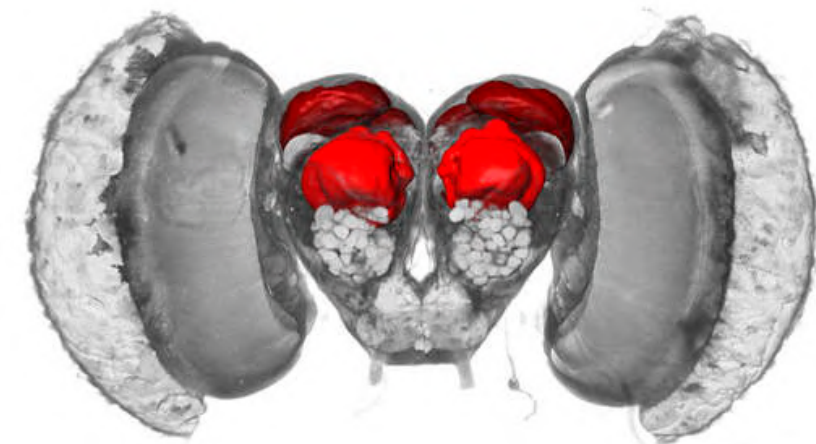


Day 9 - Comparative genomics

Whole genome alignment & Conserved Elements

F. Cicconardi, PhD

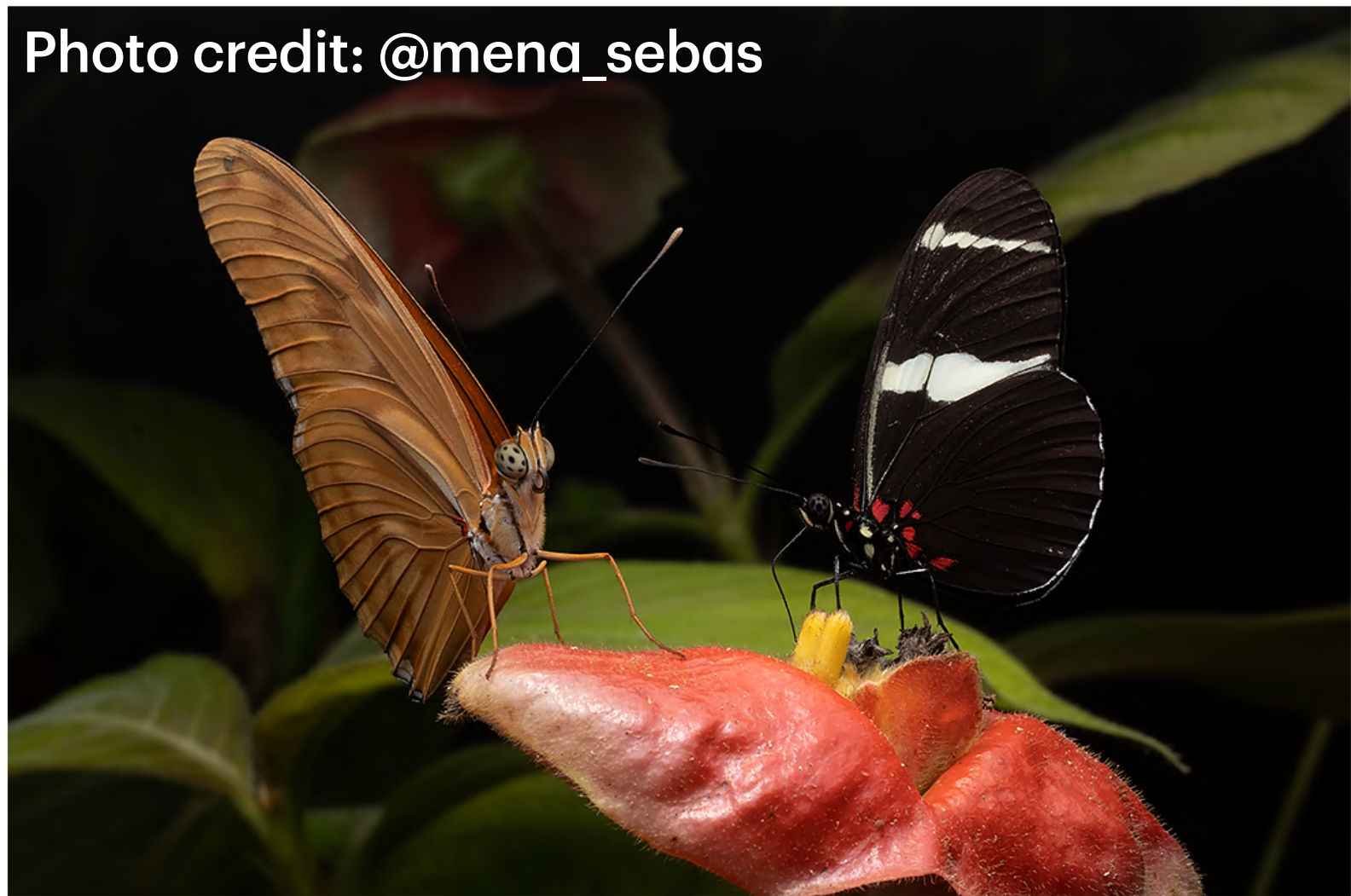


EBaB lab

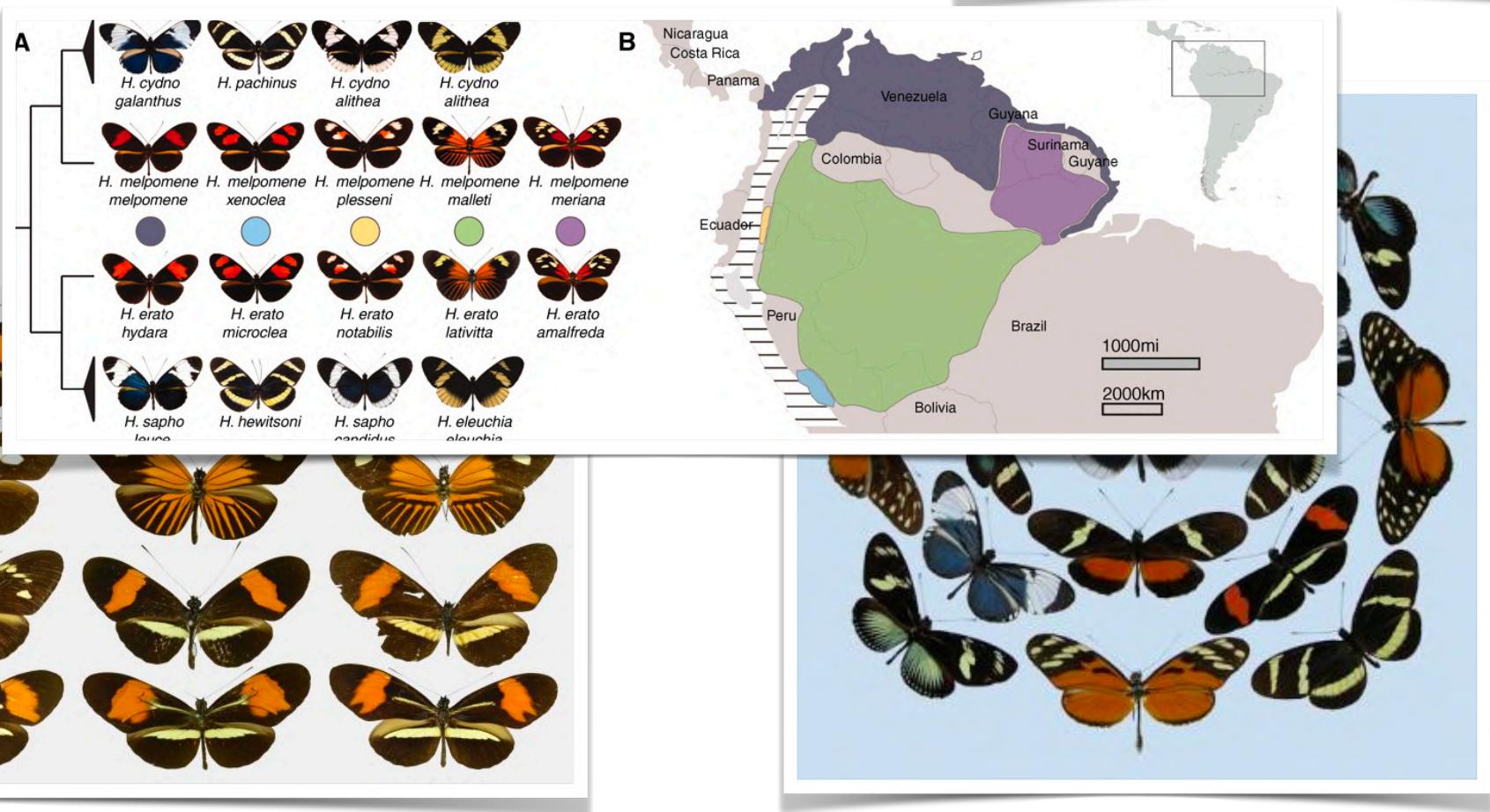
