

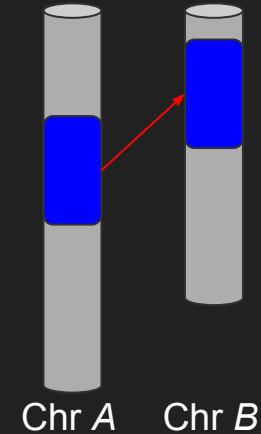
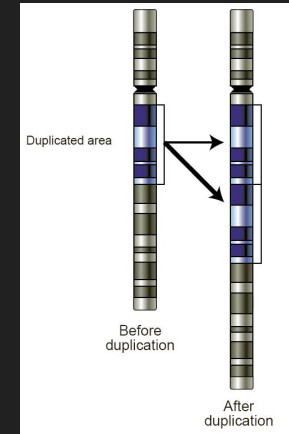
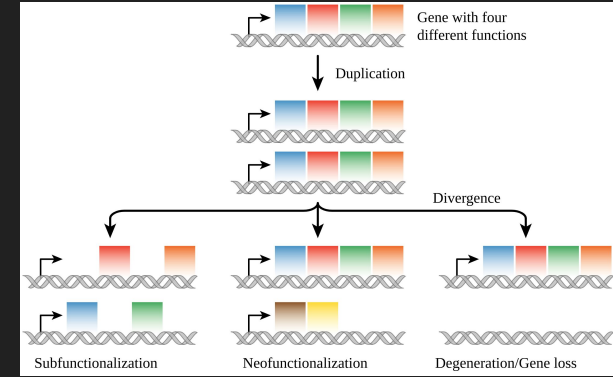
Evolution after duplication

The fate of gene function studied in eight Brassicaceae

Asis Hallab, university of applied sciences Bingen

Introduction

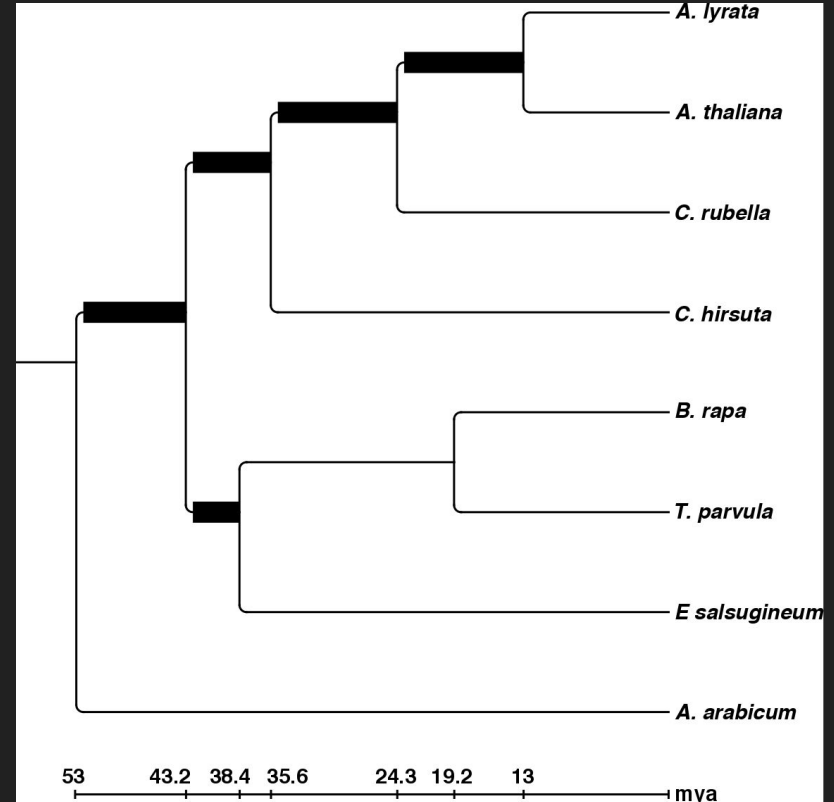
- Gene duplication is a major mechanism through which new genetic material is generated during molecular evolution.
- A duplicate can mutate and change its gene function without reducing fitness through loss of function.
- No large scale comparison of tandem and trans duplication done.



Material

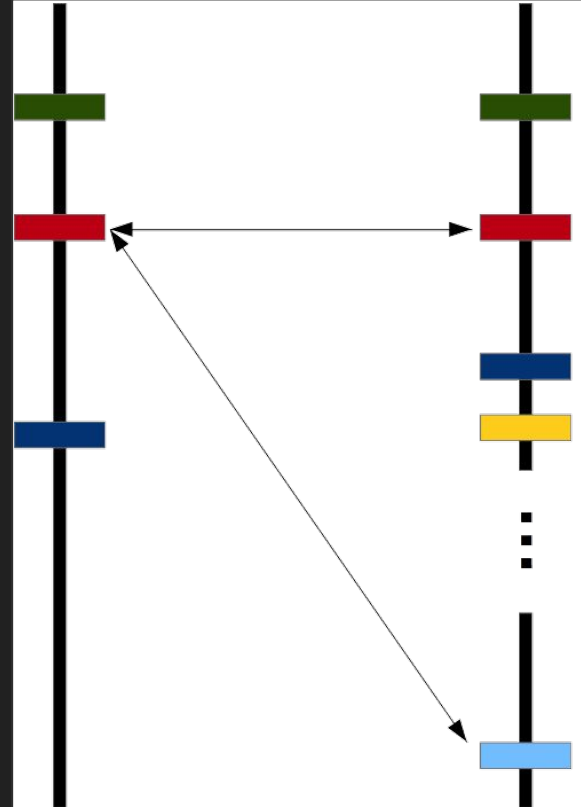
- Eight Brassicaceae genomes
- Coding Sequences

Gan, X. et al. The *Cardamine hirsuta* genome offers insight into the evolution of morphological diversity. *Nature Plants* 2, 16167 (2016).



