



Sequence and structural diversity of mouse Y chromosomes

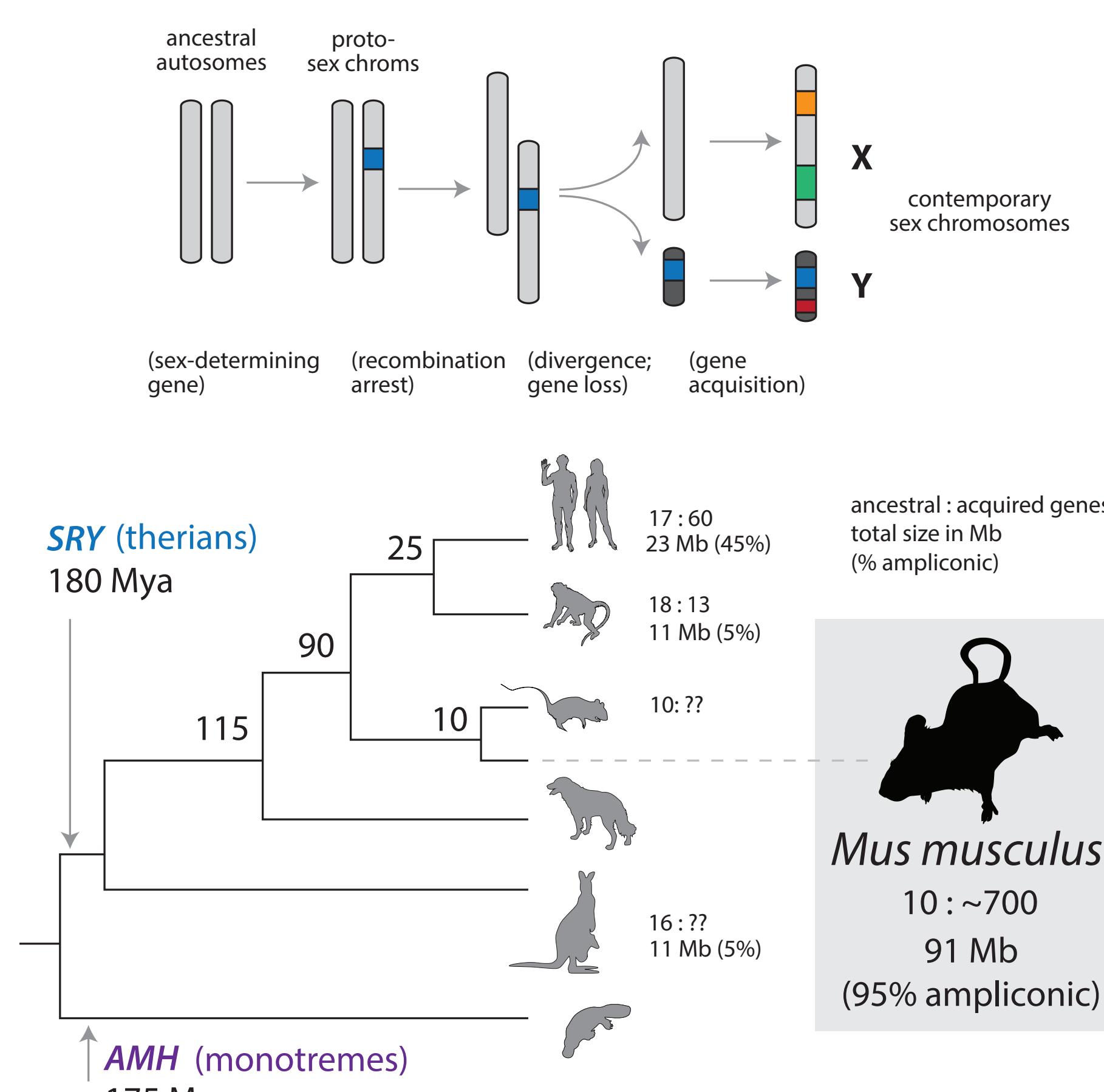
Andrew P Morgan^{1,2} and Fernando Pardo-Manuel de Villena¹

¹Department of Genetics and ²Curriculum in Bioinformatics,
University of North Carolina at Chapel Hill

Why study the Y chromosome?

- Required for sex determination and fertility.
- Fast-evolving, lineage-specific biology.
- Implicated in speciation in many taxa.
- Rich source of phylogenetic and demographic information.