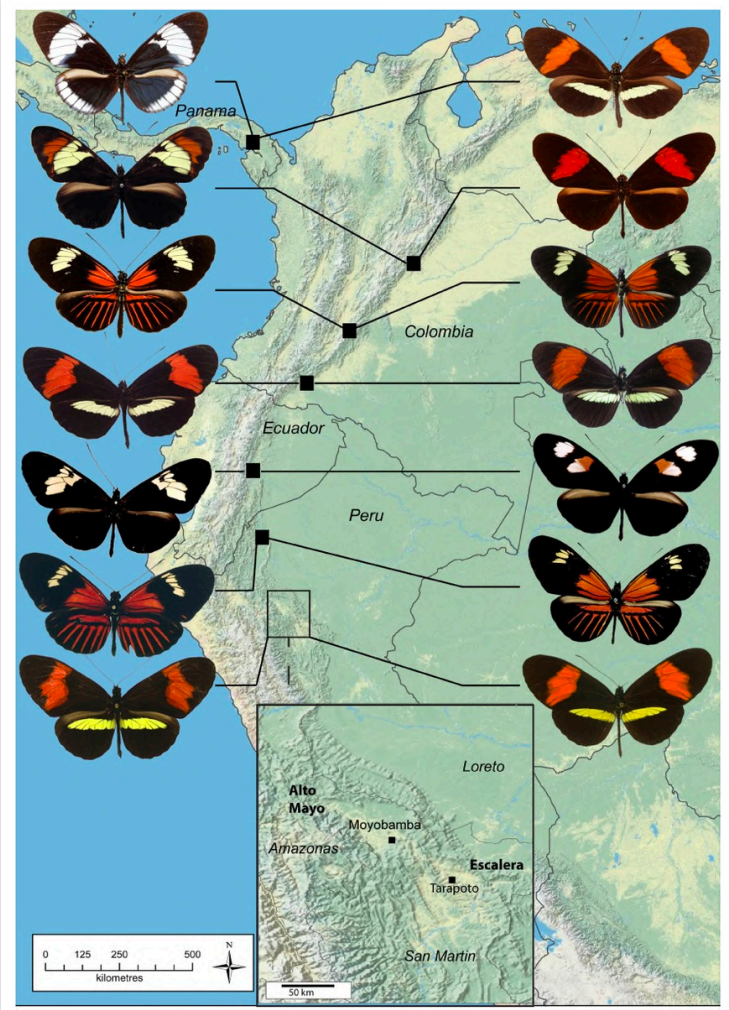
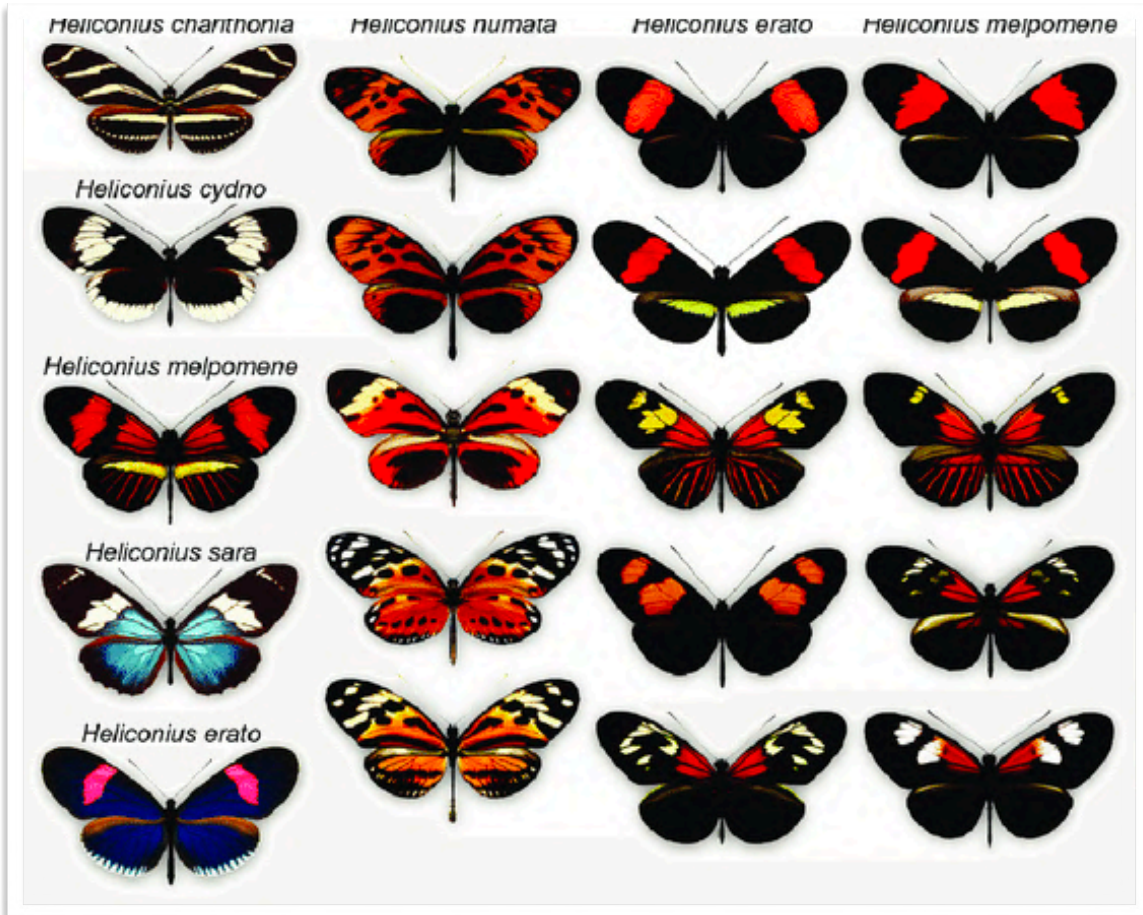
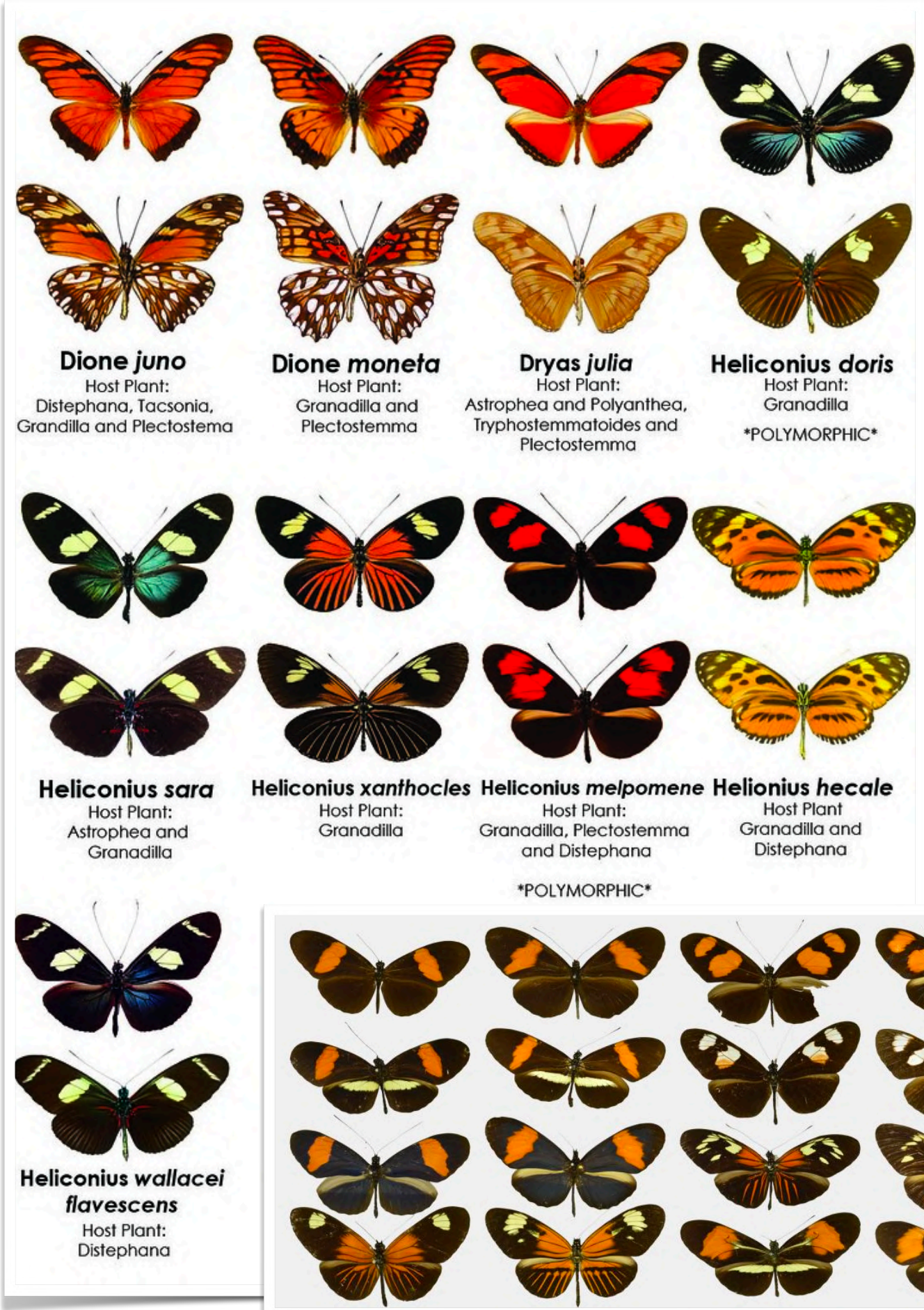
2023 WORKSHOP ON GENOMICS, CESKY KRUMLOV

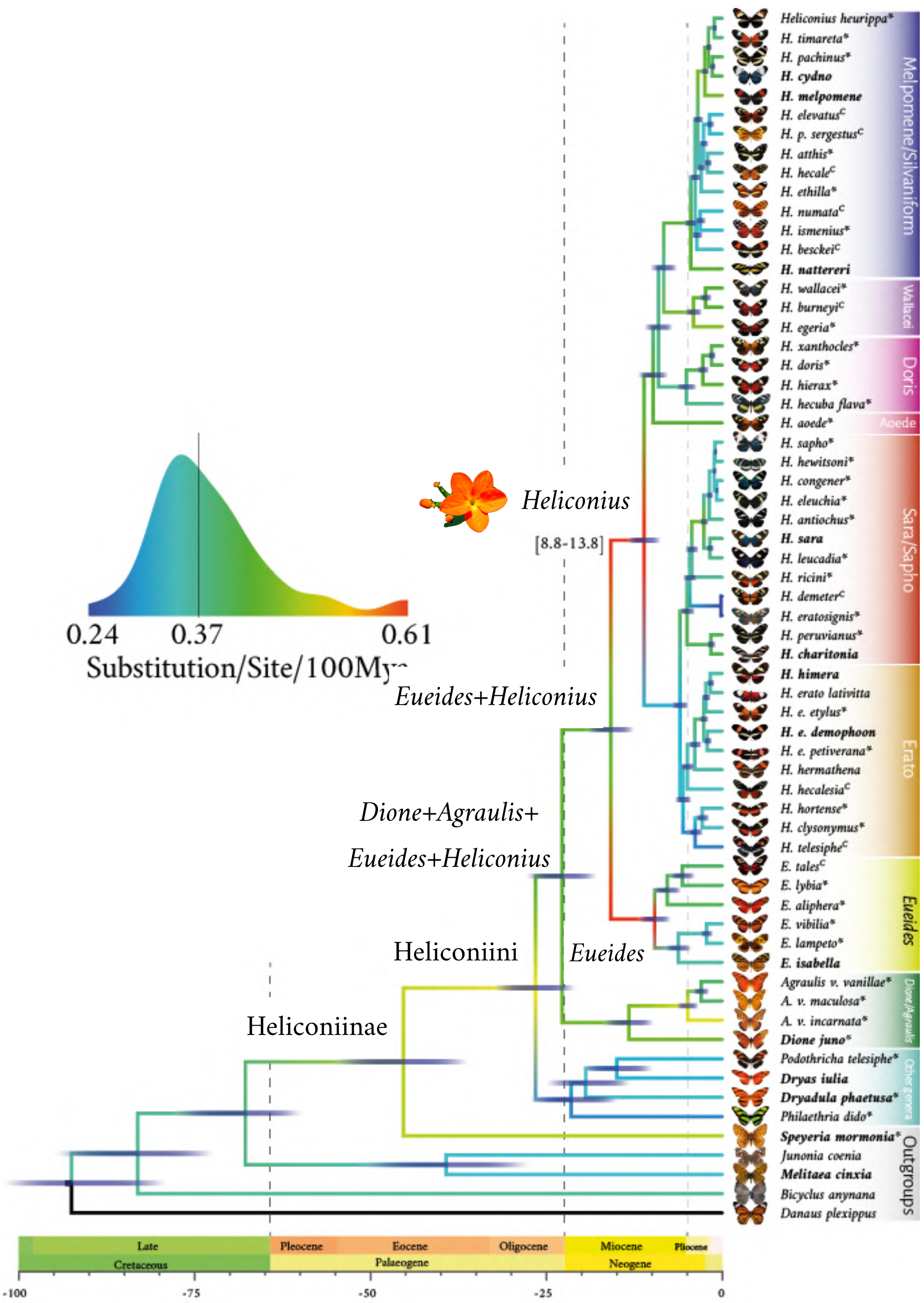
Adaptive Radiation of Heliconiini (Family: Nymphalidae)