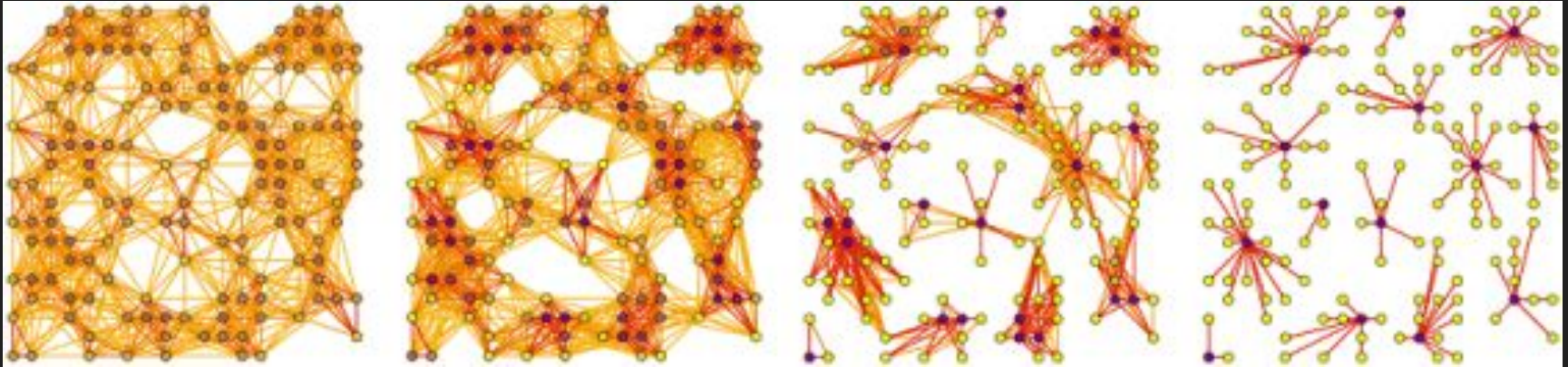
Methods: Ortholog and Tandem detection

- Reciprocal highest local pairwise sequence similarity
- Synteny: Gene neighborhood



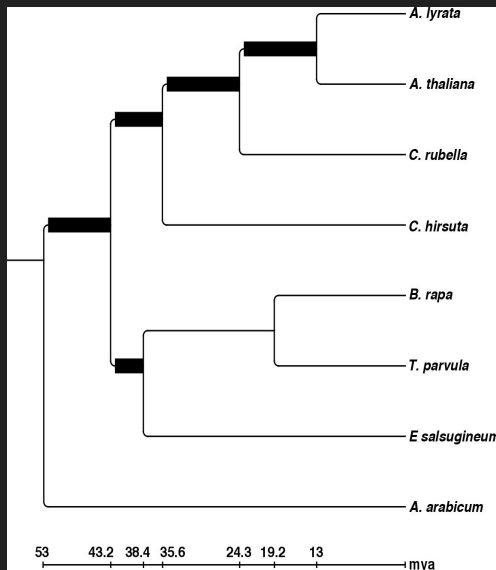
Methods: Gene family reconstruction

- Markov clustering on local pairwise sequence similarity (Blast)

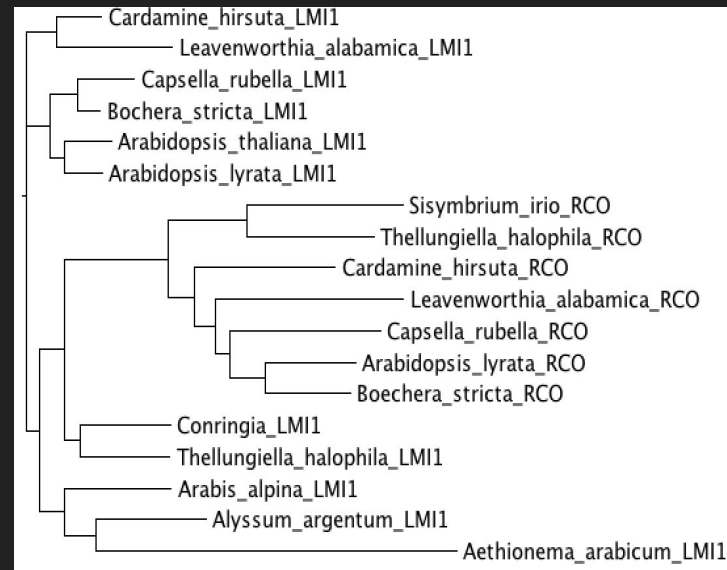


Methods: Expansion / Contraction of gene families

- Ultrametric gene tree (UST)
 - Concatonated MSA of approx. 10,000 ortholog clusters
 - A priori knowledge driven scaling
- Fit gene birth/death rate to UST
- Compare per species gene number with simulated trees



UST



Gene Family Tree



Methods: trans duplicate identification

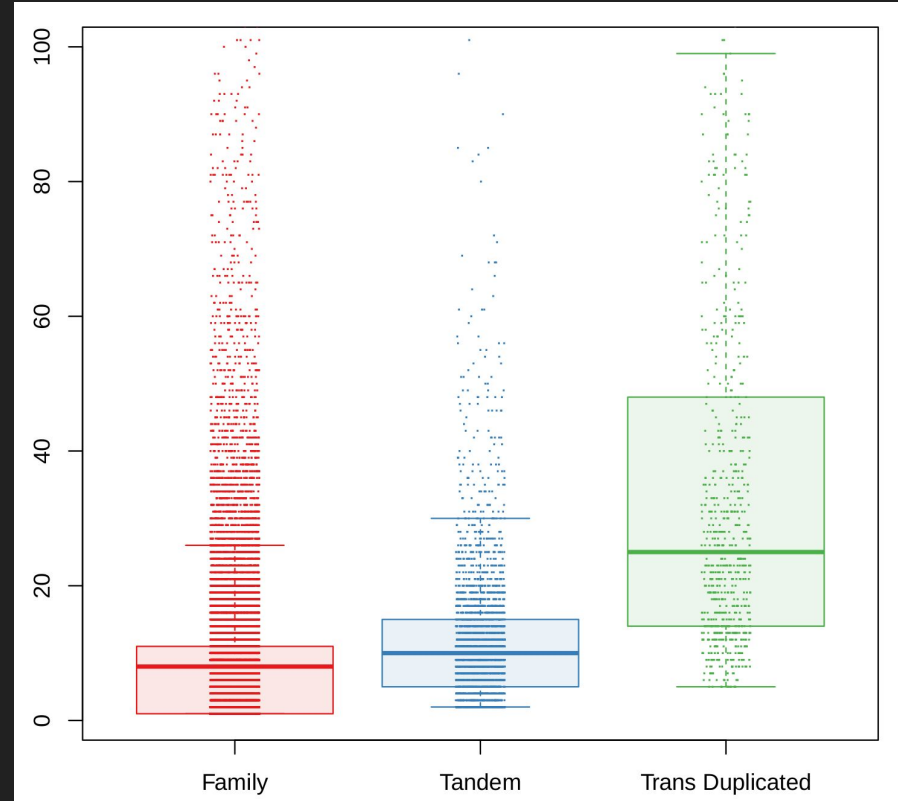
- Member of significantly expanded gene family
- By exclusion
 - Non-Ortholog
 - Non-Tandem