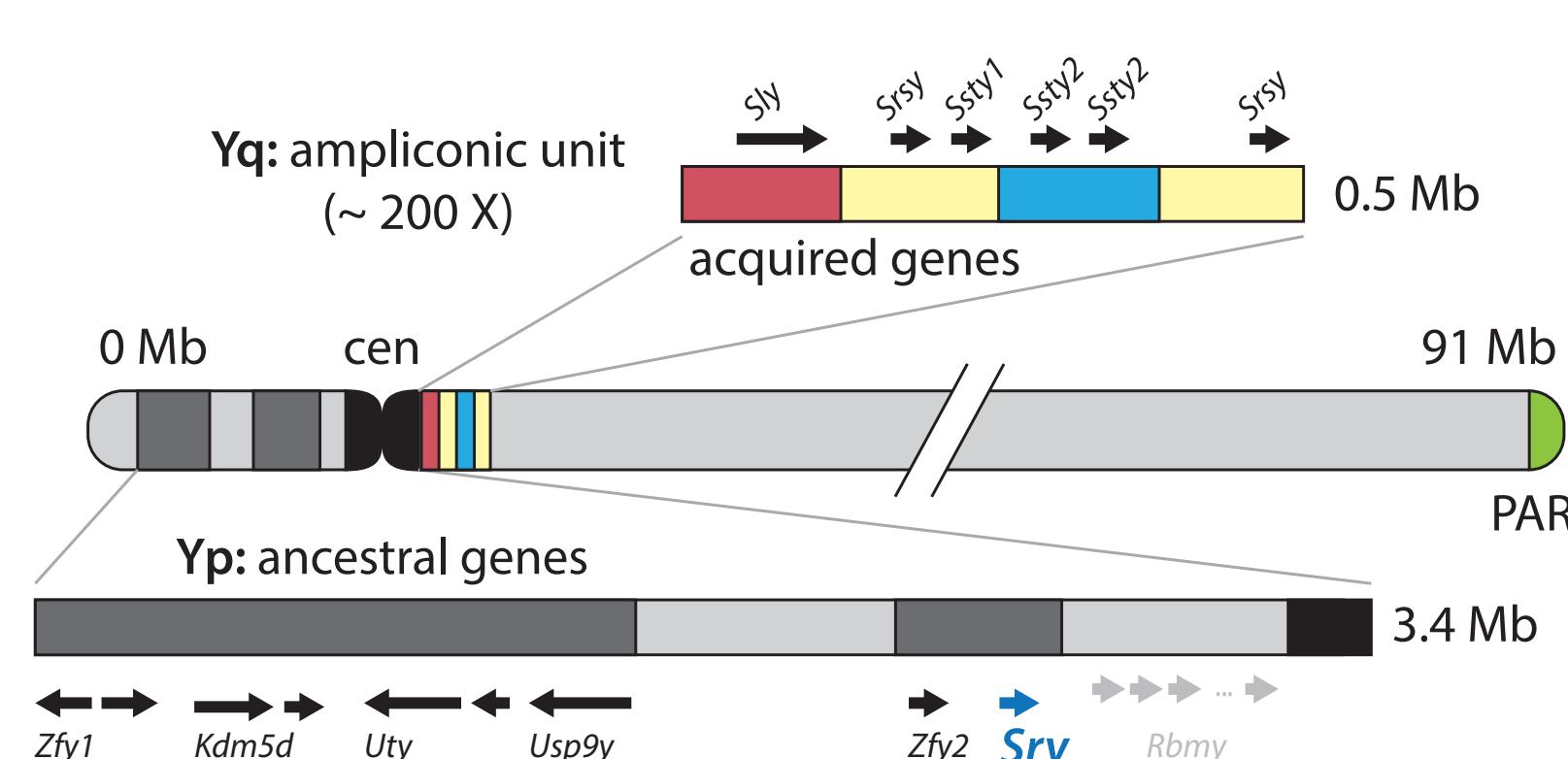
Mammalian sex chromosomes



Sex chromosomes have evolved independently in many taxa but share similar functional content and follow a predictable trajectory. [1]

Y chromosomes were long thought to have minimal importance outside the germline.

