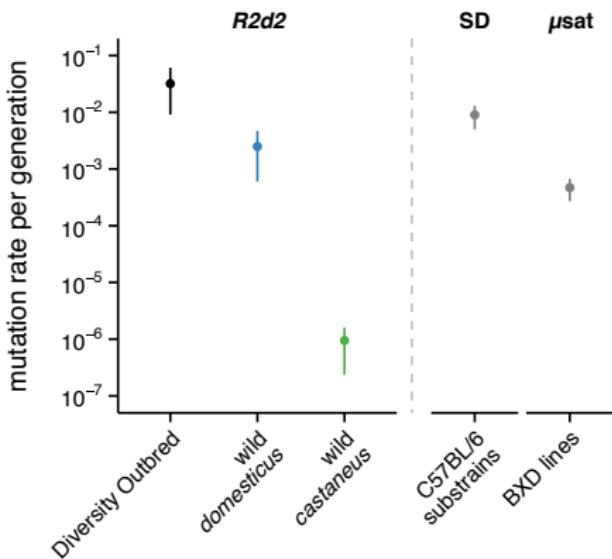
Dallas JF (1992) *Mamm Genome* // Egan C (2007) *Nat Genet*

# *R2d2* is a CNV hotspot

- 9 new alleles in 183 chromosomes from the Diversity Outbred
- 0 new alleles in > 50 generations of an *R2d2*<sup>HC/HC</sup> inbred strain
- 0 new alleles in > 2000 meiosis through *R2d2*<sup>HC/LC</sup> females

Mutations increase with:

- higher copy number
- heterozygosity
- male sex?

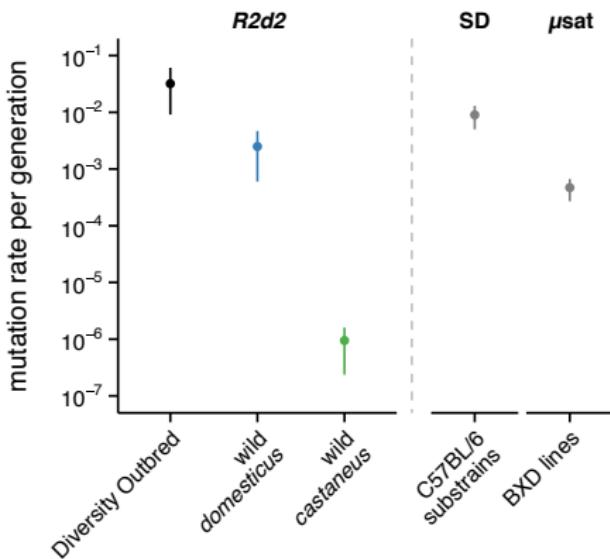


# *R2d2* is a CNV hotspot

- 9 new alleles in 183 chromosomes from the Diversity Outbred
- 0 new alleles in > 50 generations of an *R2d2*<sup>HC/HC</sup> inbred strain
- 0 new alleles in > 2000 meiosis through *R2d2*<sup>HC/LC</sup> females

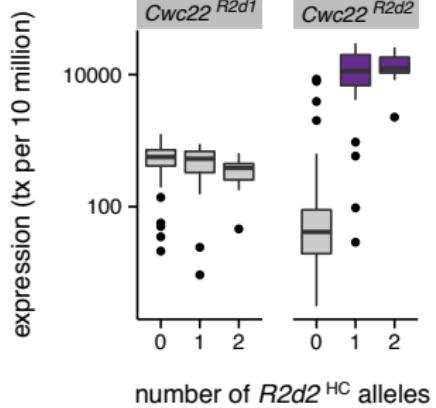
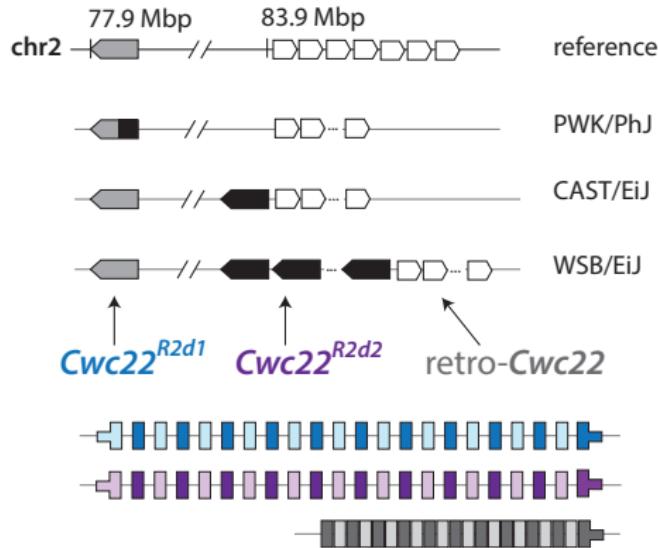
## Mutations increase with:

- higher copy number
- heterozygosity
- male sex?



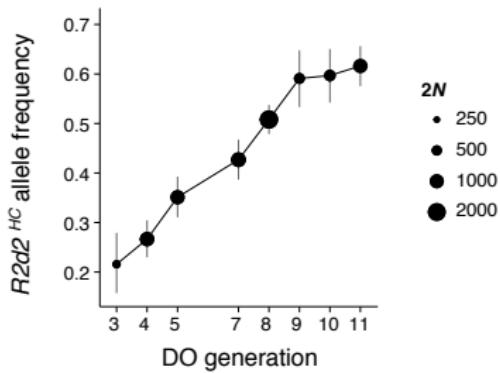
Dallas JF (1992) *Mamm Genome* // Egan C (2007) *Nat Genet*

# Impact of *R2d2* CNVs on gene expression



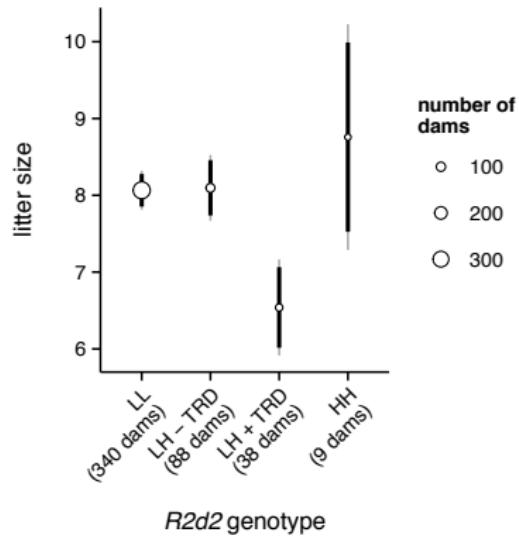
# Impact of *R2d2* CNVs on meiosis & reproduction

## “Selfish sweeps”



John Didion (poster #24)

## Embryo survival



# Summary

- *R2d2* is a meiotic drive locus derived from an ancestral duplication
- Paradoxical patterns of mutation:
  - ▶ many gene conversions . . . but no crossovers
  - ▶ many CNVs . . . but few SNVs
- Copy number impacts phenotypes
- Dependence on heterozygosity

## Future work

- mechanism of meiotic drive
- modifiers of mutation rate
- epigenetic marks associated with mutation

# Summary

- *R2d2* is a meiotic drive locus derived from an ancestral duplication
- Paradoxical patterns of mutation:
  - ▶ many gene conversions . . . but no crossovers
  - ▶ many CNVs . . . but few SNVs
- Copy number impacts phenotypes
- Dependence on heterozygosity

## Future work

- mechanism of meiotic drive
- modifiers of mutation rate
- epigenetic marks associated with mutation

# Acknowledgments

## Our group

Rachel McMullan

Liran Yadgary

Amelia Clayshulte

Tim Bell

Sarah Cates

Nicole Robinson

Ginger Shaw

Darla Miller

**Fernando Pardo-Manuel de Villena**

## Collaborators

Leonard McMillan

Daniel Pomp

Gary Churchill (Jax)