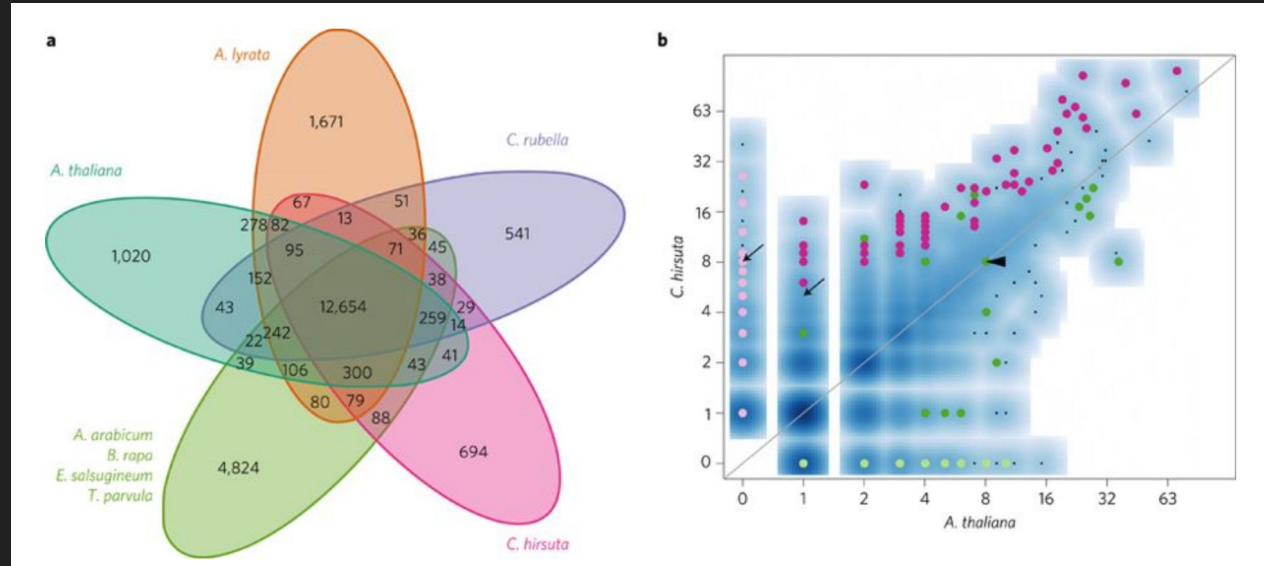
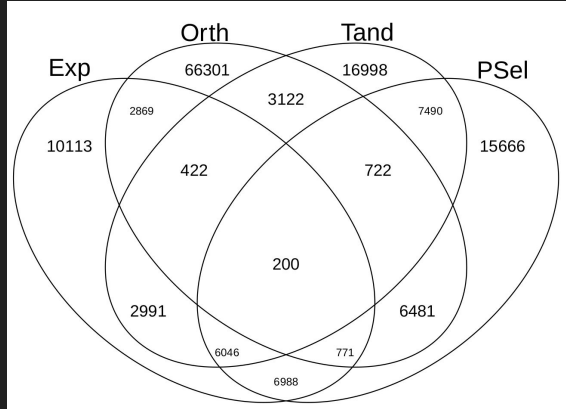
Poor guy - I mean: Gene

Results: Some numbers

- 10,111 groups of orthologous genes
- 16,910 non singleton gene families
- 758 families were significantly expanded or contracted
- 3,266 tandemly duplicated gene clusters