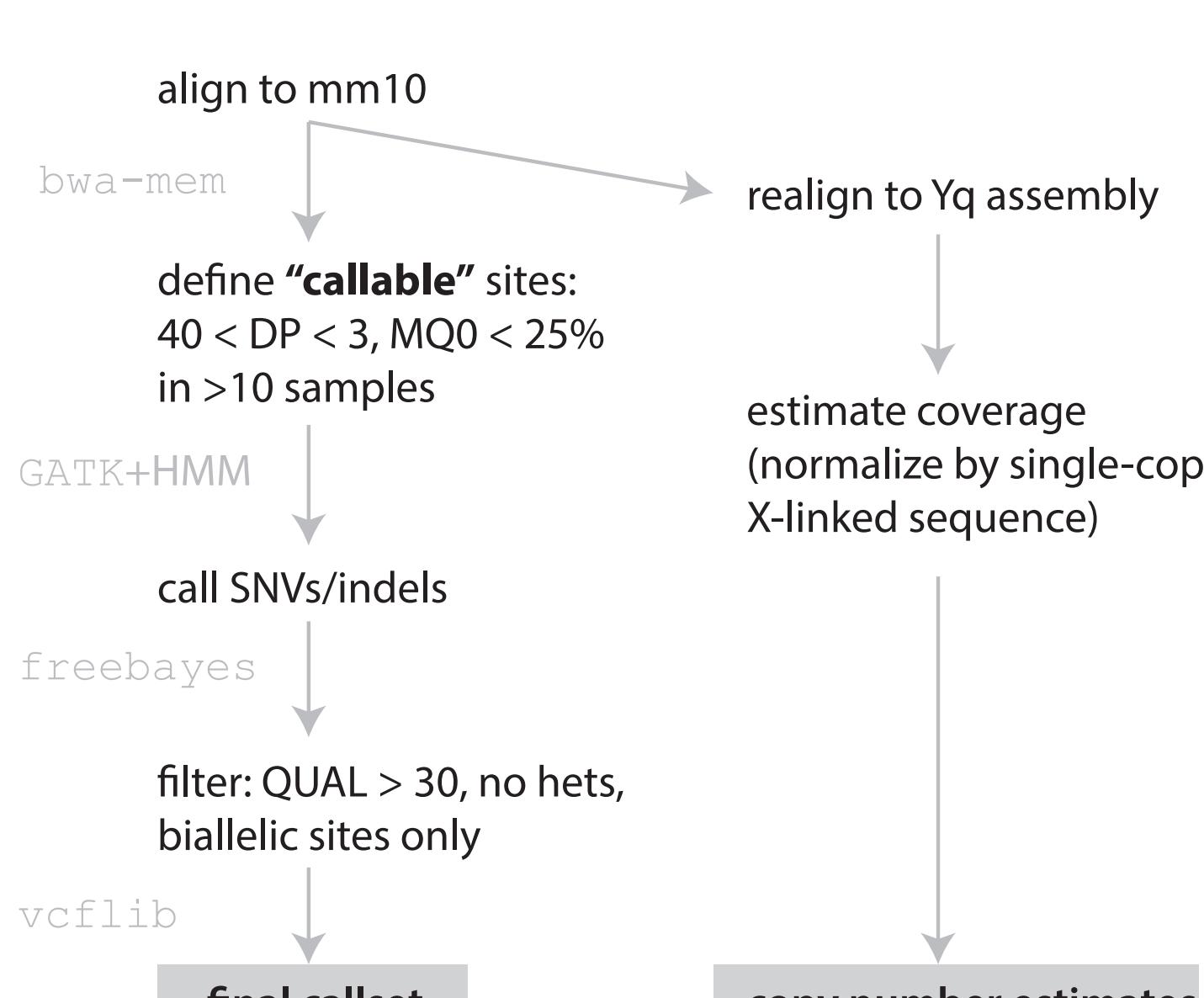
The mouse Y chromosome



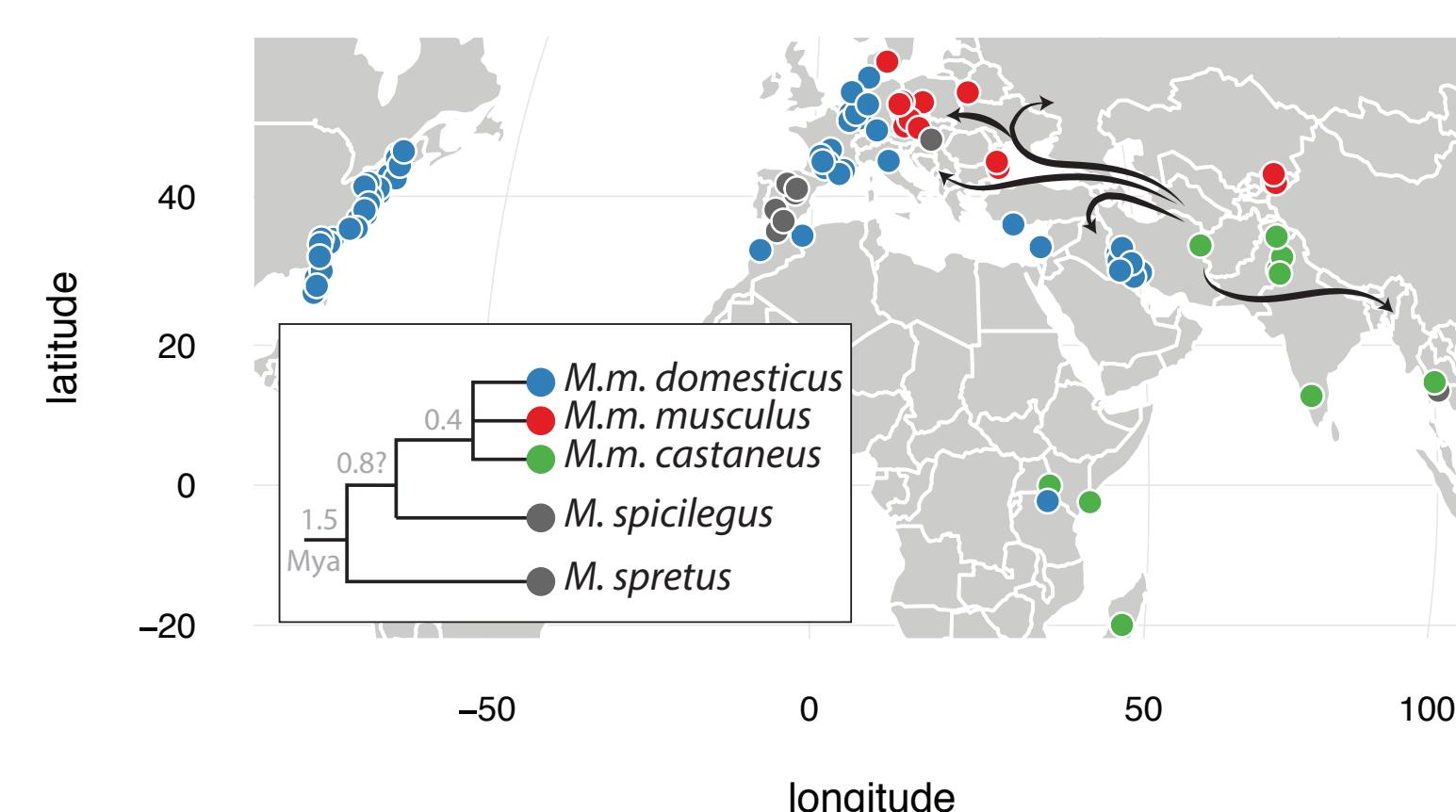
Unlike other mammalian Y chromosomes, the mouse Y is both **massively expanded** and **entirely euchromatic**. [2]

