Rampant gene duplication facilitates Y-linked gene evolution in the *Drosophila simulans* clade

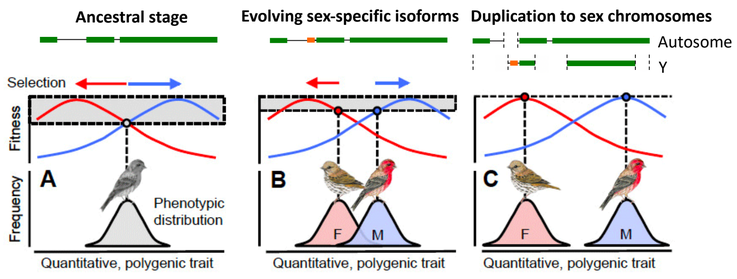
Ching-Ho Chang[1,3], Lauren Gregory[1], Colin Meiklejohn[2] and Amanda Larracuente[1]
[1]Department of Biology, U of Rochester, [2] School of Biological Sciences, U of Nebraska-Lincoln
[3] Current address: Fred Hutchinson Cancer Research Center
ching-ho.chang@rochester.edu

Summary

I. Long sequencing and *de novo* assembly of the Y chromosomes in *D. melanogaster* and *simulans* clade species
II. Rapid evolution of structure and gene content on Y chromosomes
III. Intralocus genetic conflicts and gene duplication

Conflicts and Y chromosomes

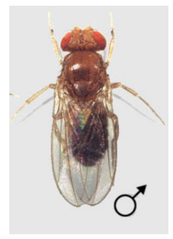
Resolving intralocus sexual antagonism

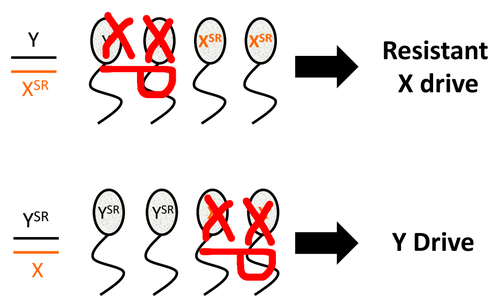


Accumulating male-specific genes

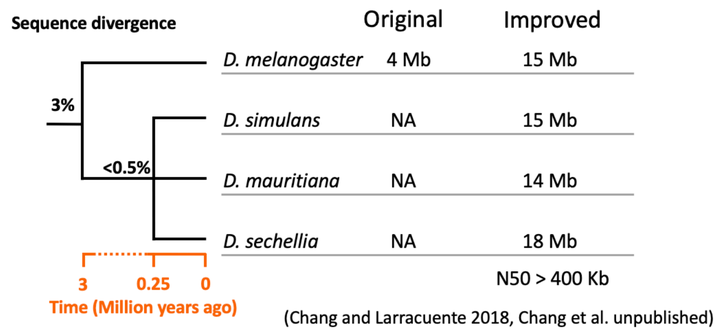
(Figure adapted from https://www.coxlabuva.org)

Sex-linked meiotic drive (sperm killer)





Improving Y assemblies

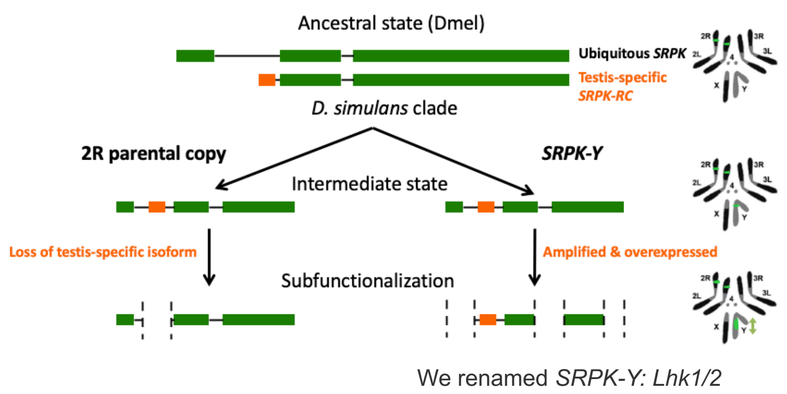


(Chang and Larracuente 2018, Chang et al. unpublished)