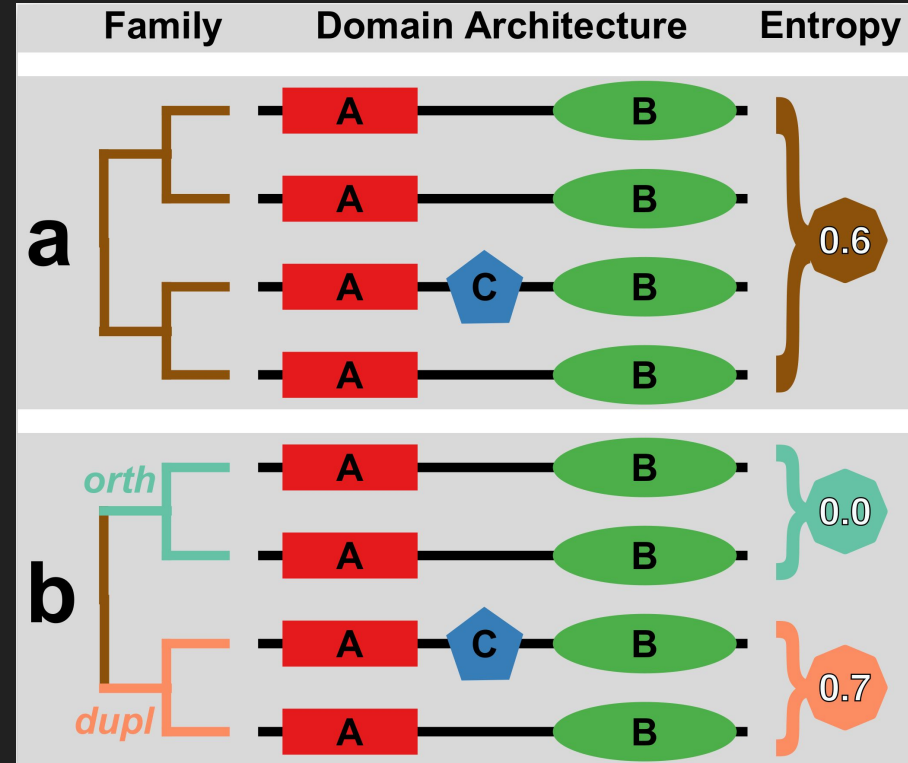


Results: Gene Families

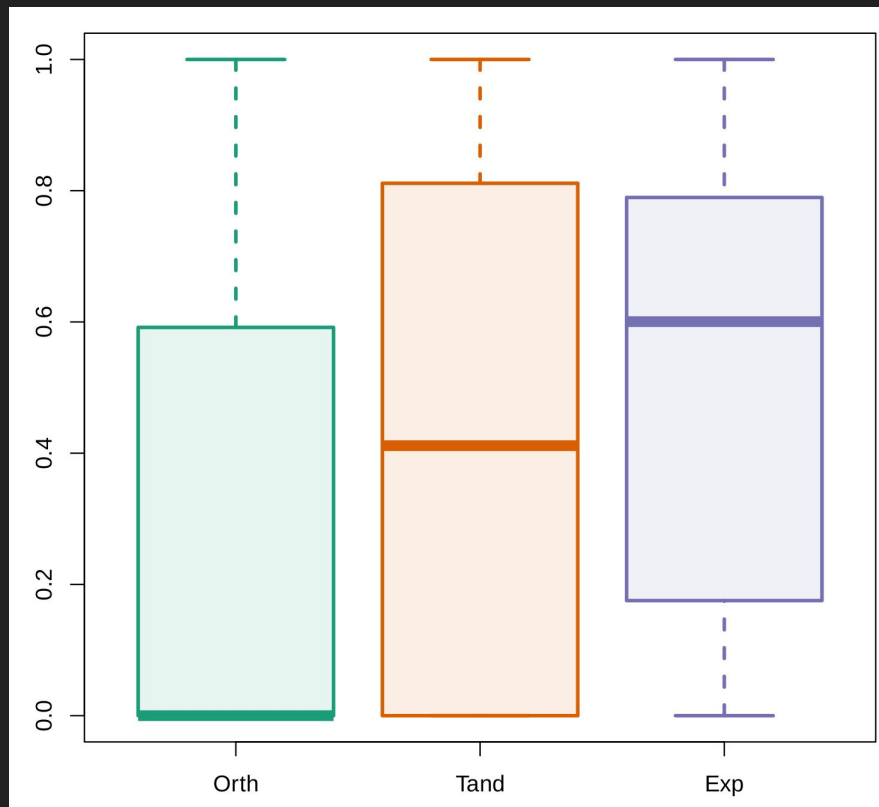
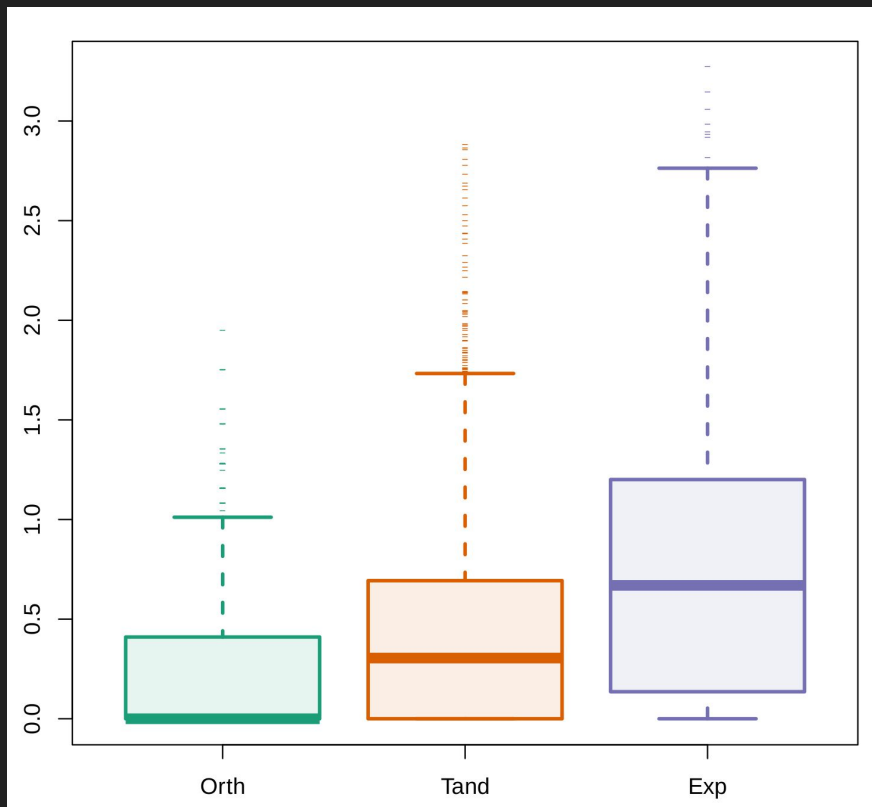


Methods: Protein domain architecture

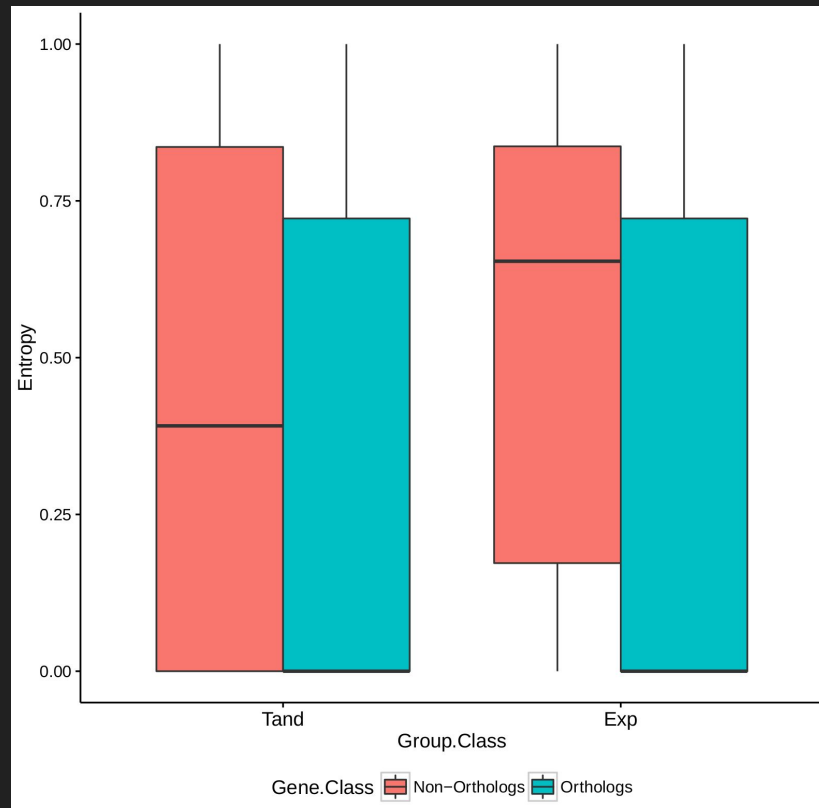
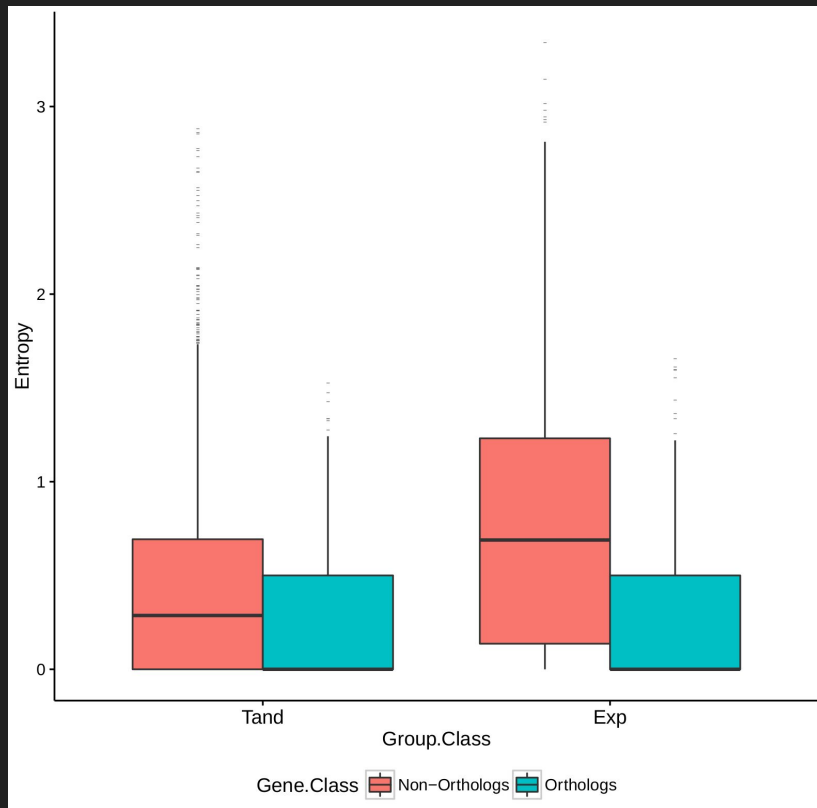
- InterProScan identified conserved protein domains
- Each protein has a unique domain architecture
- Groups of genes (families, ortholog clusters, tandem clusters) have diversity of domain architecture
 - Shannon Entropy



