

## **Supplementary Information:**

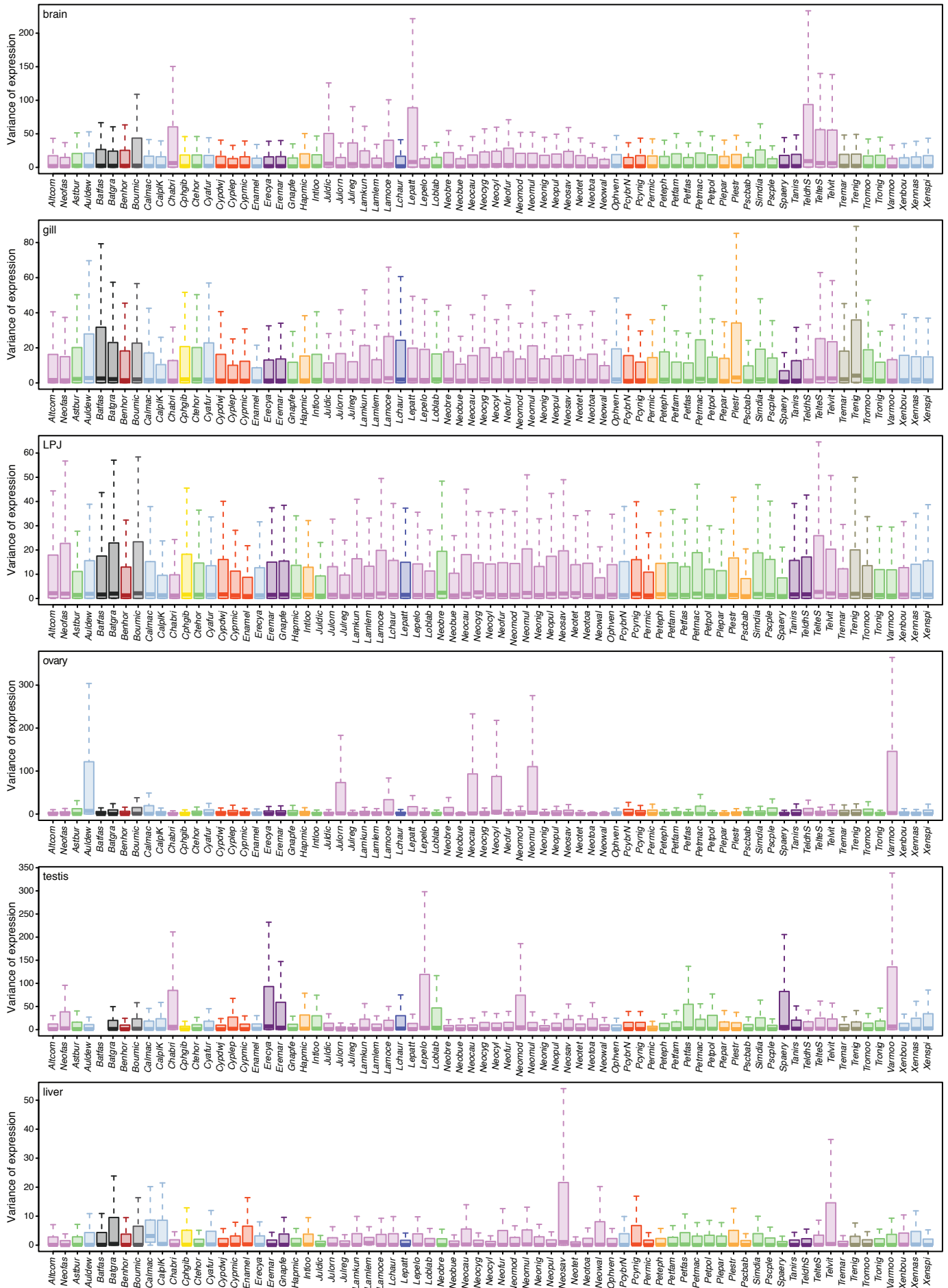
### **Gene expression dynamics during rapid organismal diversification in African cichlid fishes**

Athimed El Taher<sup>1\*</sup>, Astrid Böhne<sup>1,2</sup>, Nicolas Boileau<sup>1</sup>, Fabrizia Ronco<sup>1</sup>, Adrian Indermaur<sup>1</sup>, Lukas Widmer<sup>1</sup> and Walter Salzburger<sup>1\*</sup>