Photo credit: @mena_sebas



- 8 Genera
- 87 Species
- ~440 sub-species

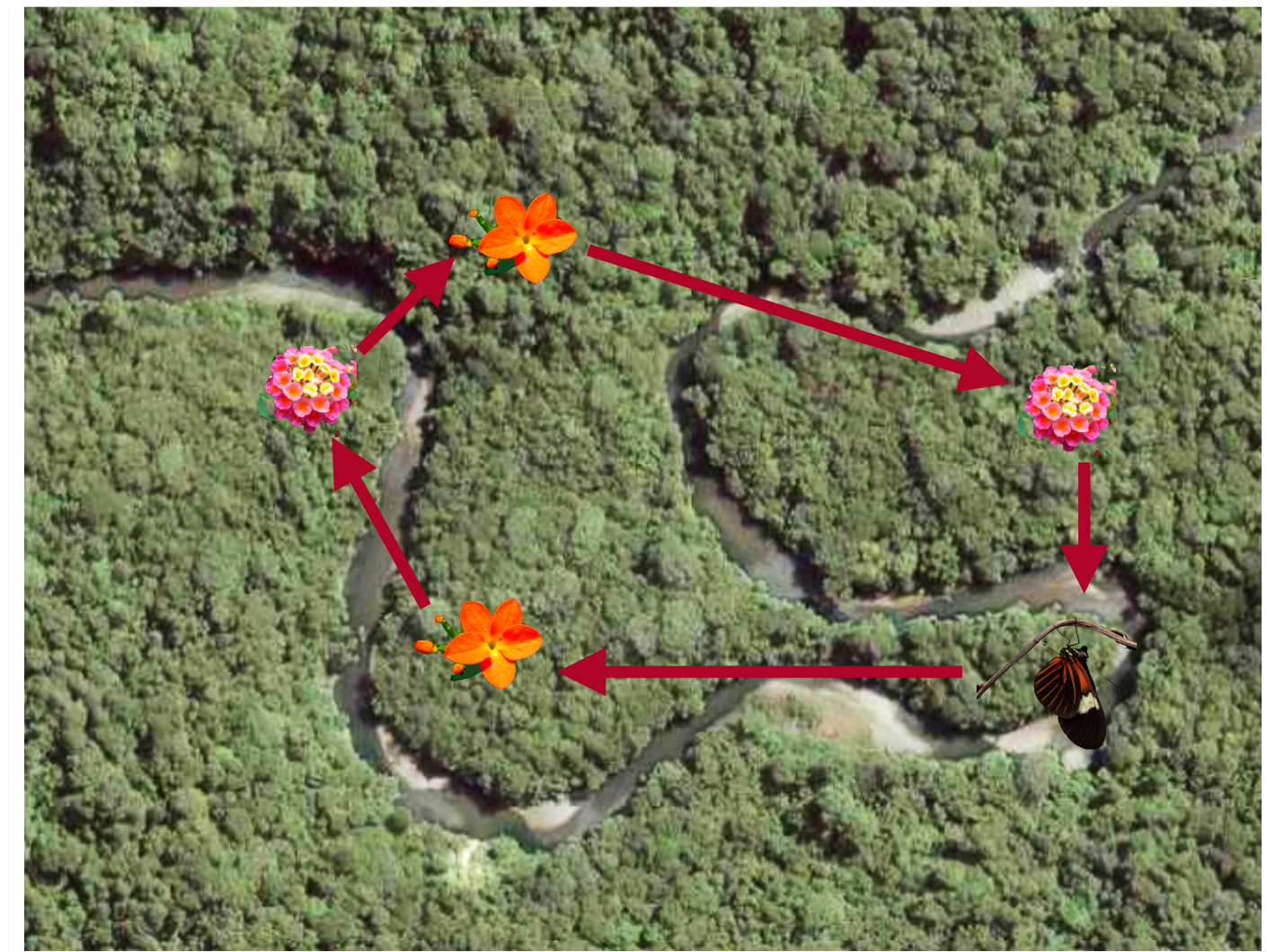




» Pollen-feeding



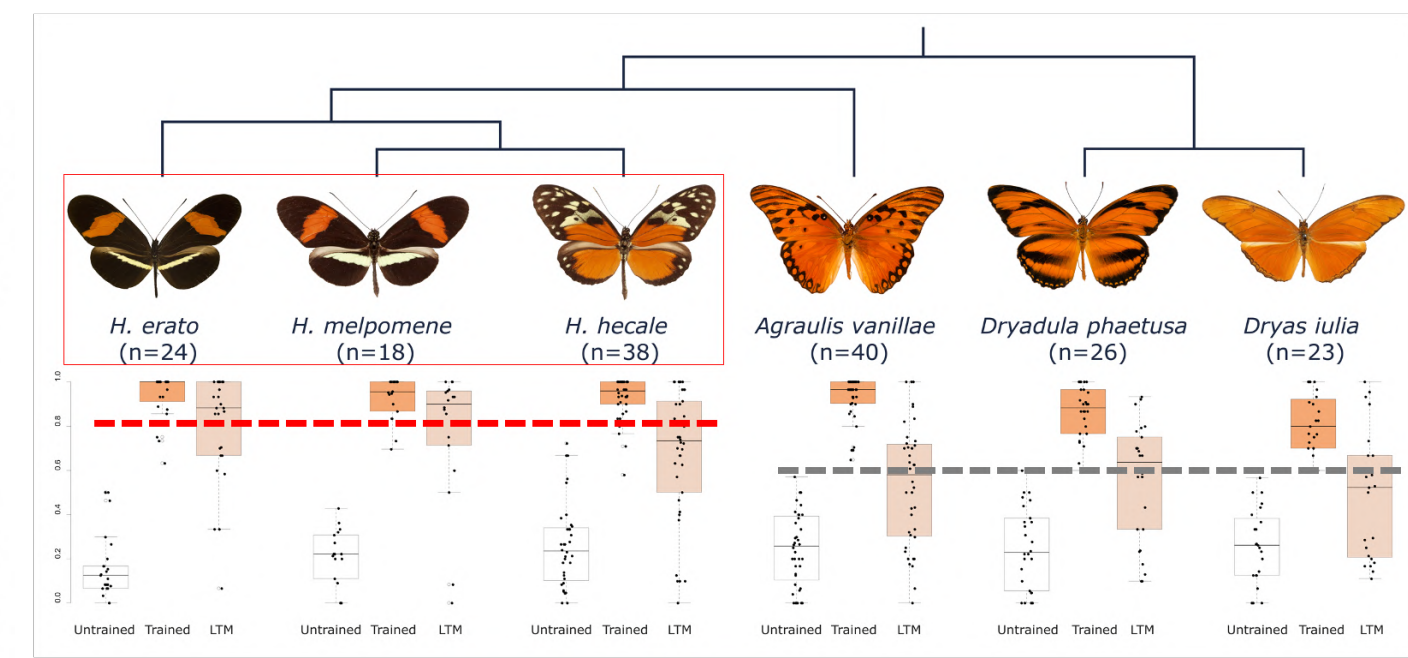
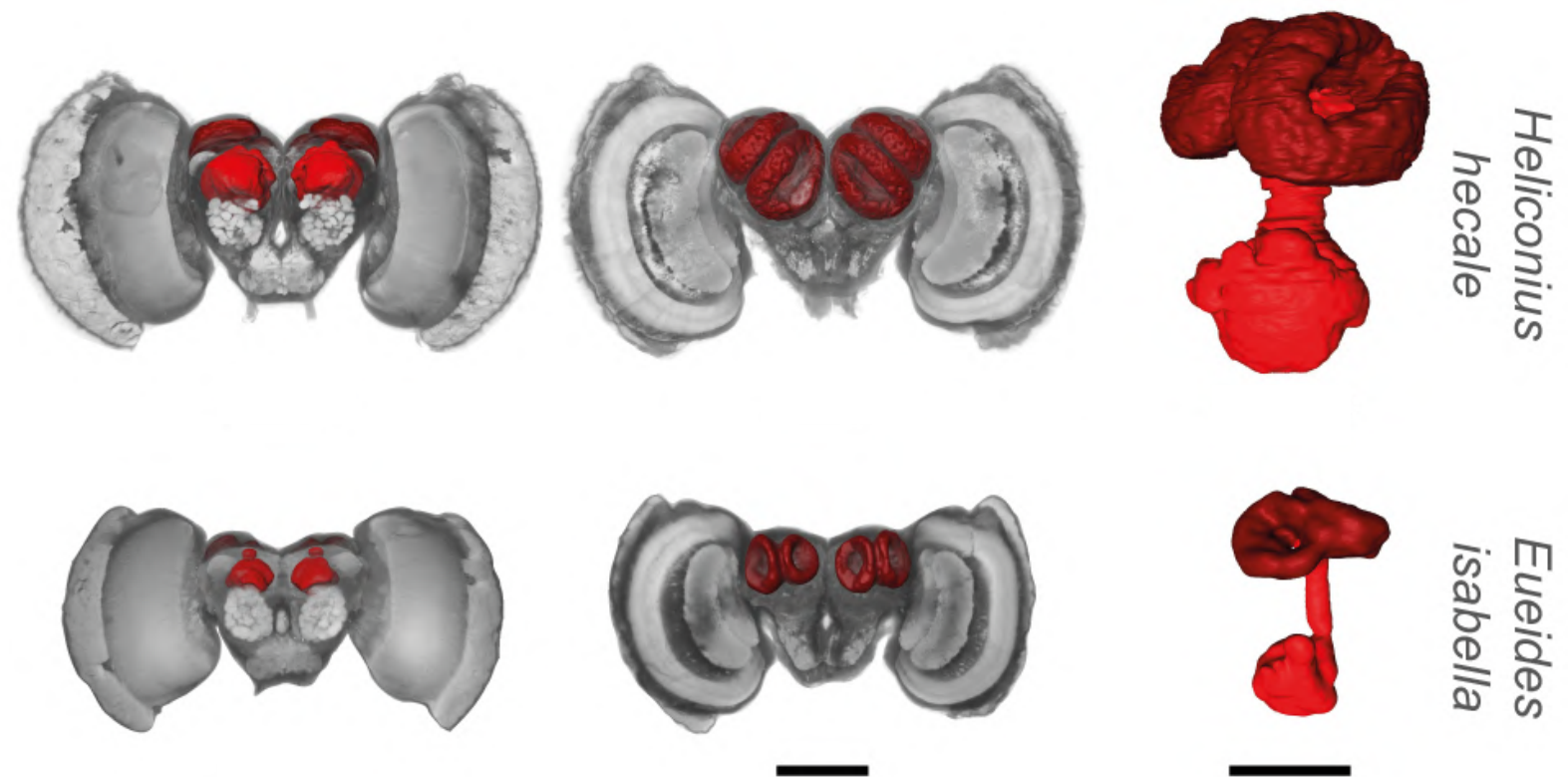
» Trap-lining behaviour



Allocentric trap-lining in home ranges up to 1 km²

Moura et al., 2021. *Functional Ecology*.

» Mushroom body expansion (Brain)



Interspecific variation: $X^2 = 76.2, p < 0.001$
 Small vs. big MBs: $X_8 = 182.8, p < 0.001$

» Whole Genome alignment »

» Conserved non-coding Elements (CNEs) »

» CNEs enrichment (extra) »

» Whole Genome alignment »



» Whole Genome alignment »

- > **Species-tree inference**
- > **Comparative gene annotation**
- > **Detection of selection**

- **Multi-species map of genomic regions to a corresponding region in each other genome.**
- **Taking into account complex rearrangements and copy number changes.**