Results: Domain Architecture Diversity

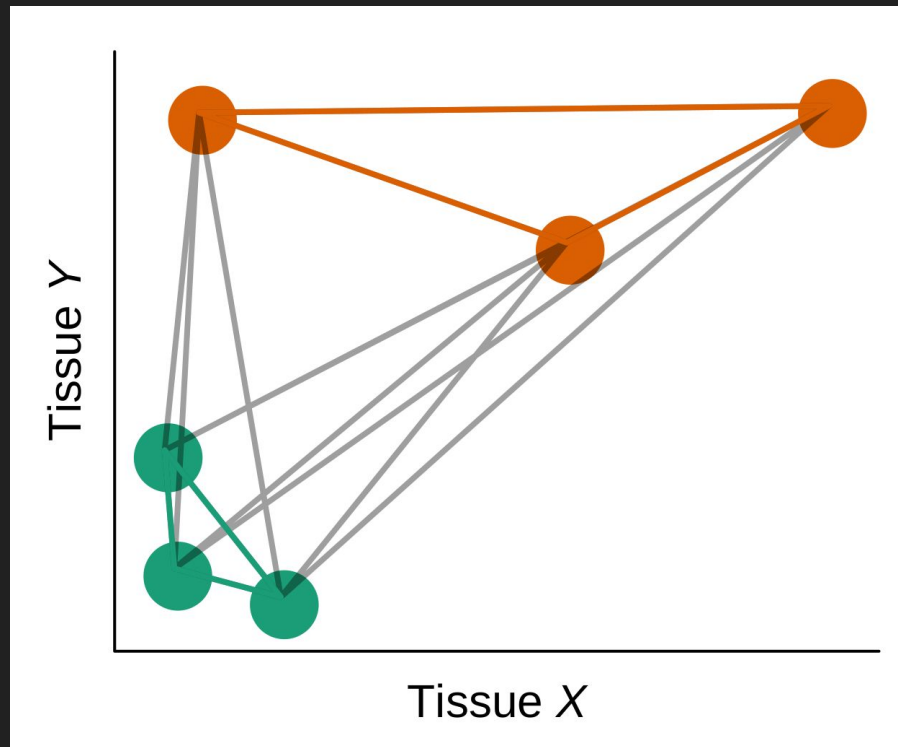


Results: Intra-Group domain architecture diversity

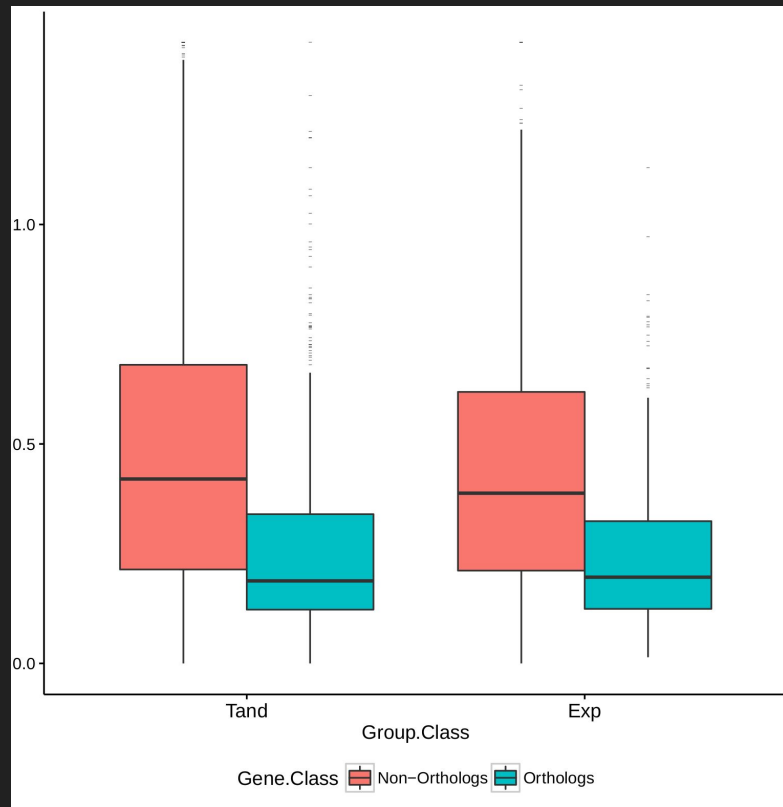
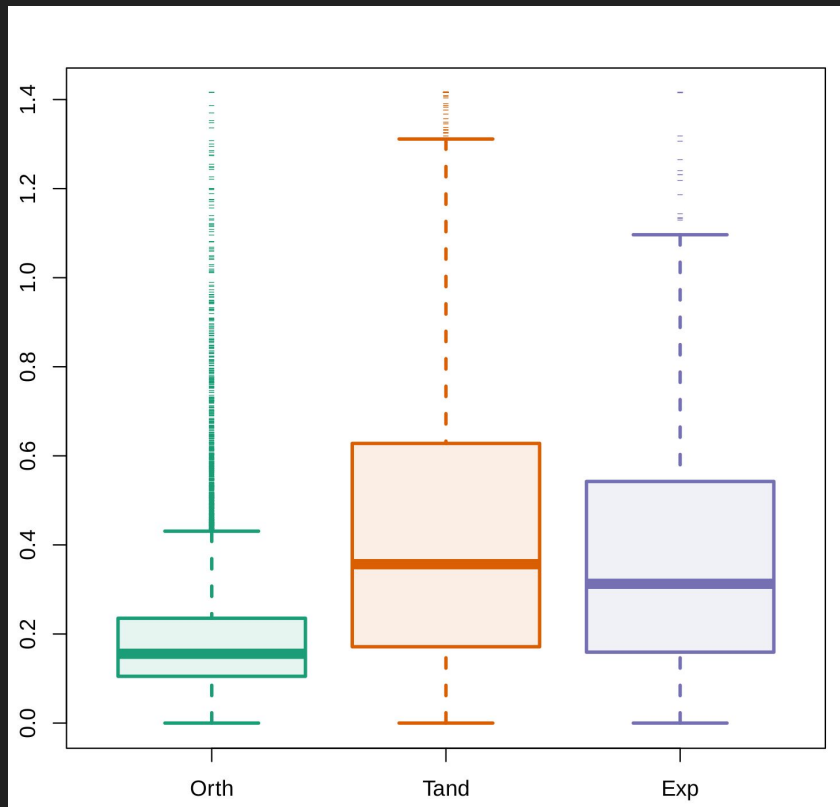