## Labmates past

**John Didion**

John Calaway

Justin Gooch

Ryan Buus

## Funding

P50GM076468

P50HG006582

R21MH09621

T32GM067553

F30MH103925

## Data

Paul Flicek (EMBL-EBI)

Duncan Odom

(EMBL-EBI)

Sanger Ctr/Wellcome

Trust

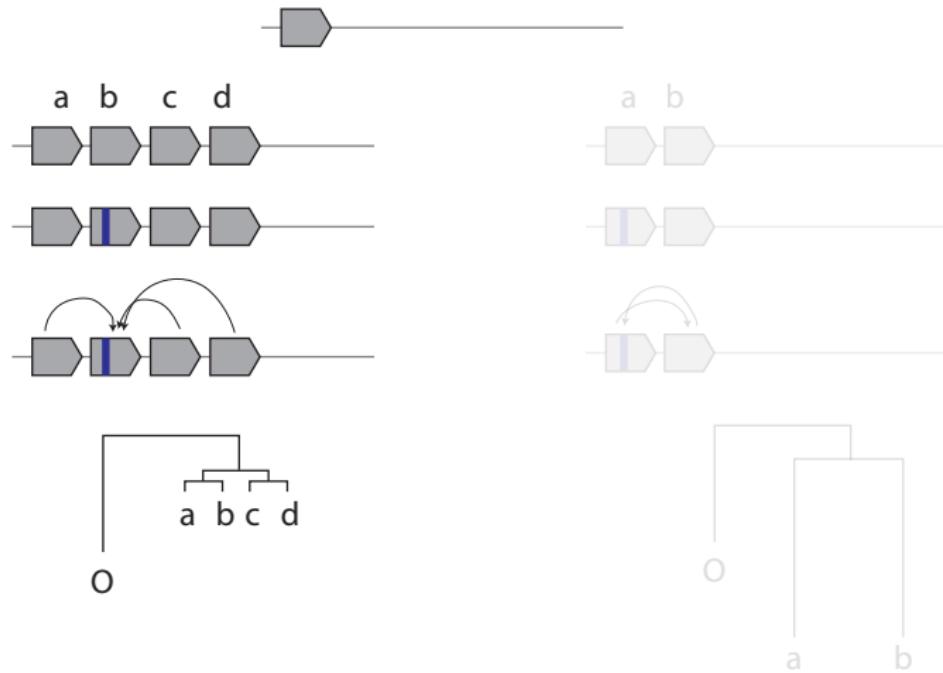
The Jackson Lab

## Wild mice

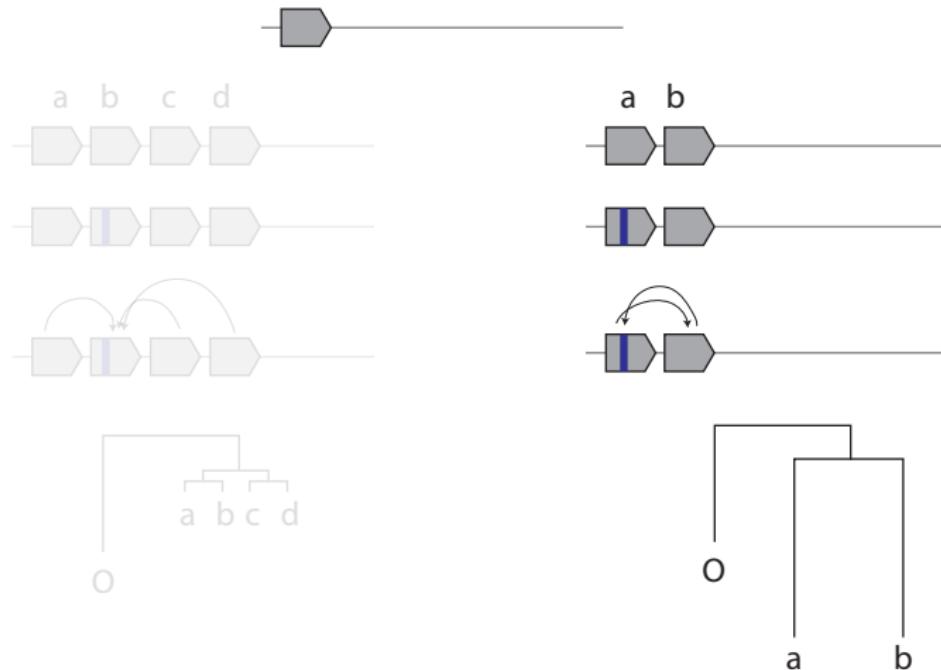
Jeremy Searle, Francois Bonhomme, Pierre Boursot, Janice Britton-Davidian, Ricardo Castiglia, Eva Giagia-Athan-asopoulou, Sofia Gabriel, Silvia Garagna, Sofia Grize, Isla Gündüz, Bettina Harr, Heidi Hauffe, Jeremy Herman, Leon Kontrimavicius, Anna Lindholm, Maria de Luz Mathias, George Mistainas, Jaroslav Pialek, Priscilla Tucker, Jacint Ventura, Jan Wojcik, Stephan Rosshart (NIDDK), Barbara Rehermann (NIDDK), Amanda Chunco (Elon) ...