Y Chromosome Assembly	# of Contigs	Total size	Contig N50
D. melanogaster	80	14,578,684	416,887
D. simulans	38	13,717,056	1,031,383
D. sechellia	63	14,899,148	555,130
D. mauritiana	55	17,880,069	1,628,994

Y-linked ampliconic genes specific to the *simulans* clade

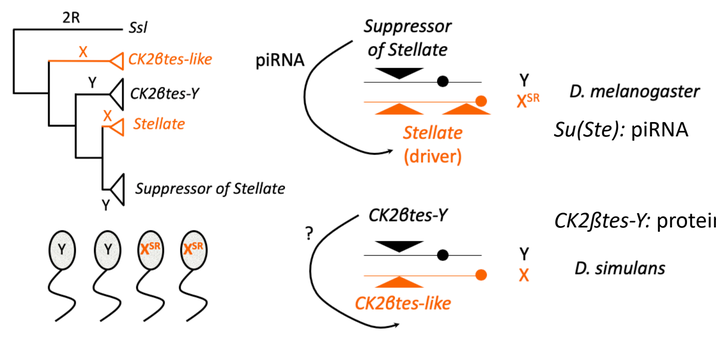
Translocation of a testis-specific isoform



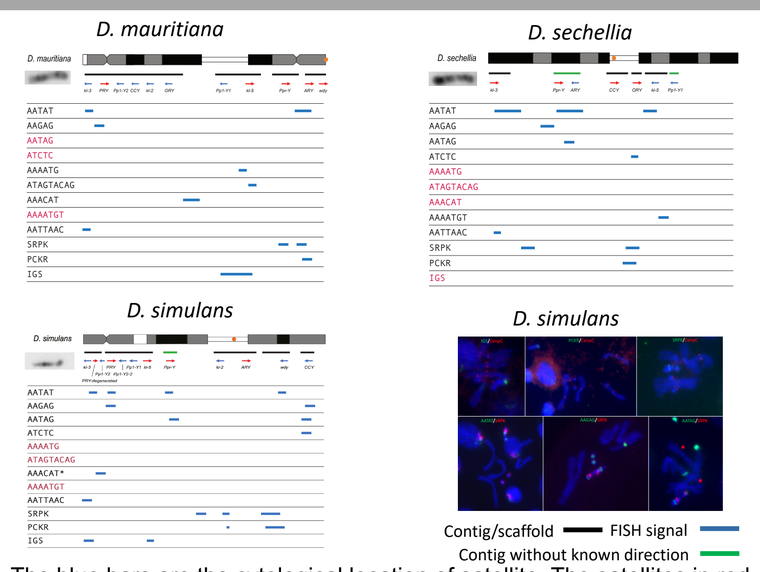
We renamed *SRPK-Y: Lhk1/2*

Co-amplification of a potential driver

Stellate (casein kinase) is an X-linked driver in *D. melanogaster*
Su(Ste) is a Y-linked suppressor of drive
Stellate and *Su(Ste)* genes are amplified. (Palumbo et al. 1994)