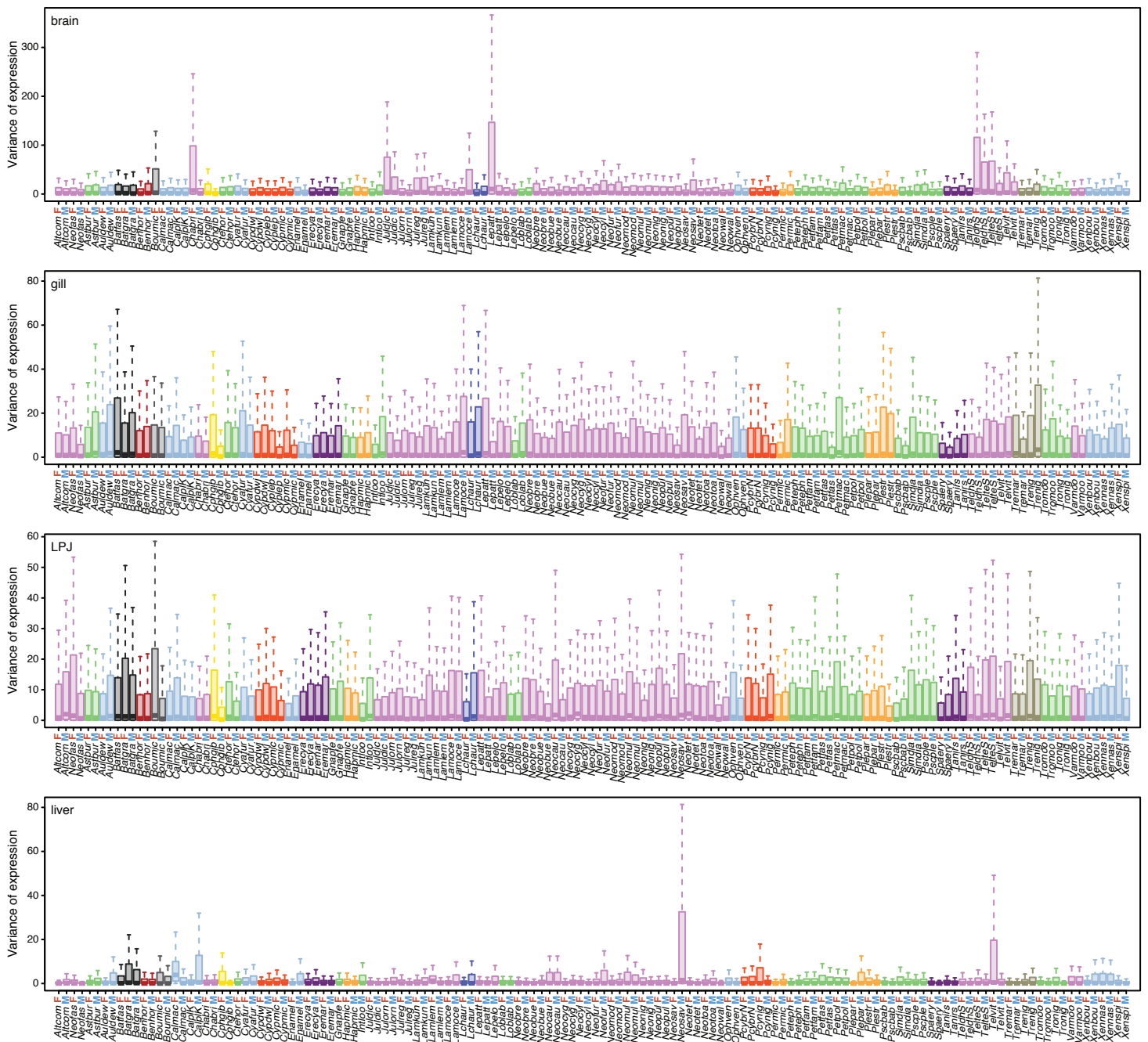
<sup>1</sup>Zoological Institute, Department of Environmental Sciences, University of Basel, Basel, Switzerland

<sup>2</sup>Present address: Center for Molecular Biodiversity Research, Zoological Research Museum Alexander Koenig, Bonn, Germany

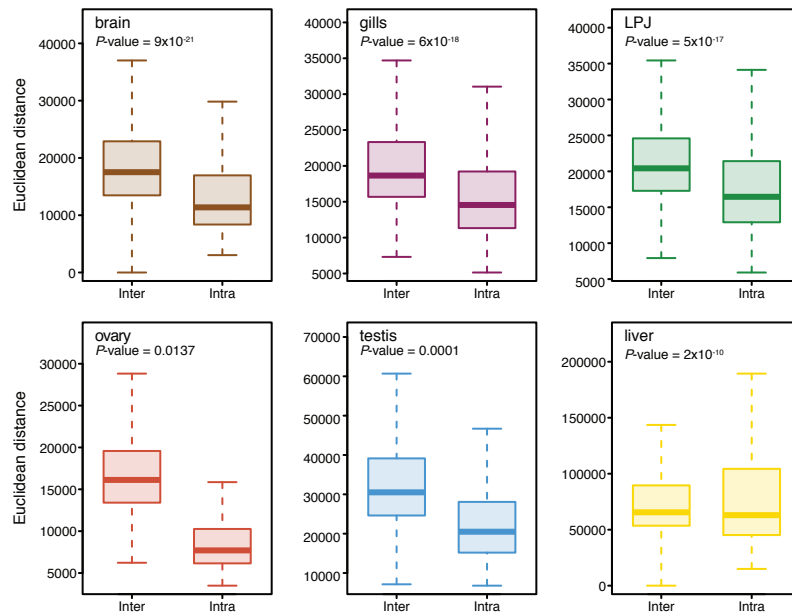
\*e-mail: athimed.eltaher@unibas.ch, walter.salzburger@unibas.ch