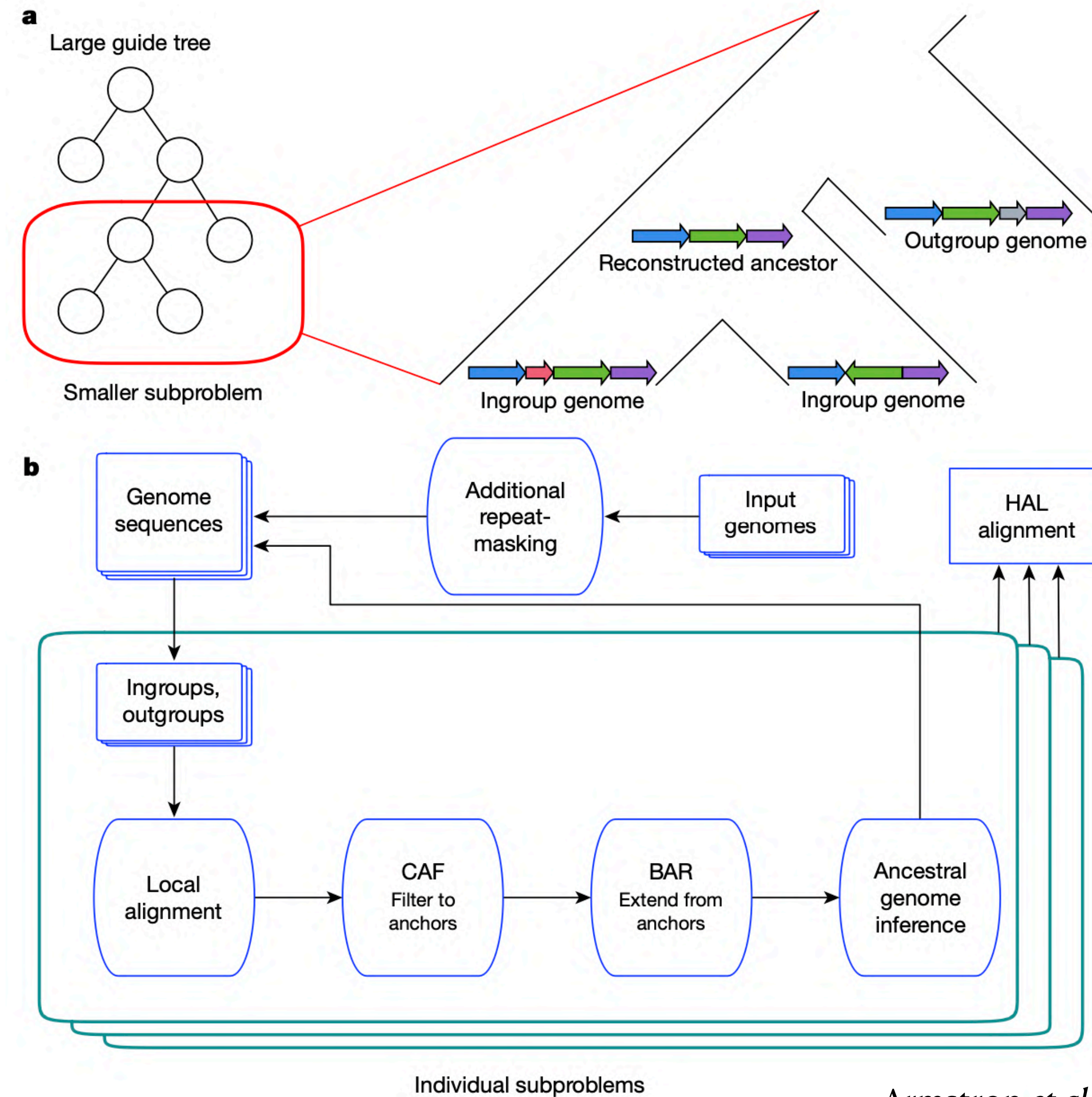
Common limitations are ‘reference bias’:

- Constrains a multiple alignment to only regions present in reference genome.
- Restricting the alignment to be ‘single-copy’, determining miss multiple-orthology relationships.

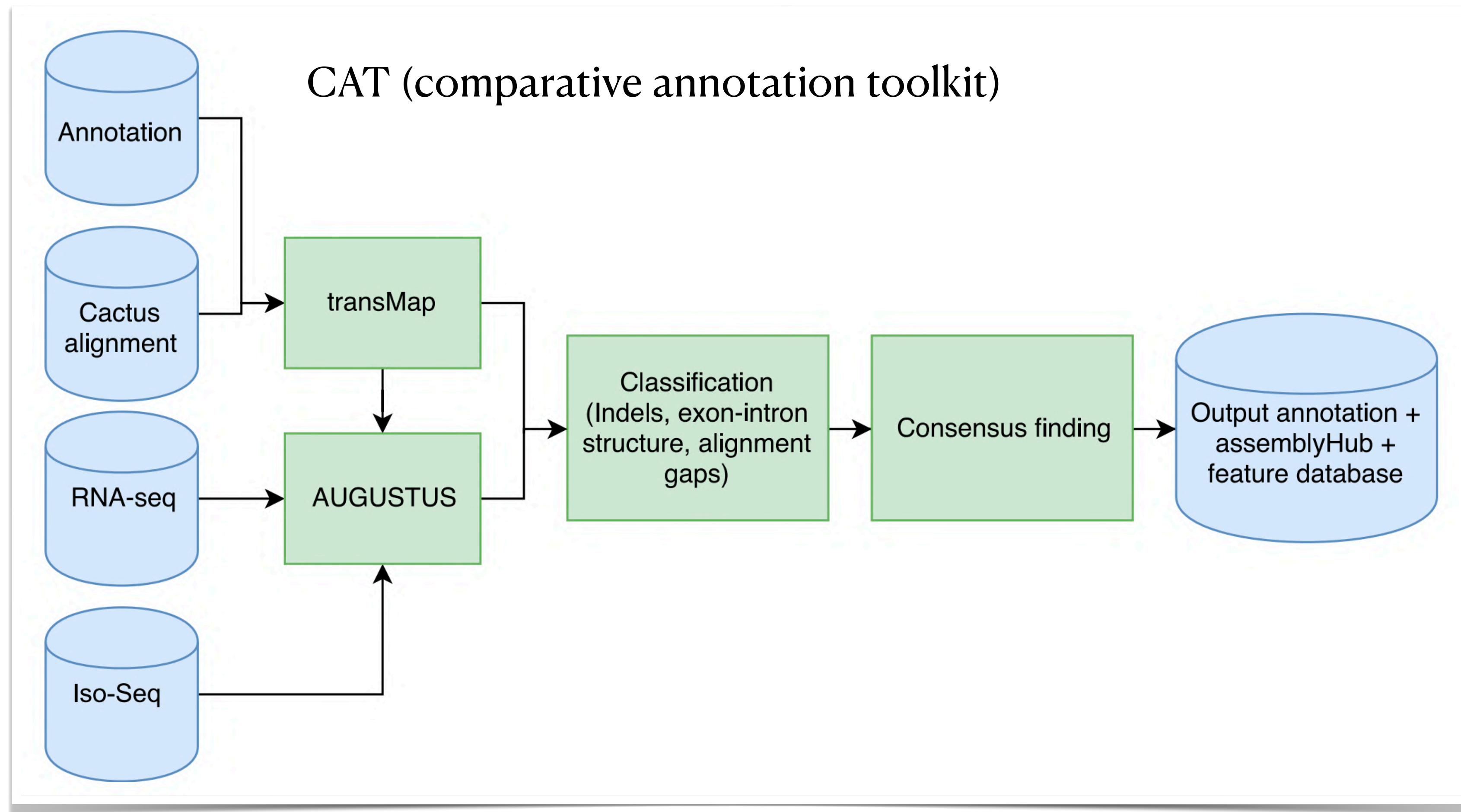
Cactus (ProgressiveCactus) is a “reference *free*” whole genome aligner.



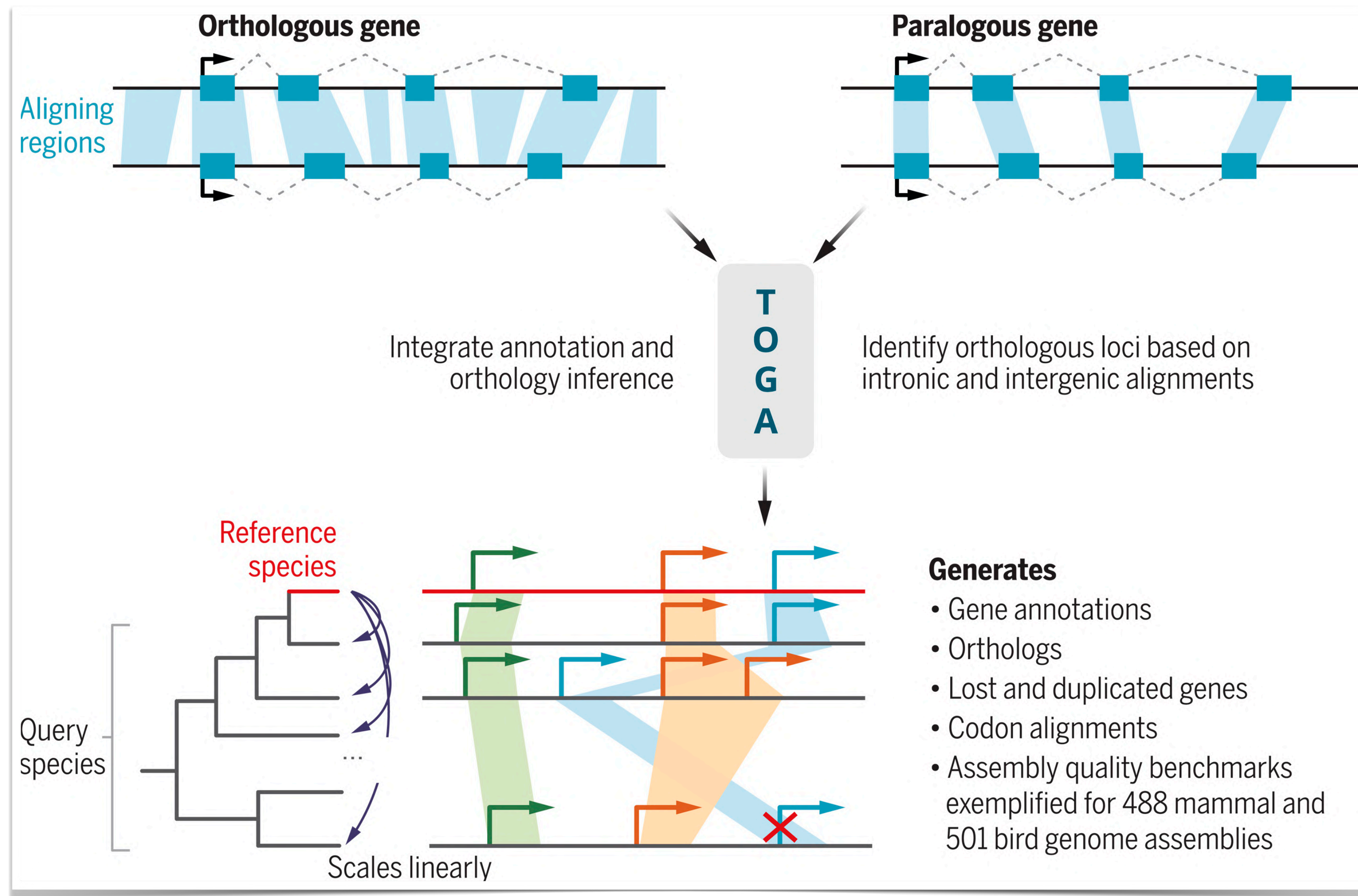
» Whole Genome alignment »