The long arm (Yq) has three rodent-specific gene families (*Sly*, *Srsy*, *Ssty*) whose expression is restricted to germ cells.

Sequencing workflow



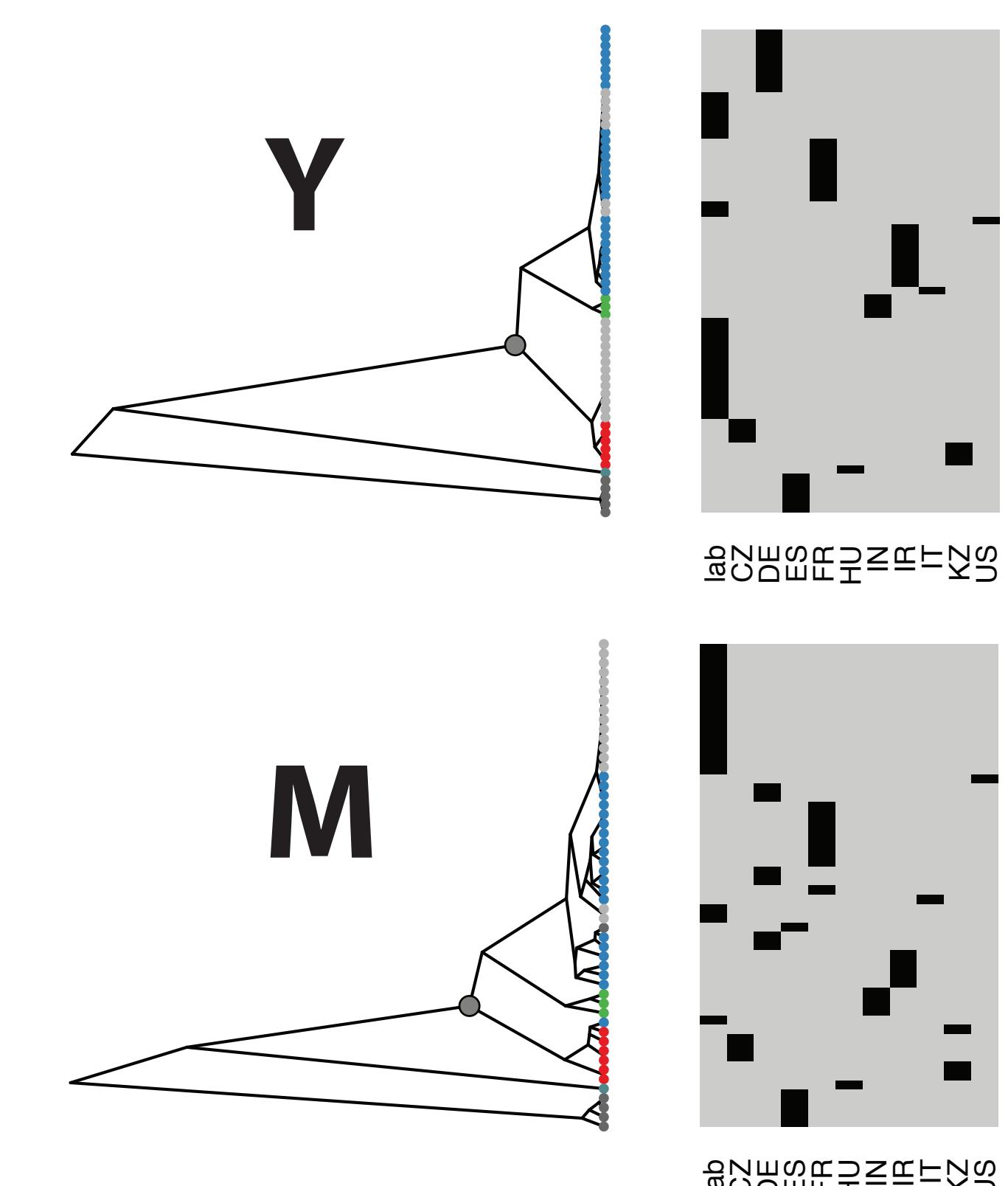
Global sample of Y chromosomes



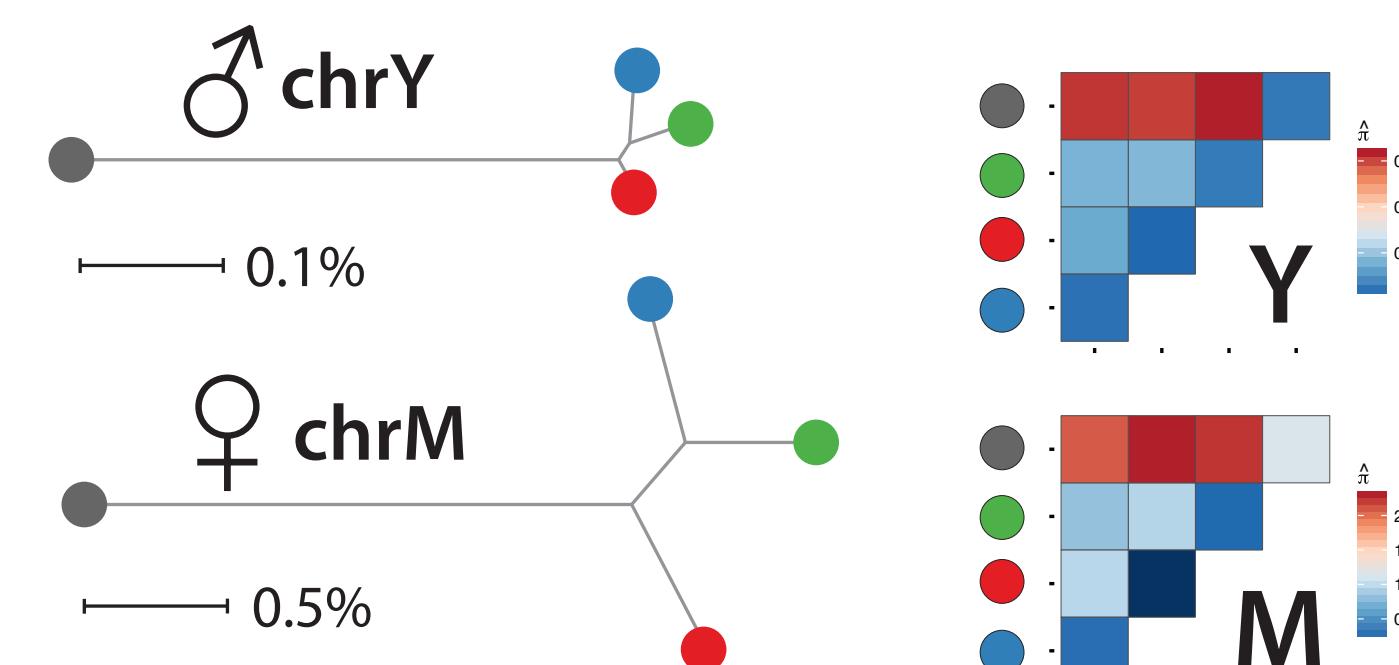
Our sample comprises 142 Y chromosomes from laboratory and wild house mice (*M. musculus*) and outgroup species *M. spicilegus* and *M. spretus*.