Methods: Expression Diversity

- RNA-Seq for five tissues:
 - Two species only:
 - *A. thaliana*
 - *C. hirsuta*
 - Seedling
 - Cotyledon
 - Developing Leaf
 - Flower stage 9
 - Flower stage 16
- Expression vector space on RPKM values
 - Relative for euclidean distance measures
 - Median of in-group pairwise euclidean distances

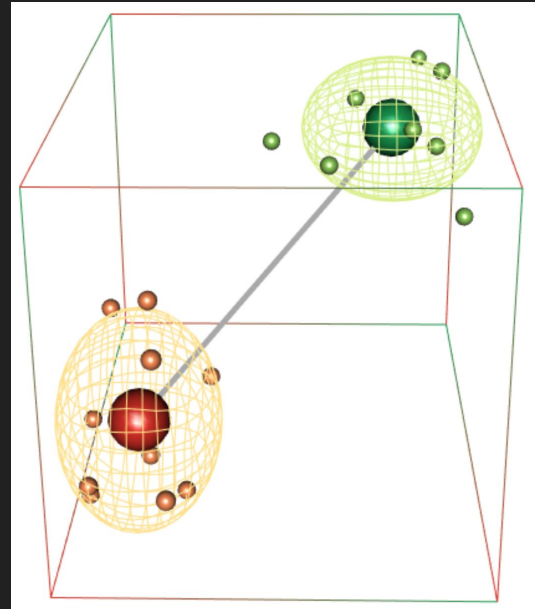
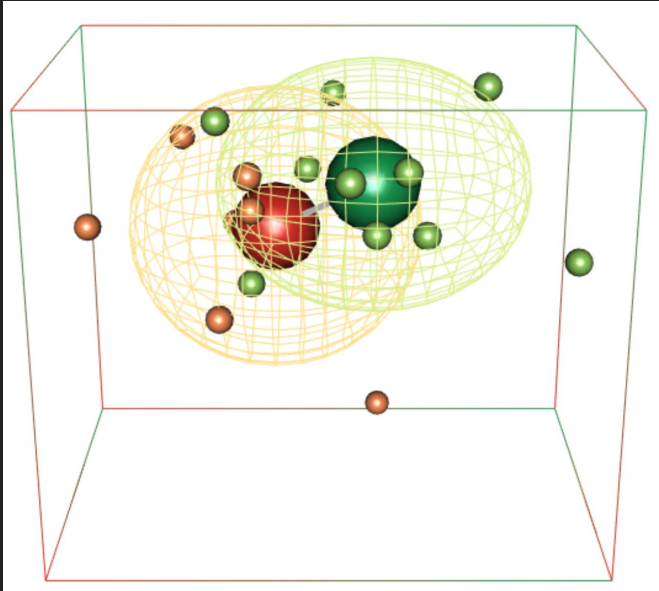


Results: median Expression Distances



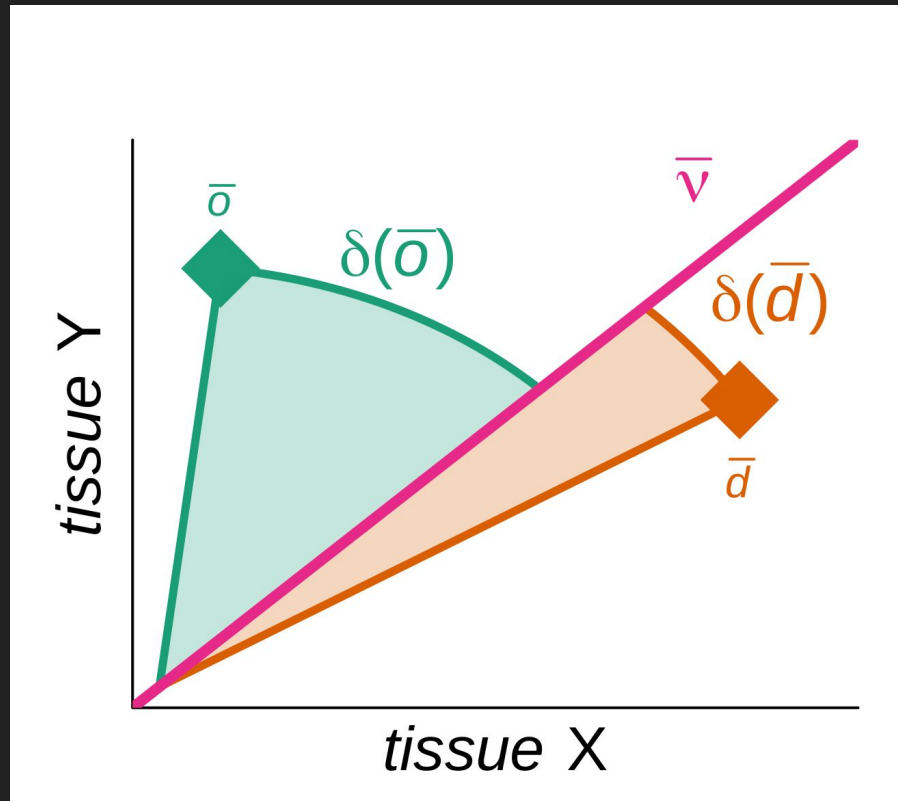
Methods: Significant diversification of expression

- Group-wise mean expression: Orthologs vs Duplicated
- Standard deviation space
- Separation or overlap?

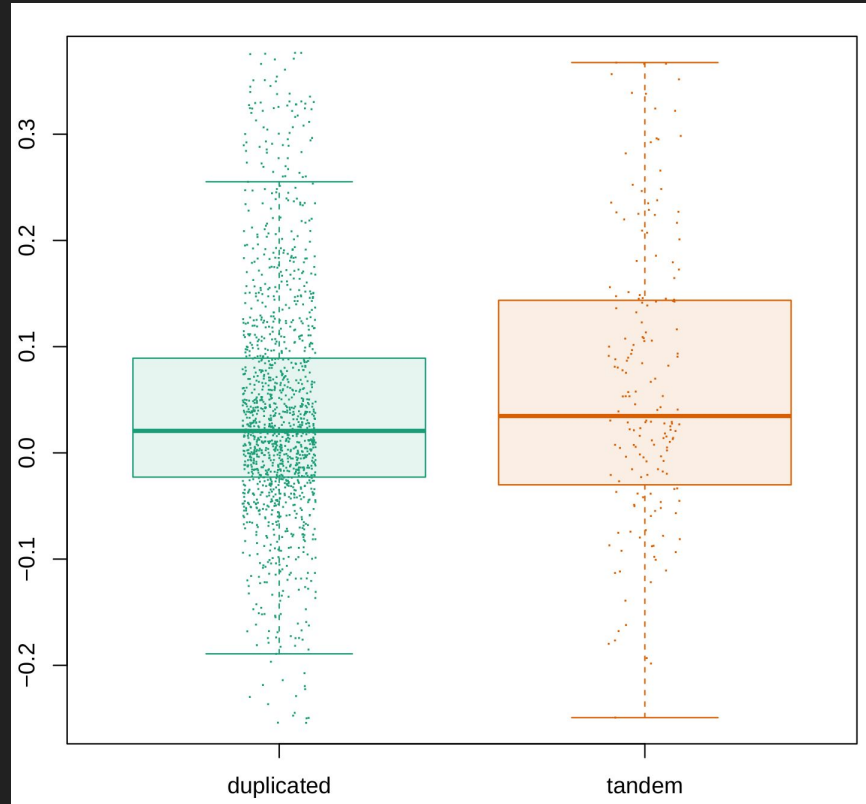


Methods: Change in tissue versatility

- Groups with significant separation of expression
- Gain or loss in number of tissues?