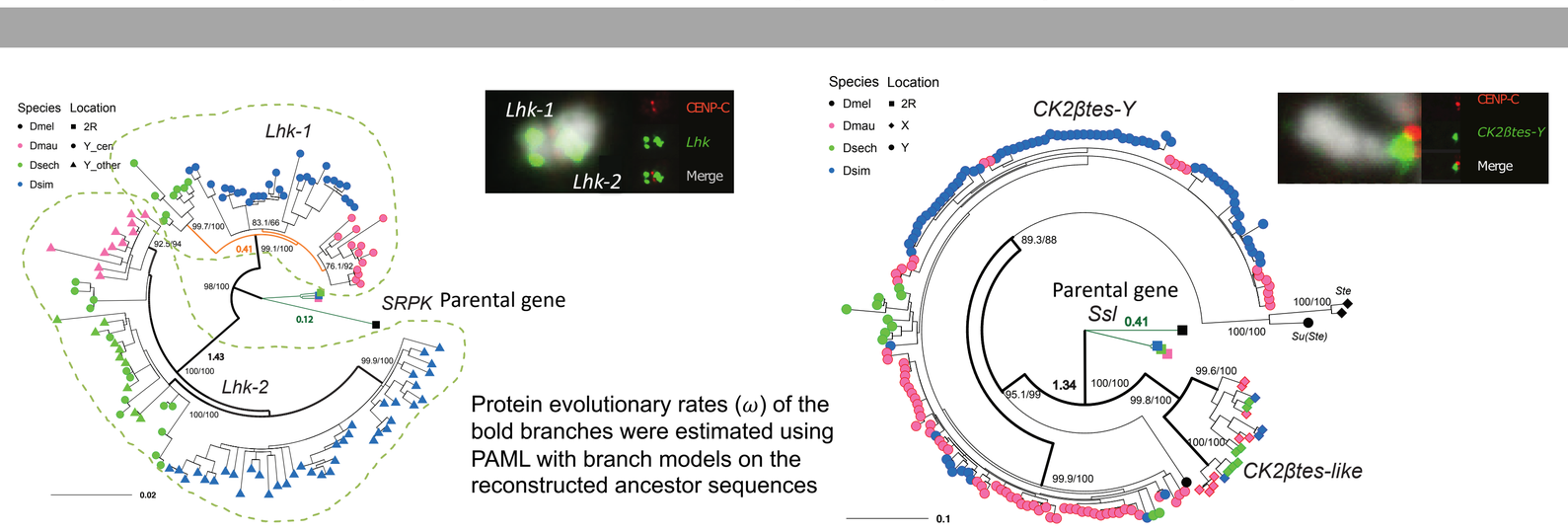


Placing contigs on the Y using FISH



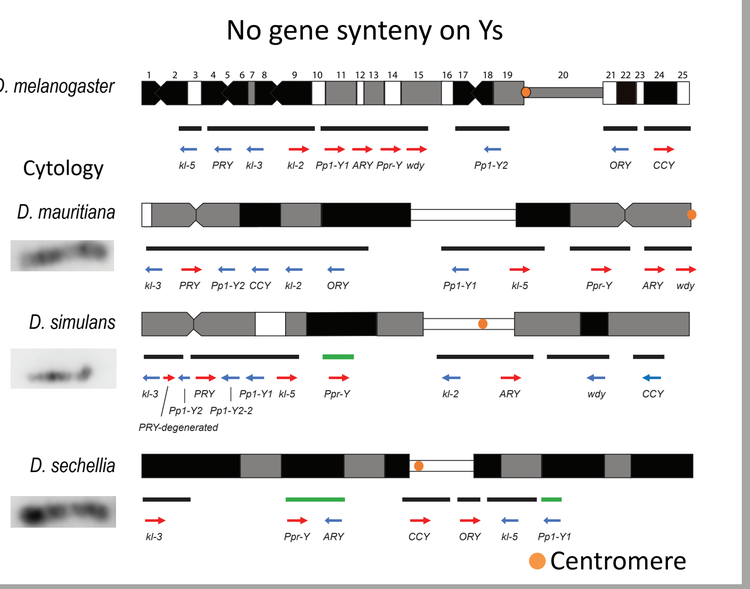
The blue bars are the cytological location of satellite. The satellites in red are sequences we cannot detect on Y chromosomes using FISH.