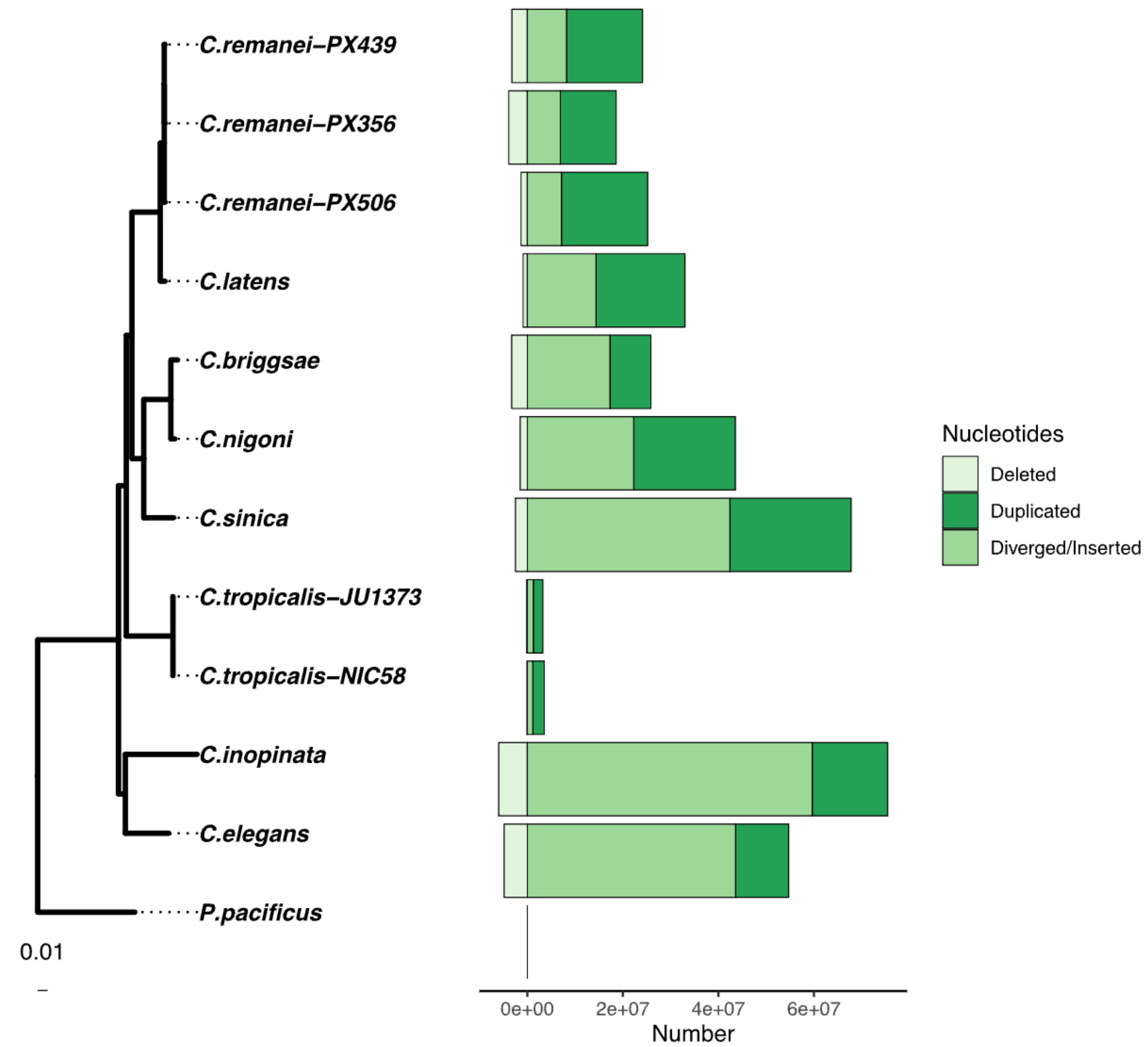
» Some of the applications (annotation) »



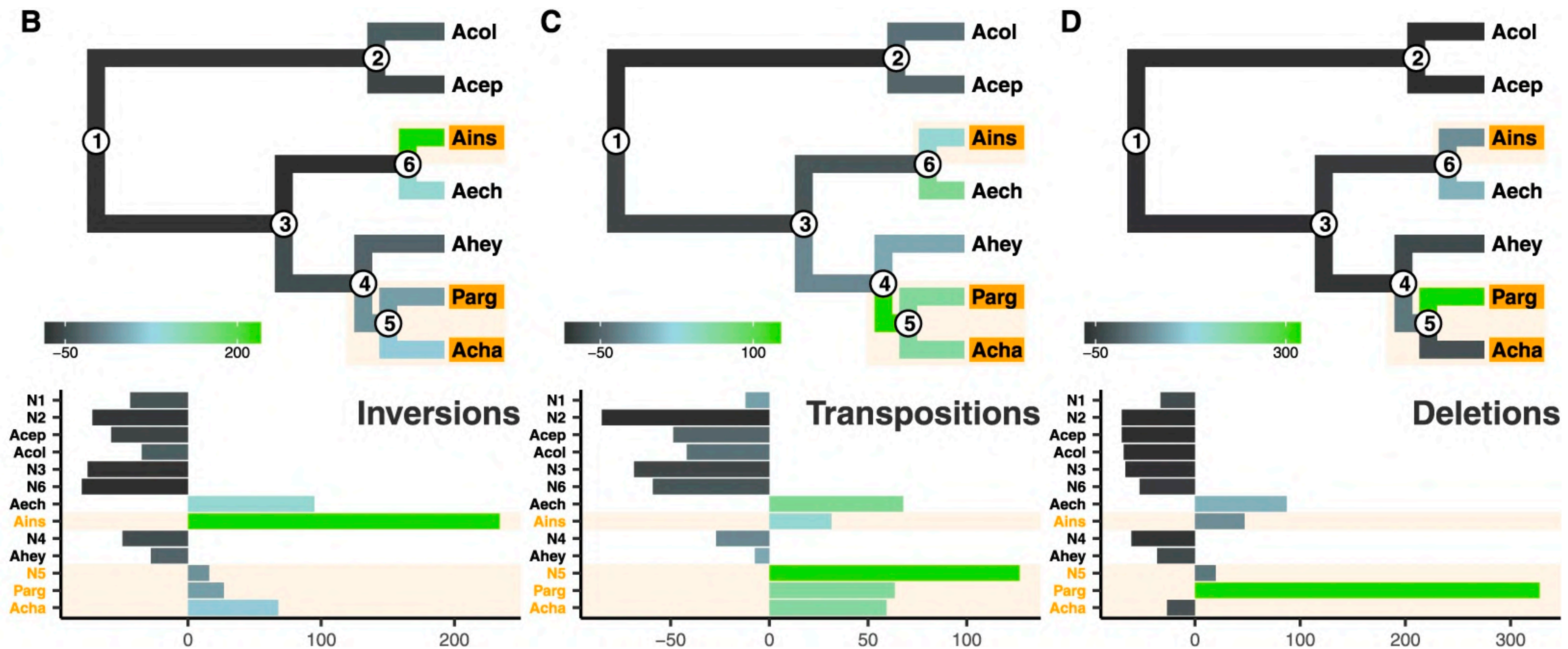
» Some of the applications (orthology inference) »



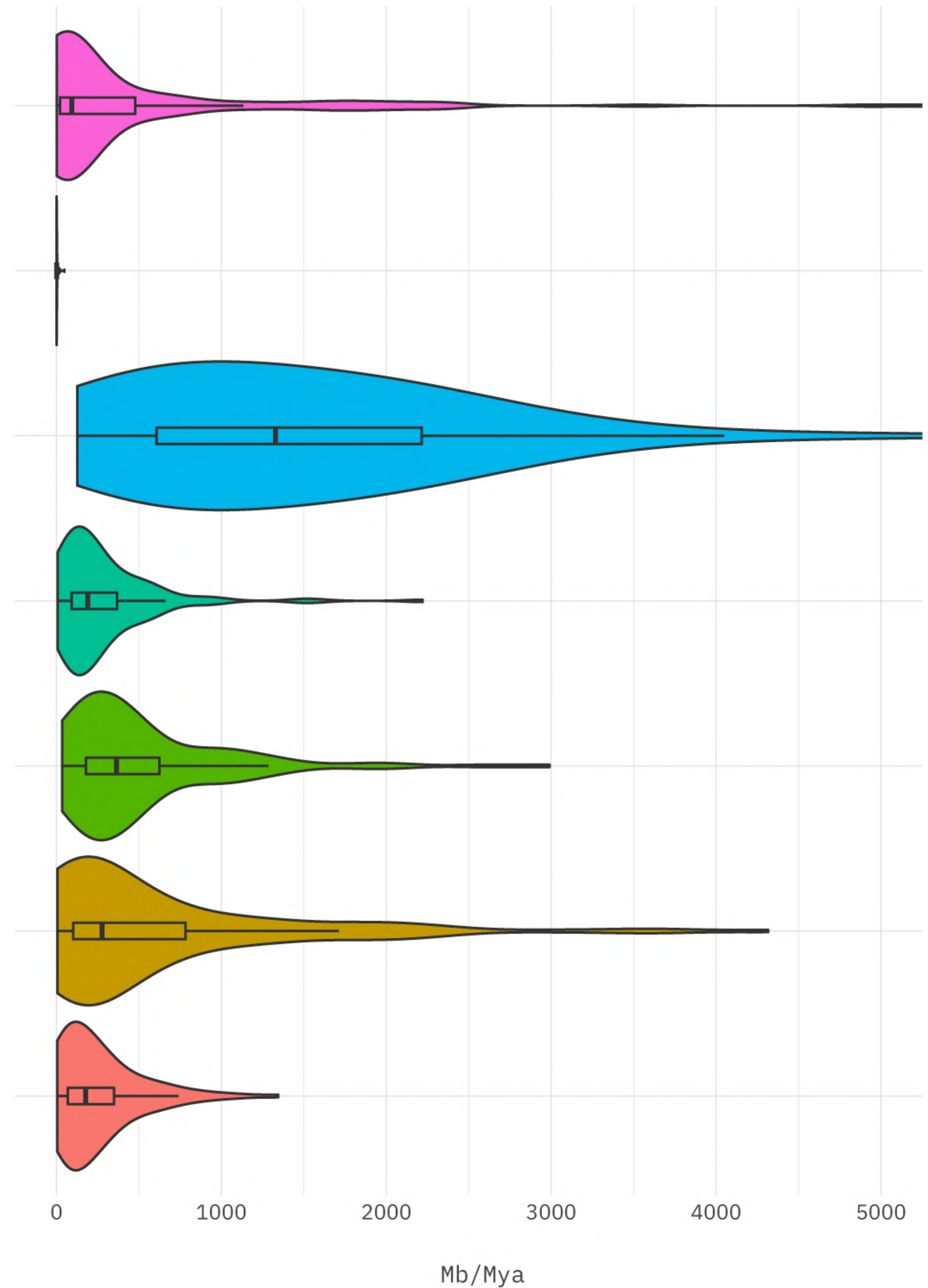
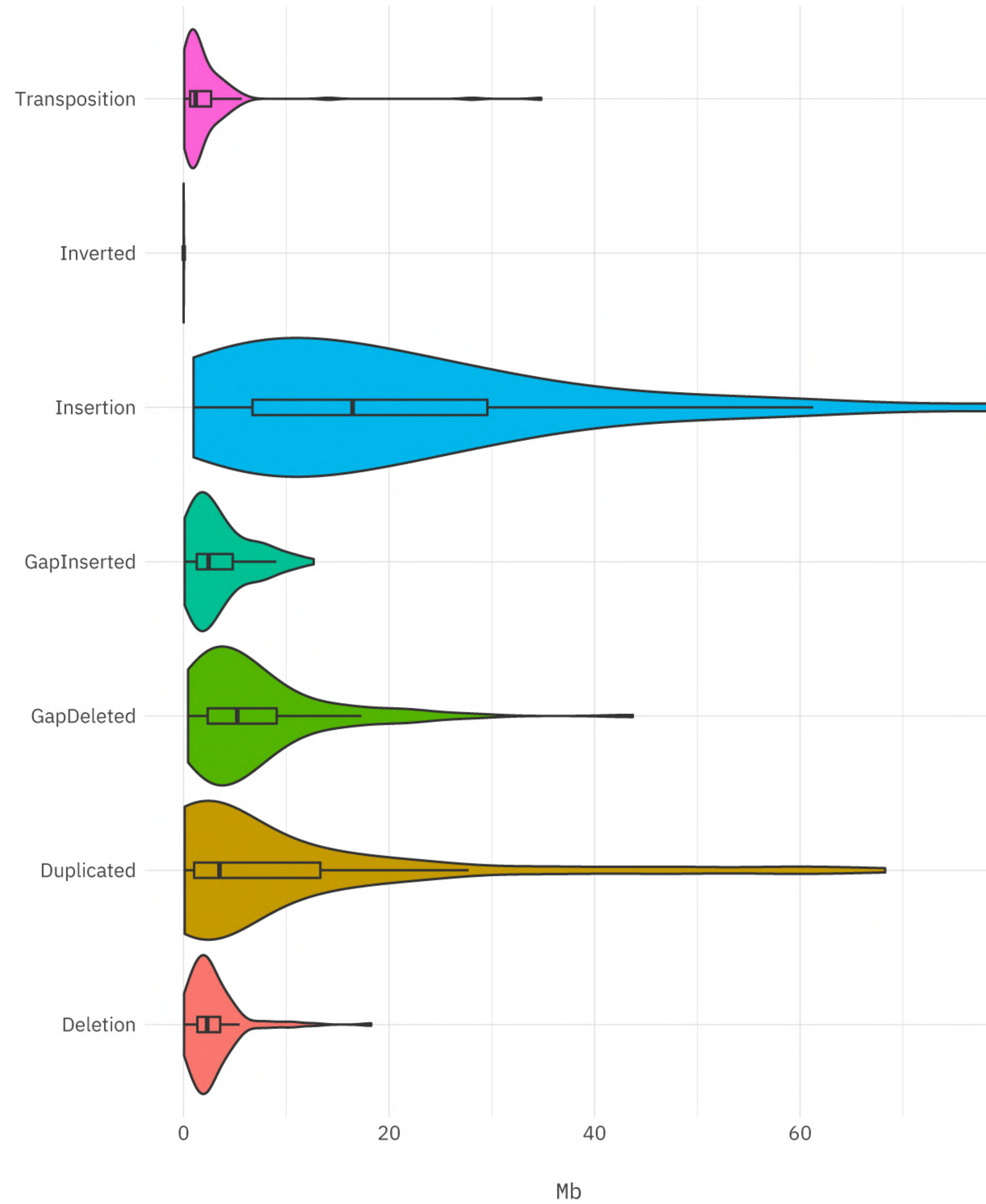
» Structural Rearrangements »