28 738 non-coding SNVs
283 coding SNVs
3 398 indels
32 419 total sites

Sex-specific phylogeography



Low patrilineal diversity

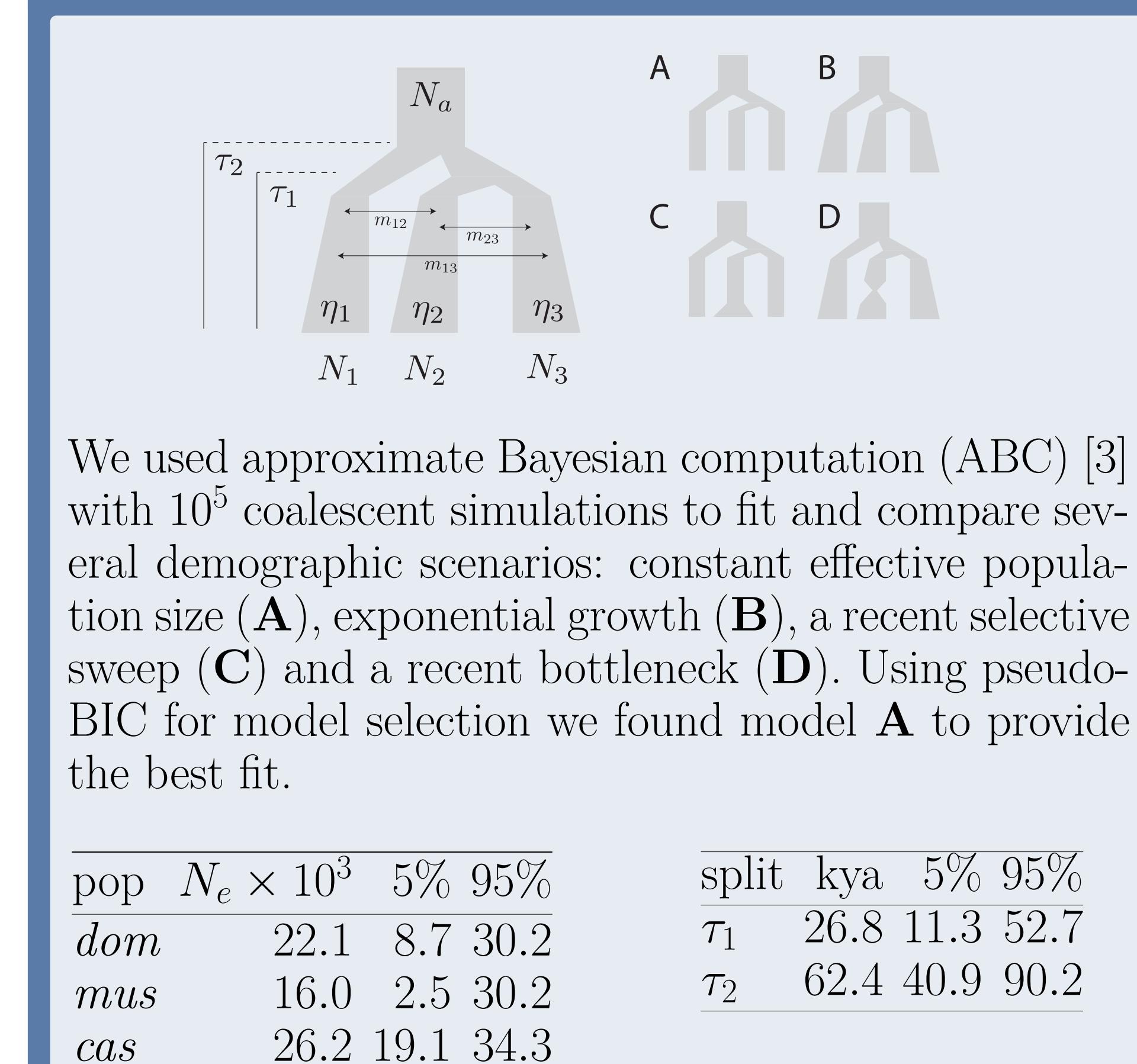


After standardizing by the divergence to *M. spretus* (grey), within-population diversity in *M. musculus* is **4-fold reduced** on the Y relative to the mitochondria.

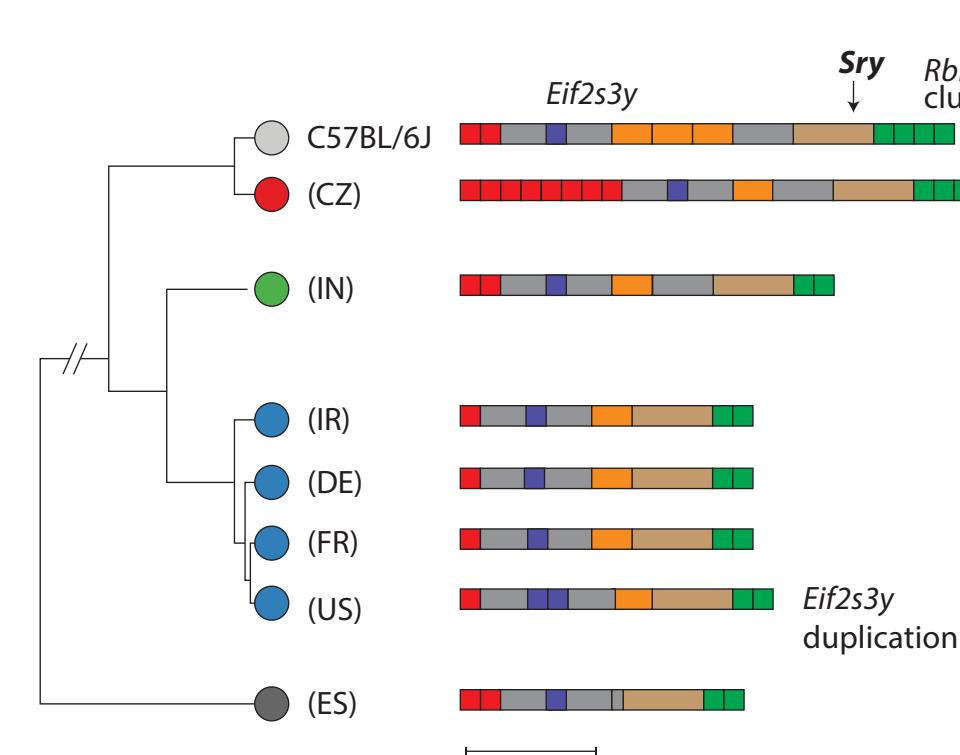
Diversity within *M. m. musculus* (red) is **3-fold lower** than within the other two *M. musculus* subspecies.