Strong positive selection on Y ampliconic genes ($\omega = 1.3–1.5$)



***Lo-han-kha (Lhk)* –** sleep in temples
The homeless male migrants that moved from the mainland China to Taiwan.

Extensive rearrangement of Y-linked genes



SRPK has different roles in each sex

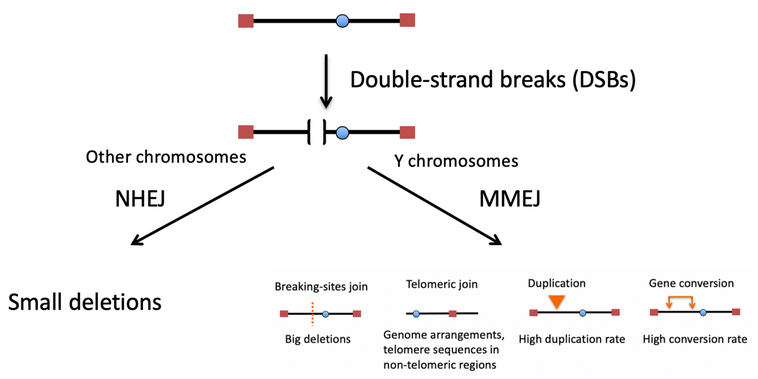
I. Testis-specific isoform, *SRPK-RC,* is recently acquired in melanogaster group
II. *SRPK* knockout is sterile in both sexes
III. Two *SRPK* mutants are only sterile in females
IV. Ubiquitous *SRPK* isoforms are expressed premeiotically and are localized to cytoplasm and nucleolus in testes
V. *SRPK-RC* is only expressed in primary spermatocytes

Common features of ampliconic Y genes

I. Positive selection after duplicating to the Y
II. Massive copies (>40) on Ys
III. High gene conversion between Y copies (10^{-4} to 10^{-6} event per bp/generation)
IV. Nuclear localization and function in serine phosphorylation

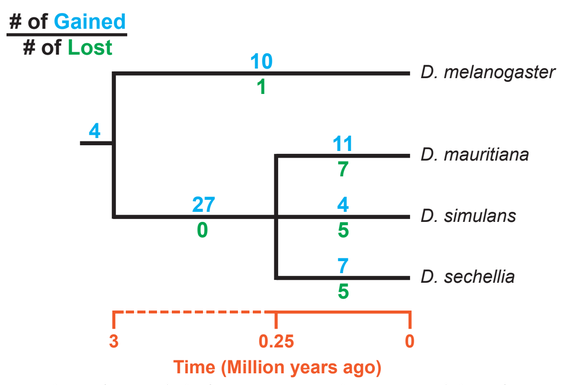
Y-specific mutation patterns?

Model: Microhomology-mediated end joining
I. NHEJ is rare in heterochromatin regions
II. No homologous recombination



Rapid turnover of duplicated genes on Ys

Many genes duplicated to Y from other chromosomes
>50% of new Y-linked genes are degenerated



*whole mitochondrial genome inserted in the ancestral *simulans* clade Y

Ongoing

I. Exploring functional divergence between *SRPK-RC, Lhk-1, and Lhk-2*
II. What drives positive selection and amplification of multicopy Y-linked genes?
III. Do the sequence and copy number divergences between species contribute to hybrid sterility in *D. simulans* clade?

References

Chakraborty et al. 2016. Nucleic Acids Res 44(19):e147.
Chin et al. 2013. Nat Methods 10:563–569.
Hoskins et al. 2015. Genome Res 25(3):445–458.
Koren et al. 2017. Genome Res 27(5):722–736.
Loh et al. 2012. J Cell Sci 125(Pt 19):4457–62.
Palumbo et al. 1994. Genetics 138:1181–1197.

Acknowledgements

This work was supported by NIH R35 GM11951501 and University of Rochester funds to AML. We thank our collaborators in the Emerson lab (UC Irvine), for generating PacBio reads in *simulans* clade, Dr. Barbara Mellone for the CENP-C antibody, and Dr. Casey Bergman and Dr. Kevin Wei for helpful discussion. We also thank the U of Rochester CIRC for computing cluster resources.