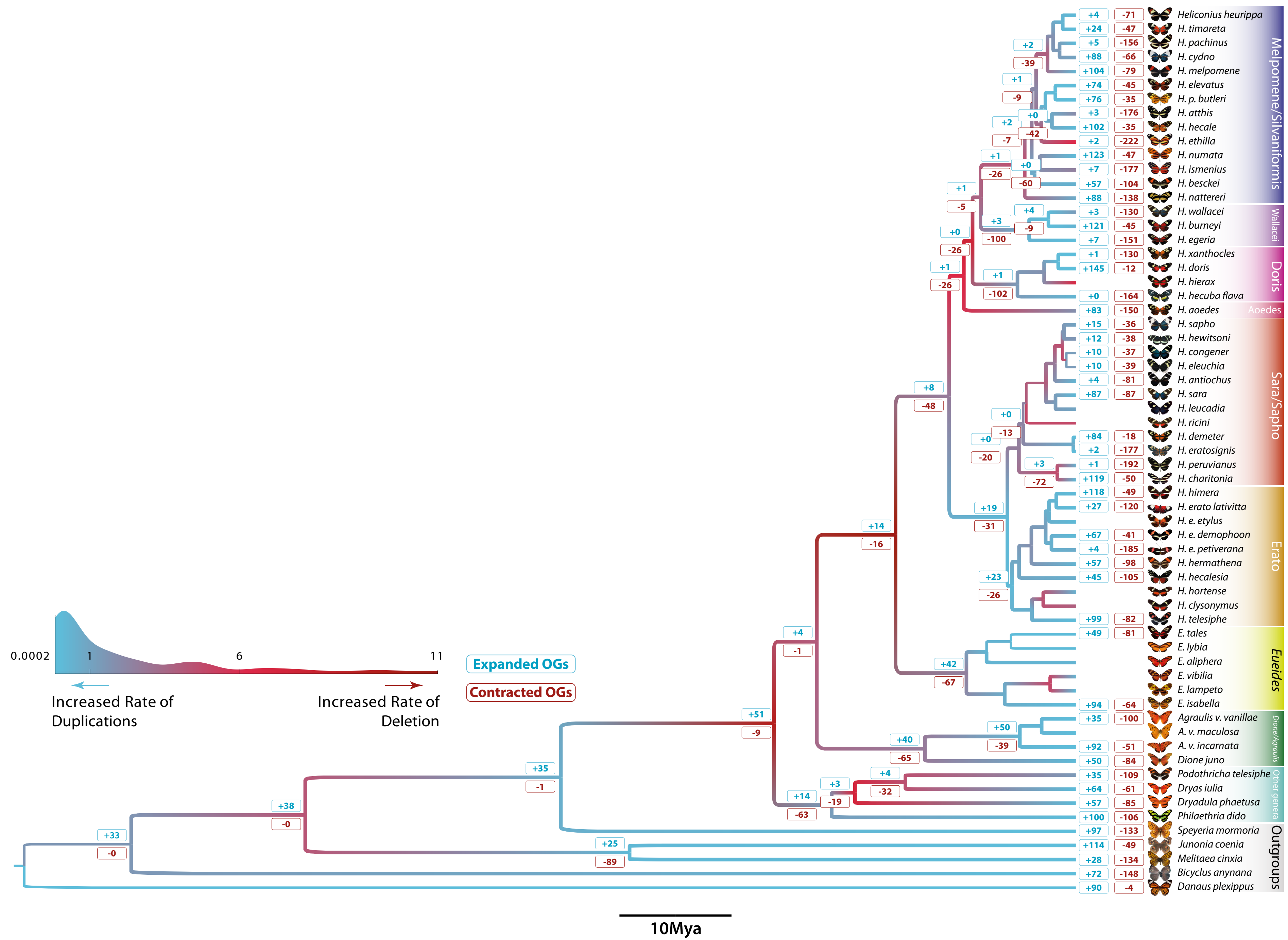


Adams *et al* 2023 MBE



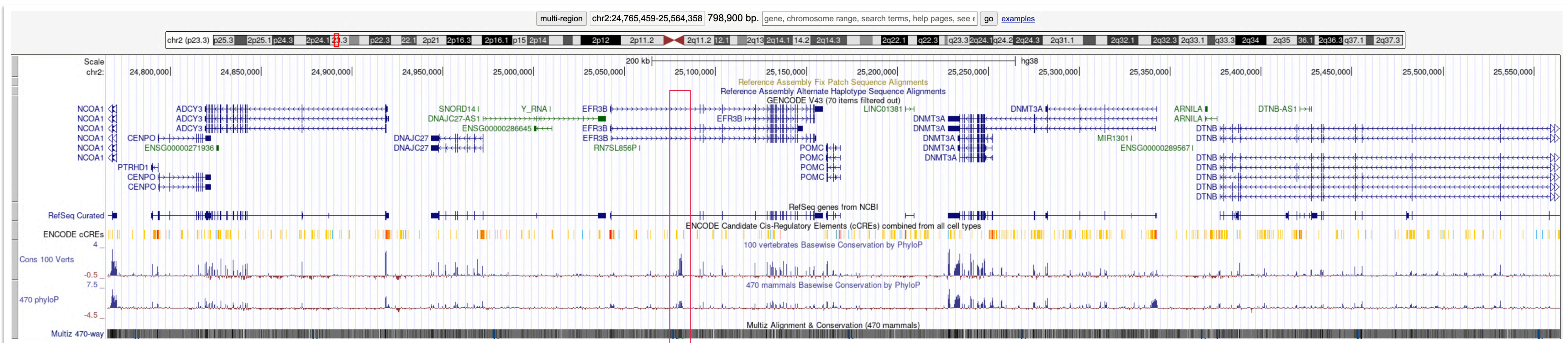
Schrader *et al* 2021 *Nature Comm.*





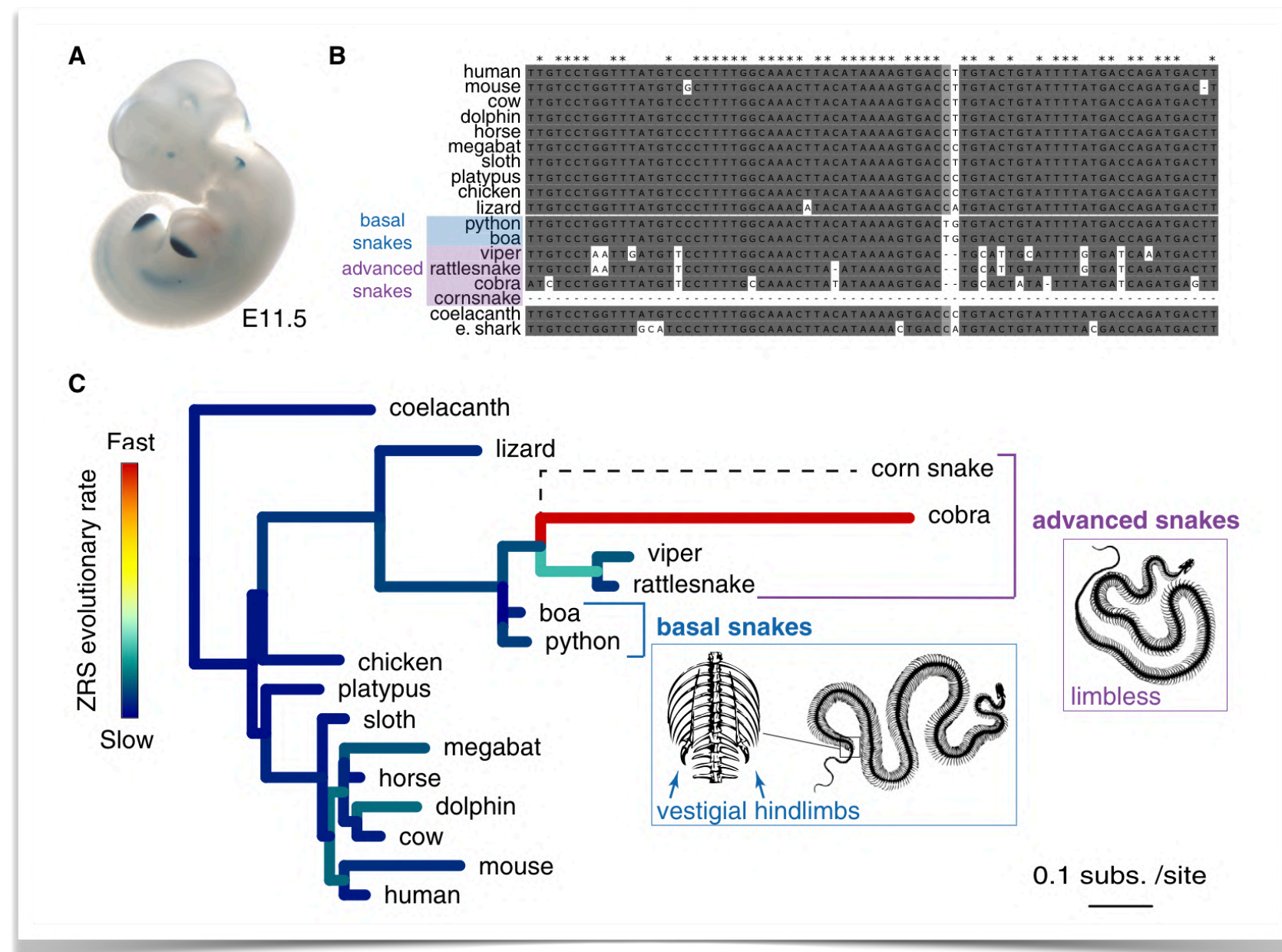
» Conserved non-coding Elements (CNEs) »

- > A class of *non-protein-coding* genomic sequences with elevated degree of conservation.
- > CNEs are non-randomly distributed, clustering in the vicinity of genes with regulatory roles.
- > Organised into functional ensembles (regulatory blocks), which coordinate the expression of shared target genes.
- > The disruption of these elements contribute to diseases linked with development, and cancer.