Hypothesis: X-Y conflict in *musculus-domesticus* hybrid zone imposes selective pressures that depress N_e in *musculus*.

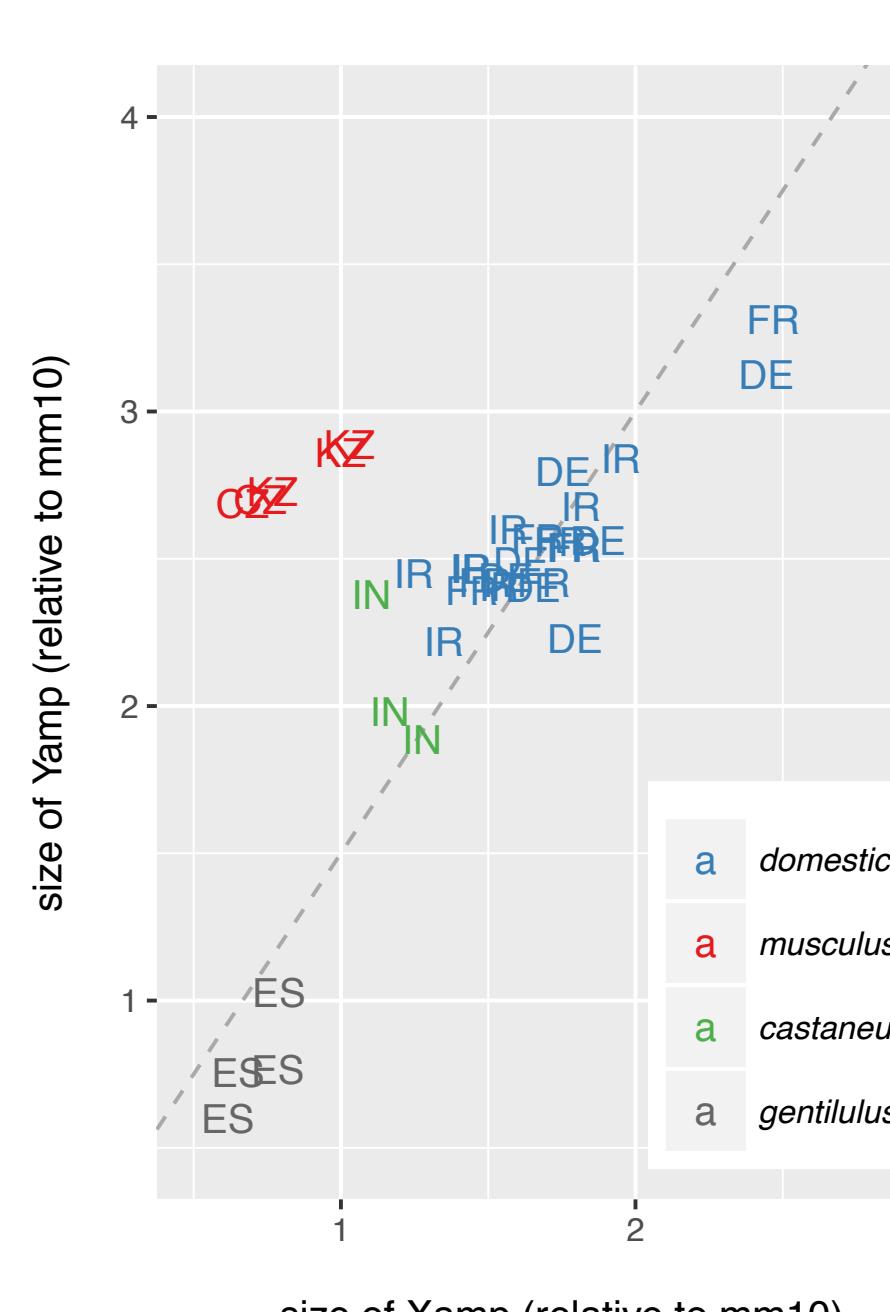
Demographic modeling



Large-scale structural differentiation



We identified ≥ 861 kb of copy-number variable sequence on short arm of Y (25.1% of its nominal length), of which 779 kb overlaps segmental duplications in the reference sequence. Alleles follow the SNV-based tree.