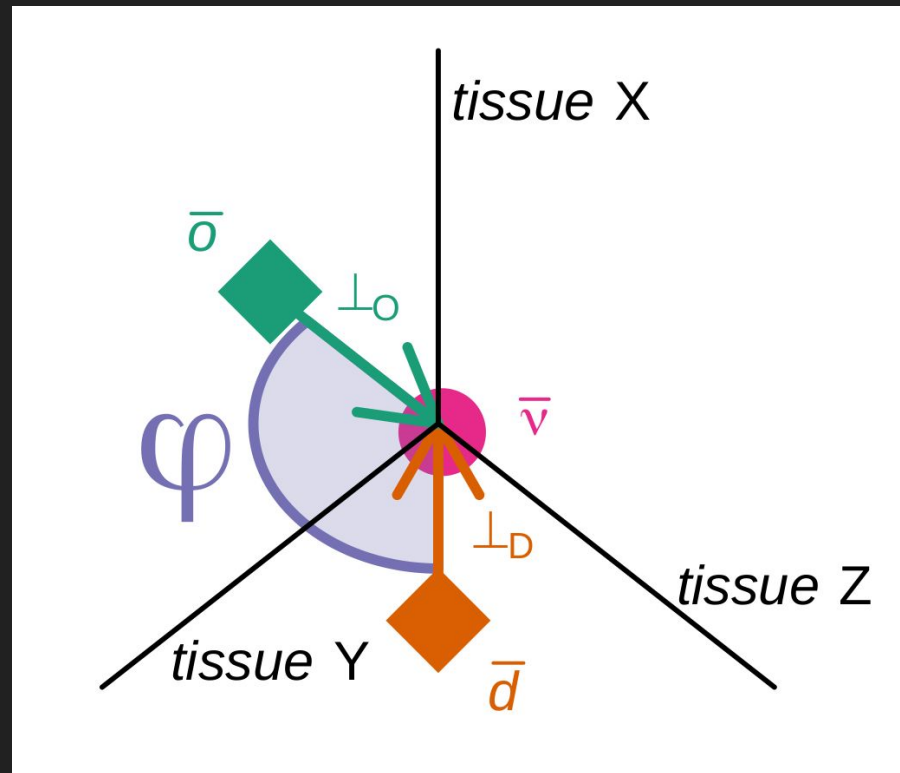


Results: Change in tissue versatility

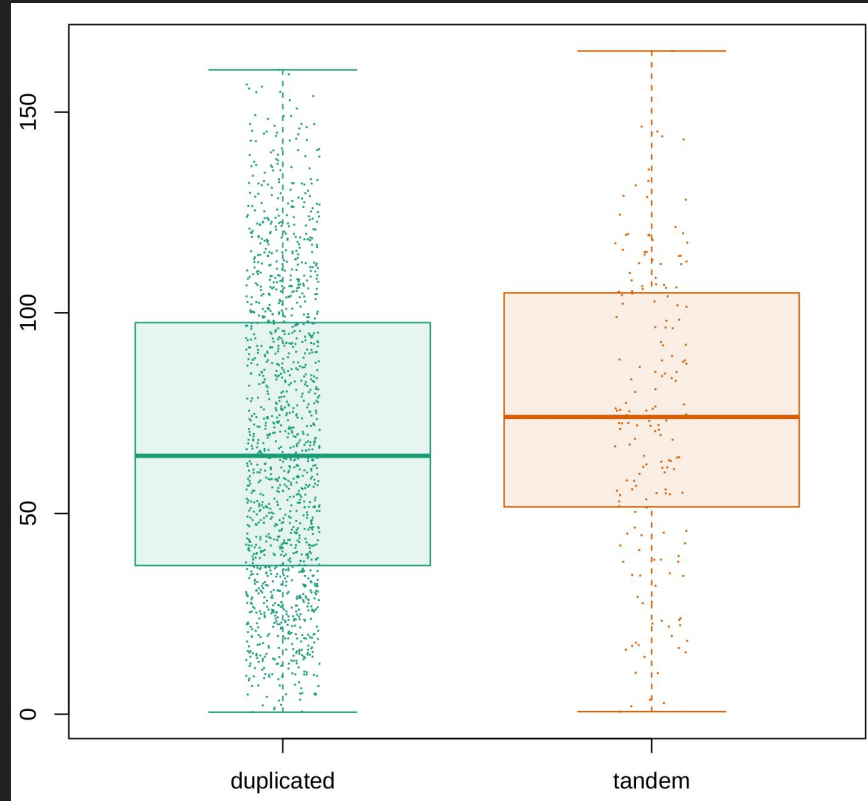


Methods: Change in tissue specificity

- Groups with significant separation of expression
- Switch to other tissues?



Results: Change in tissue specificity



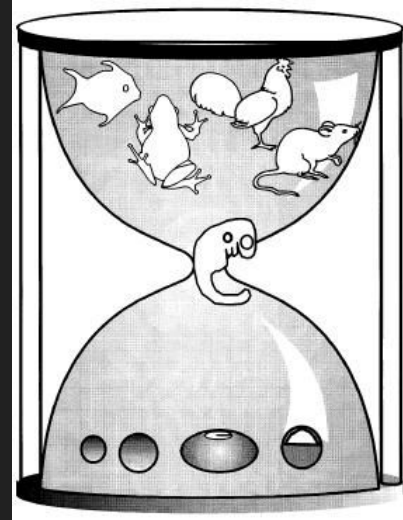
Conclusion

After gene duplication

- **Tandem** Duplicates tend to **expressional** diversification
- **Trans** Duplicates tend to **molecular** diversification

Side story: The developmental hourglass

- Expression diversity is highest in early and late developmental stages of embryos
- Conserved in animals and plants
- What kind of genes contribute to this effect?
- Plants develop through their whole life-span