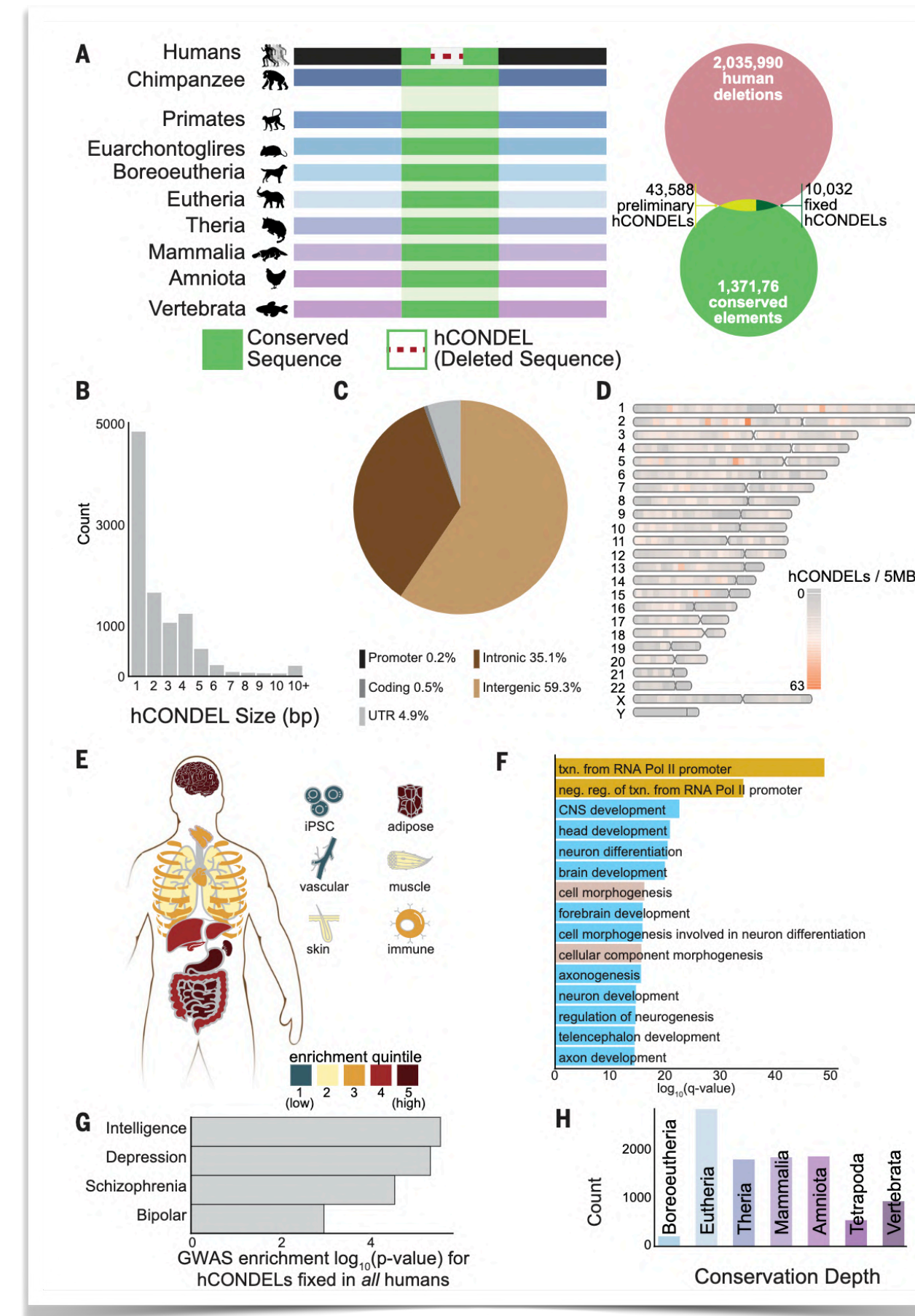


Candidate Cis-Regulatory Elements (cCREs)

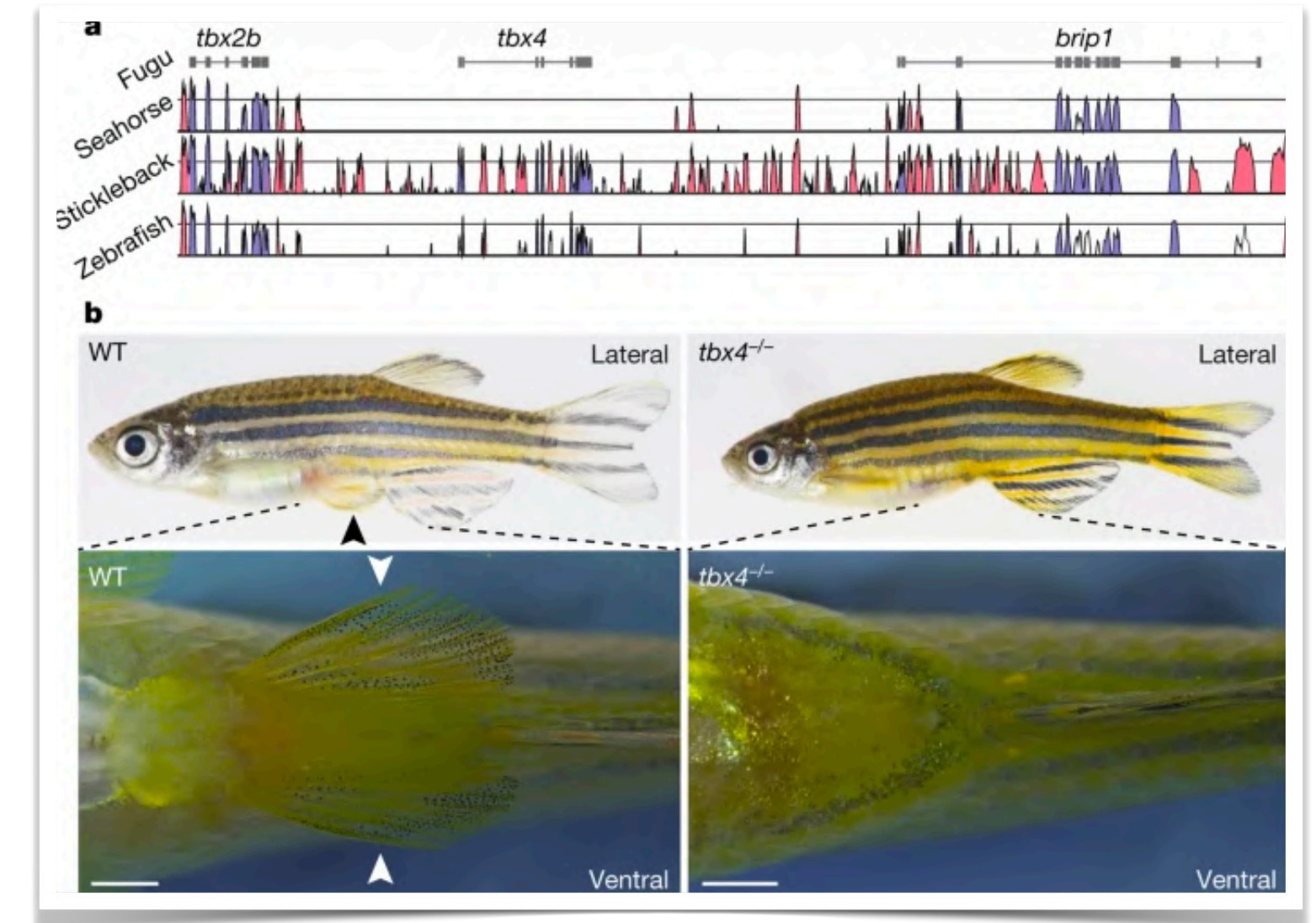
» Conserved non-coding Elements (CNEs) »



Kvon et al 2016 Cell



Xue et al 2023 Science



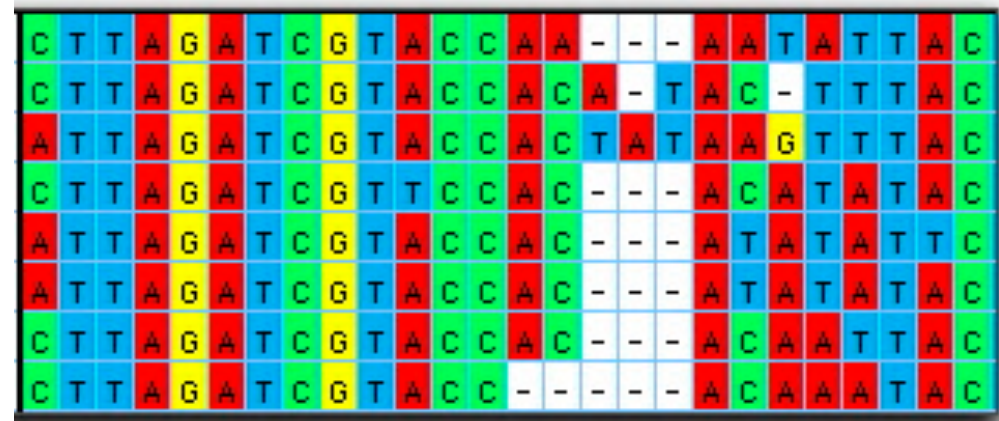
Lin et al 2016 Nature

» How to identify these regions »

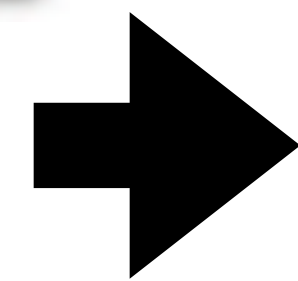
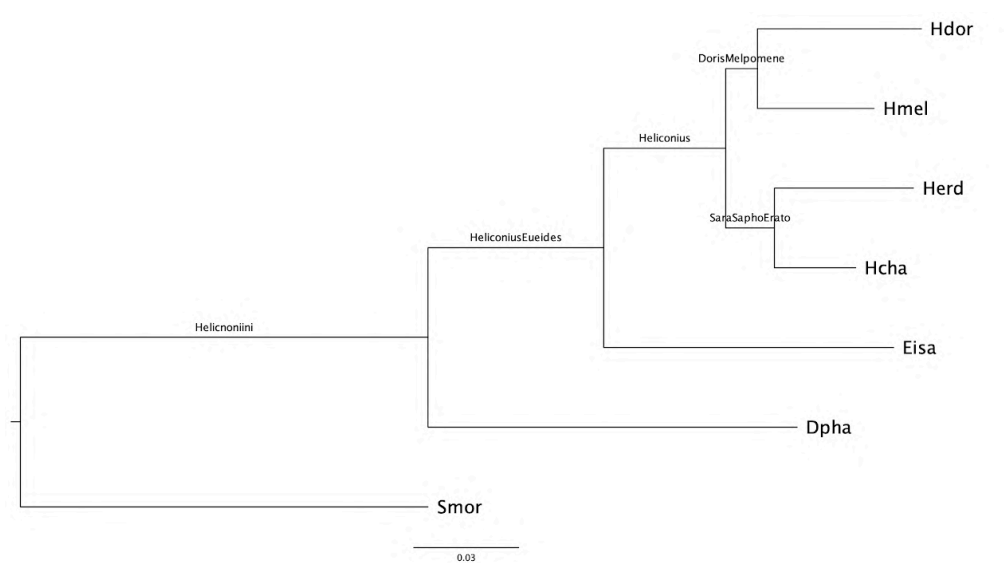
- 1. A Model of DNA sequence evolution**
- 2. Phylogenetic tree (wait until tomorrow)**

» How to identify these regions »