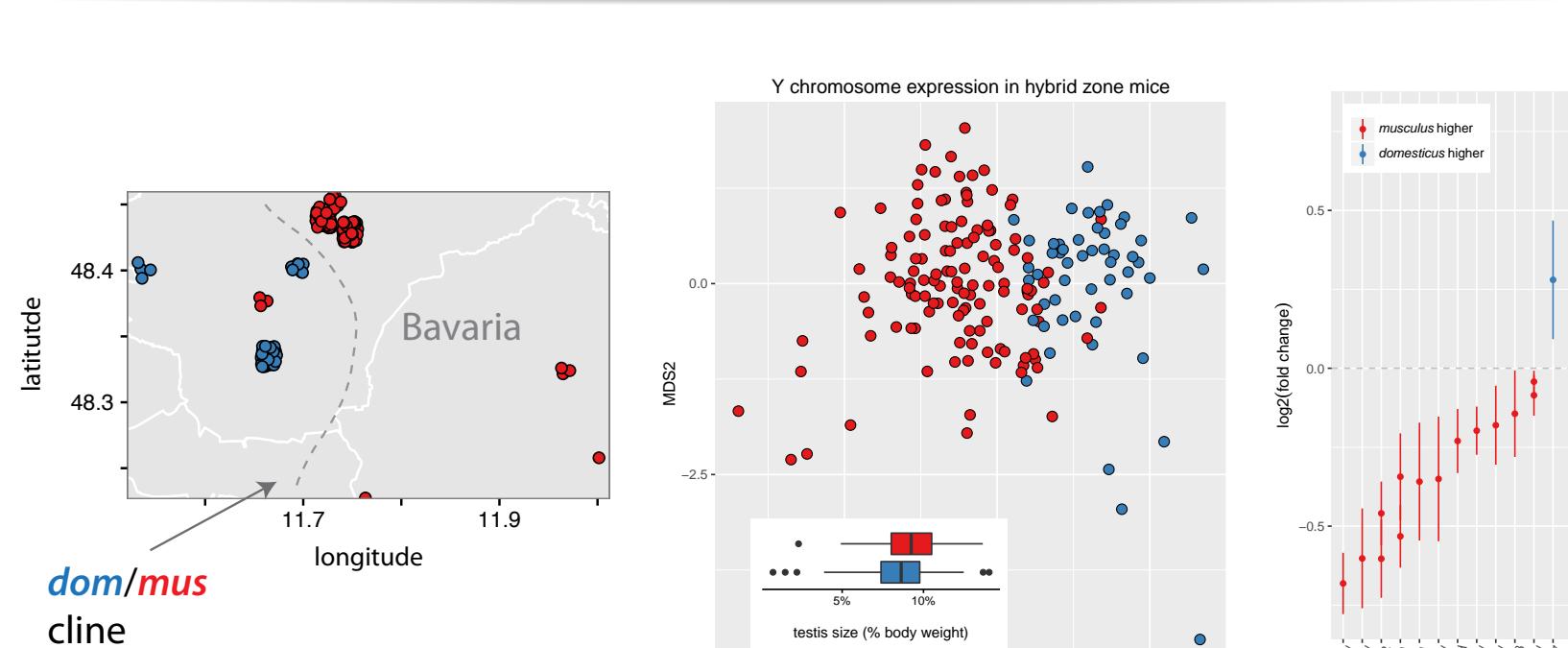


Gene families on the long arm (Yq) are co-amplified with their X-linked gametologs within the last 2 My in *Mus*.

X-Y dosage balance for co-amplified genes is different in *musculus* versus *domesticus* or *castaneus*.

Dosage balance between *Slx* and *Sly* is known to be required for spermiogenesis. [4]

Differentiated Y-linked expression



In mice from the *musculus-domesticus* hybrid zone [5], Y-linked expression pattern is determined by Y-chromosome origin independent of admixture on autosomes or X chromosome.

Acknowledgements

We thank B. Harr, L. Turner, D. Tautz, M. Phifer-Rixey and M. Nachman for their assistance in obtaining access to sequencing datasets.

Funds from P50GM076468 (FPMV), F30MH103925 (APM).

References

- Hughes (2015) *Annu Rev Genet* **49**: 507.
- Soh (2014) *Cell* **159**: 800.
- Pritchard (1999) *Mol Biol Evol* **16**: 179.
- Cocquet (2010) *Mol Biol Cell* **21**: 3497.
- Turner LM (2014) *eLife* **3**: e02504.