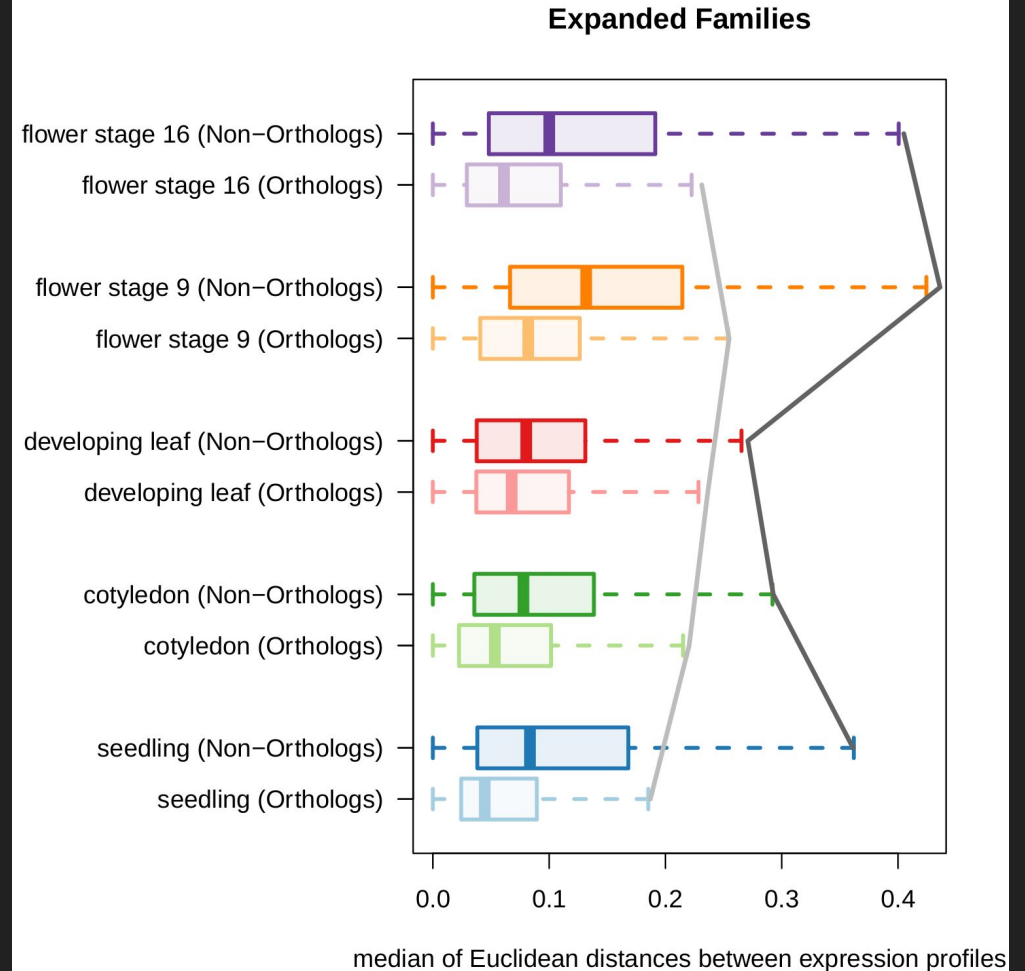


Results: Tandems - Tissue specific expression diversity



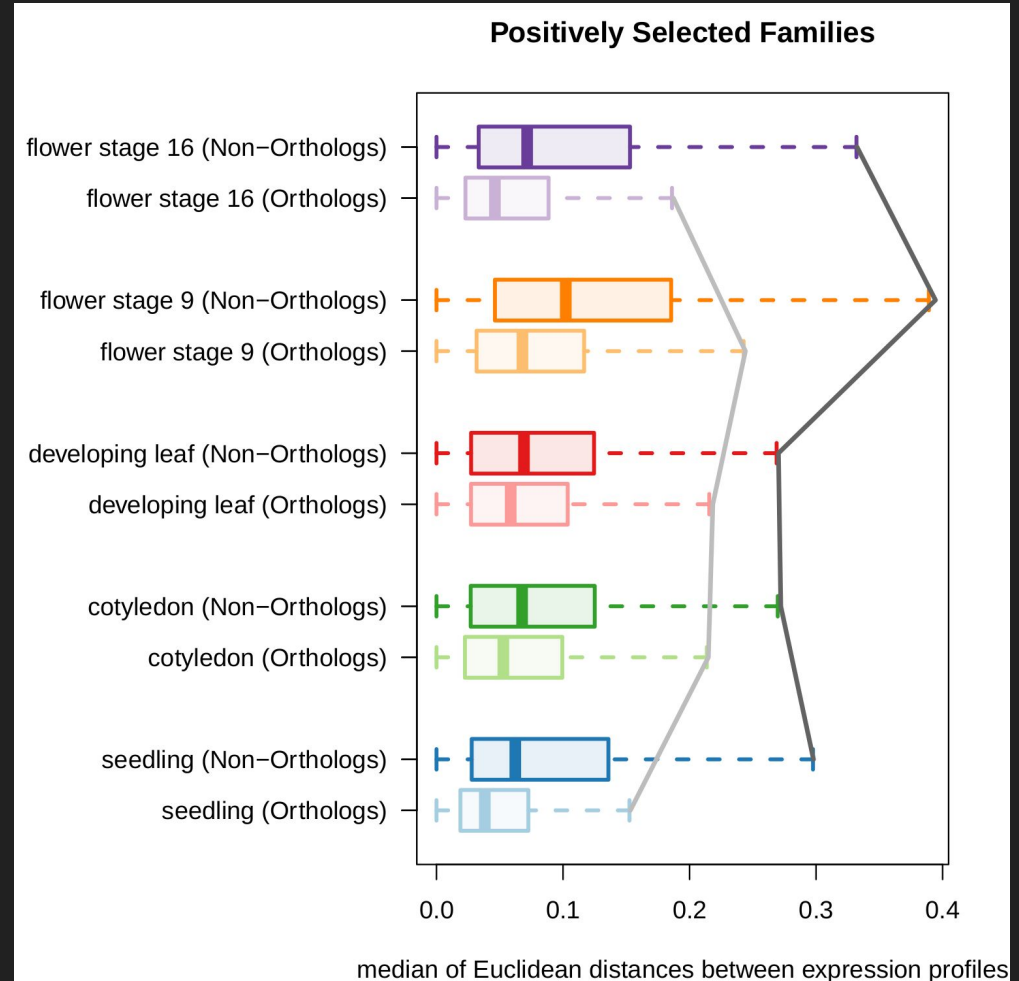
Results:

trans duplicates - Tissue specific expression diversity



Results:

Pos. selected genes - Tissue specific expression diversity



Side Story Conclusion

- Indication for a developmental hourglass throughout the plant's life-span?
- Tandem duplicates, trans duplicates, and positively selected genes appear to underpin this expression diversity.
- Gene duplication and diversification of expression leads to differential morphogenesis?
- Evidence might be too weak.

Thank you very much



Yokozuna Harumafuji gives you all a big big hug