# Supplementary material

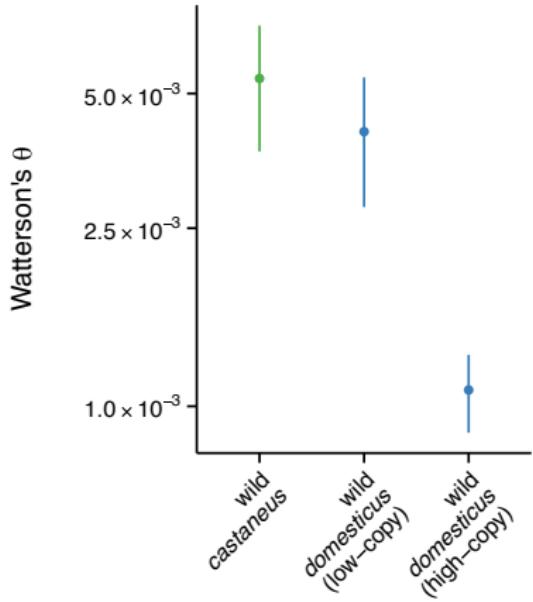
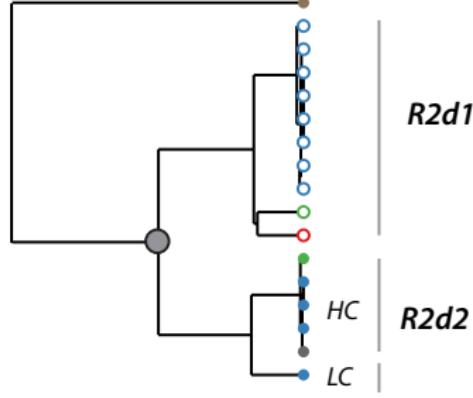
# Copy number influences rate of genetic drift



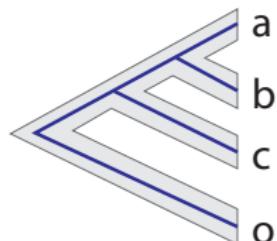
# Copy number influences rate of genetic drift



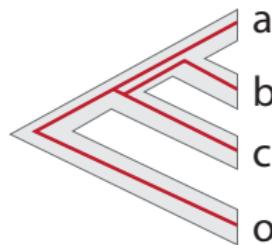
# Sequence diversity in $R2d2^{HC}$ alleles is relatively low



# Phylogenetic discordance in the rest of the genome



*concordant*



*discordant*

0.8% of mouse genome has similar properties:

- discordant phylogeny ( $T_{MRCA} > 2$  Mya)
- low recombination rate
- CNV/SD-rich

These loci are enriched for:

- olfactory receptors
- immunoglobulins