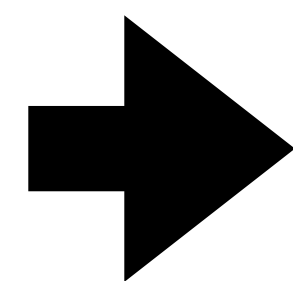
Neutral evolving regions



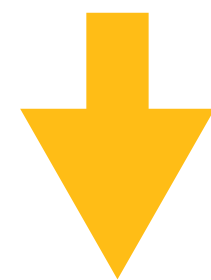
+



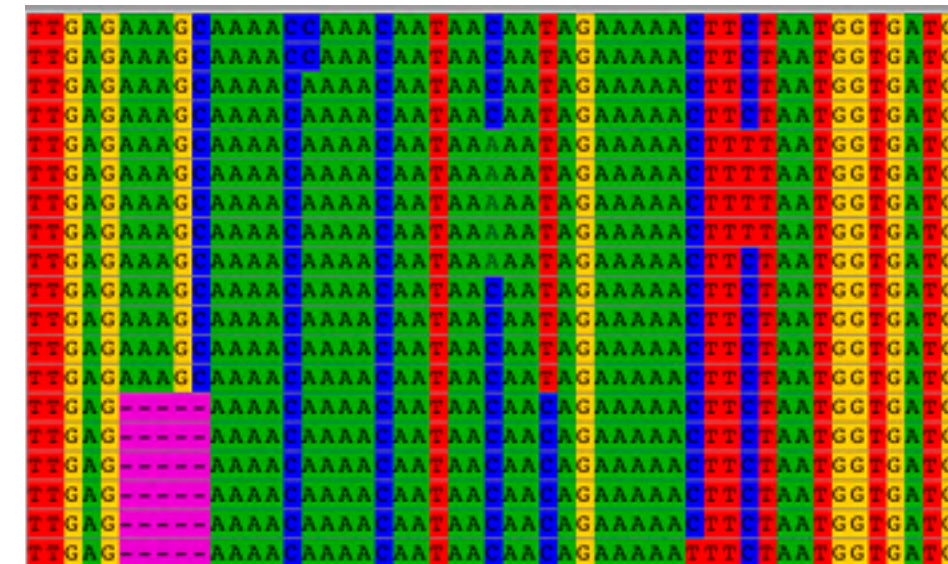
Neutral model



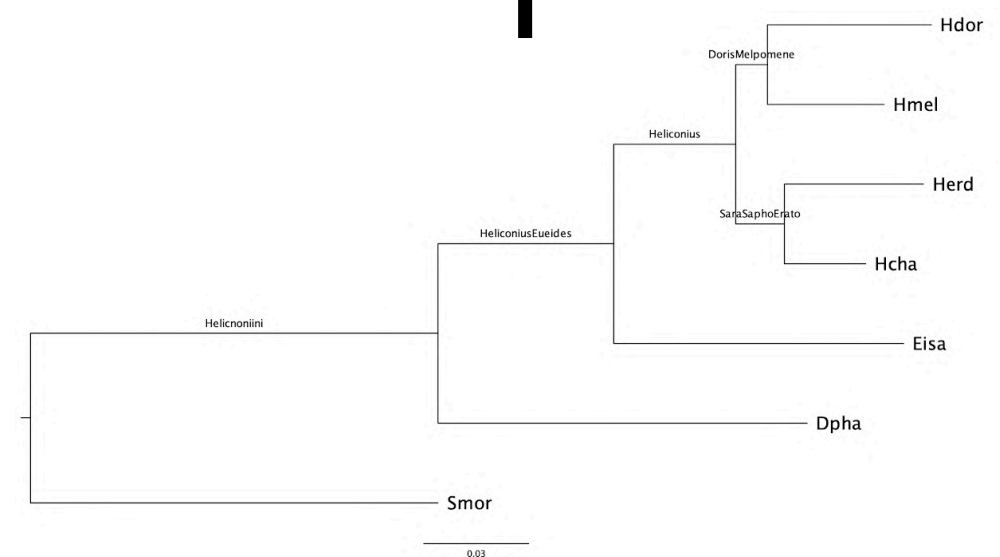
Other metrics
(*PhyloP*)



WGA

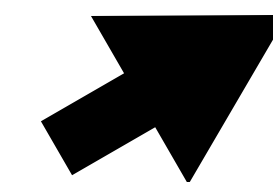


+

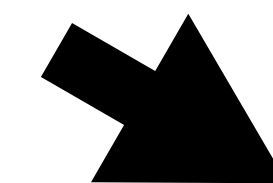


+

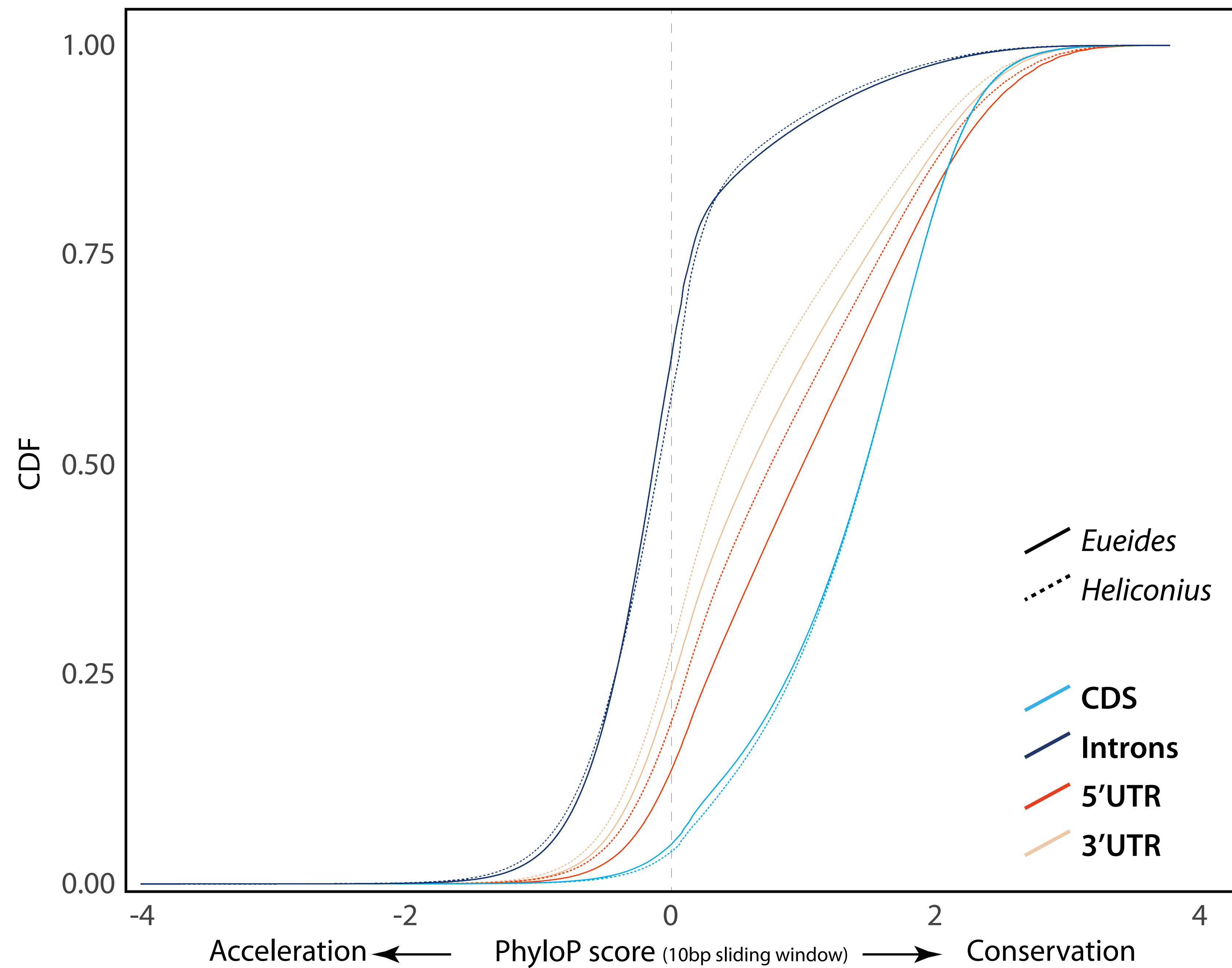
neutral model



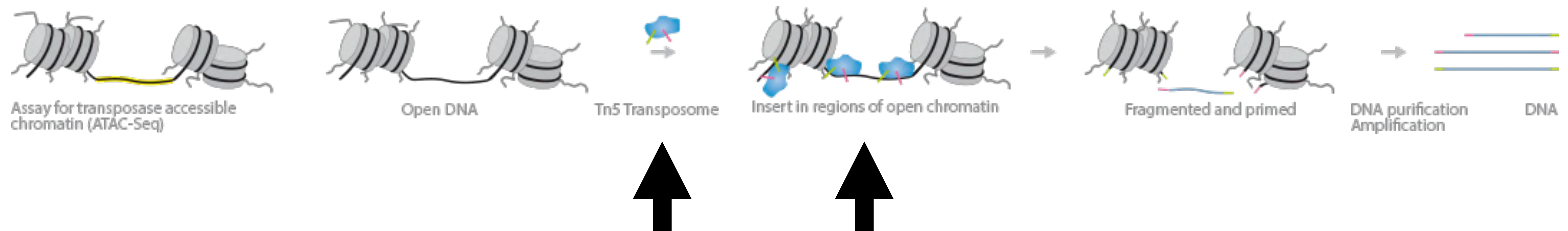
Conservation model



Non-conserved model



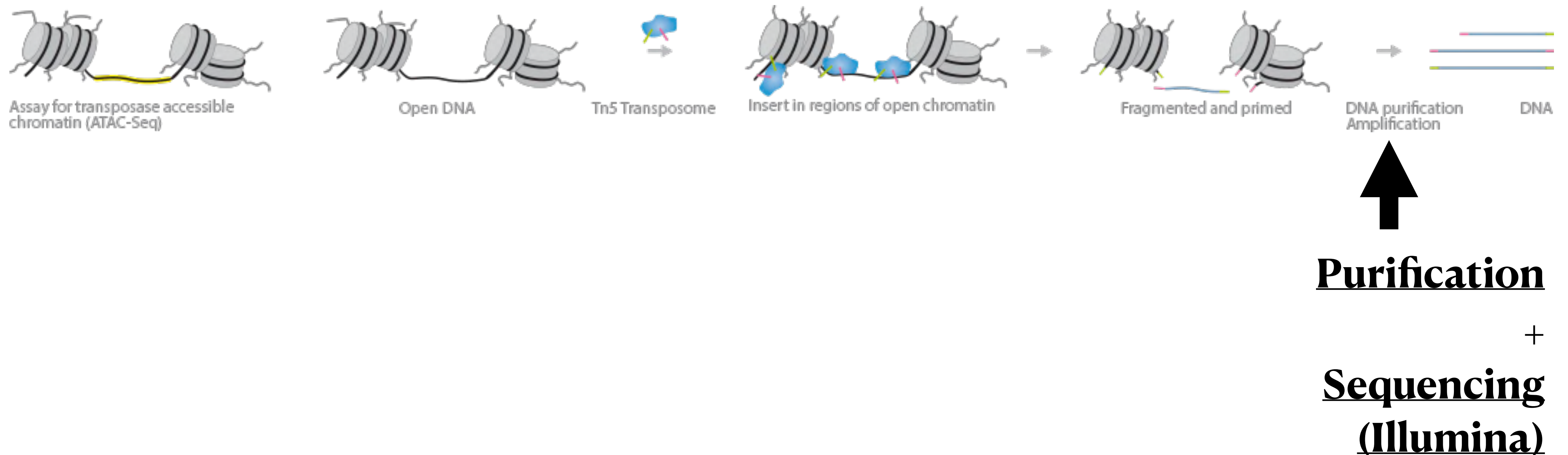
» ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) »



Tagmentation

The activity Tn5 (hyperactive) Transposase that inserts sequencing adapters into open regions of the genome and cleaves

» ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) »



» ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) »