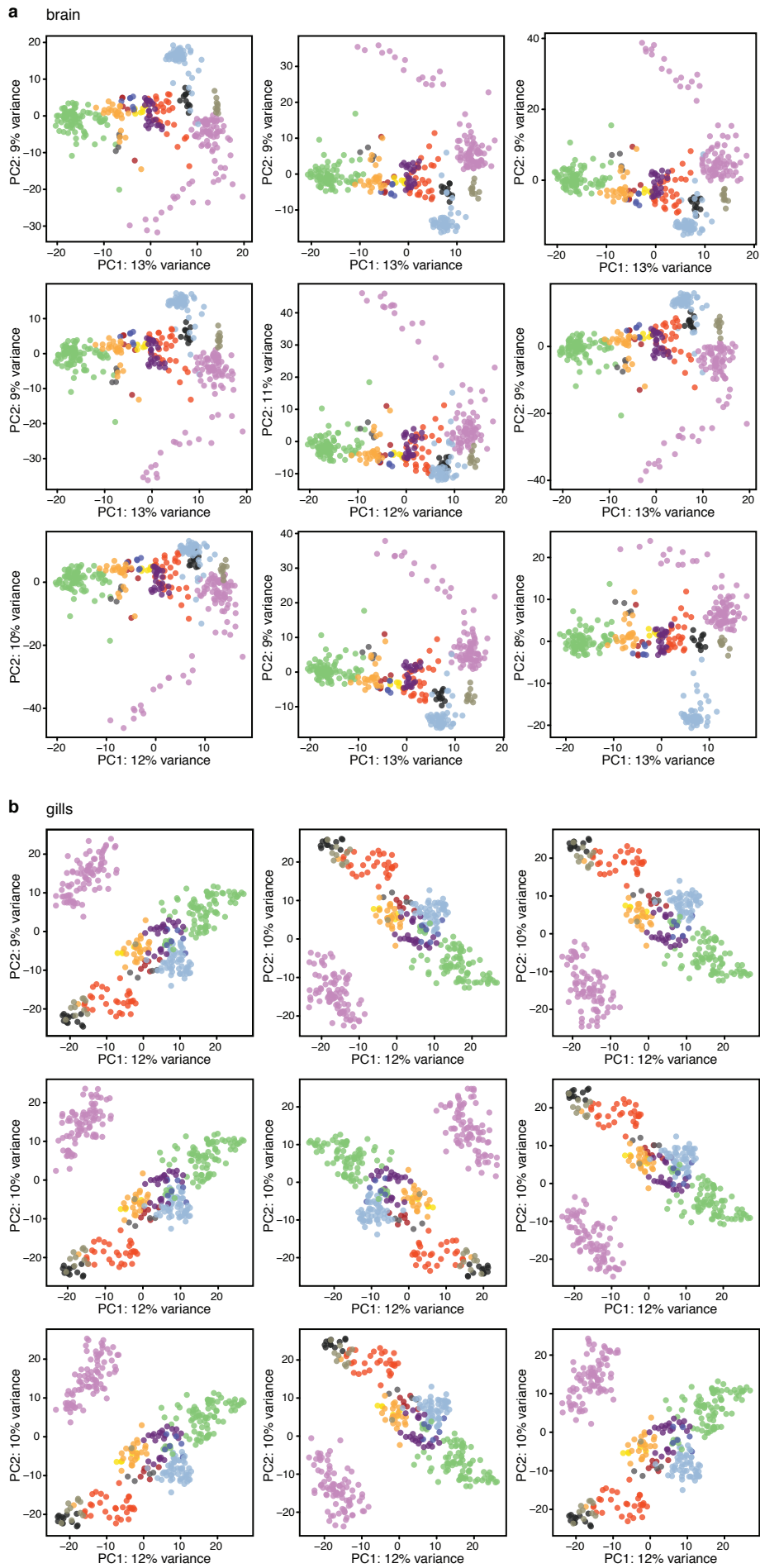
**Supplementary Figure 1 | Within-species variance of gene expression.** Boxplots summarise median variance per gene across species replicates. Outliers are not shown. Boxplot centre lines represent the median, box limits the upper and lower quartiles, and whiskers the 1.5x interquartile range. Boxplots are coloured according to tribes as defined in Fig. 1a.

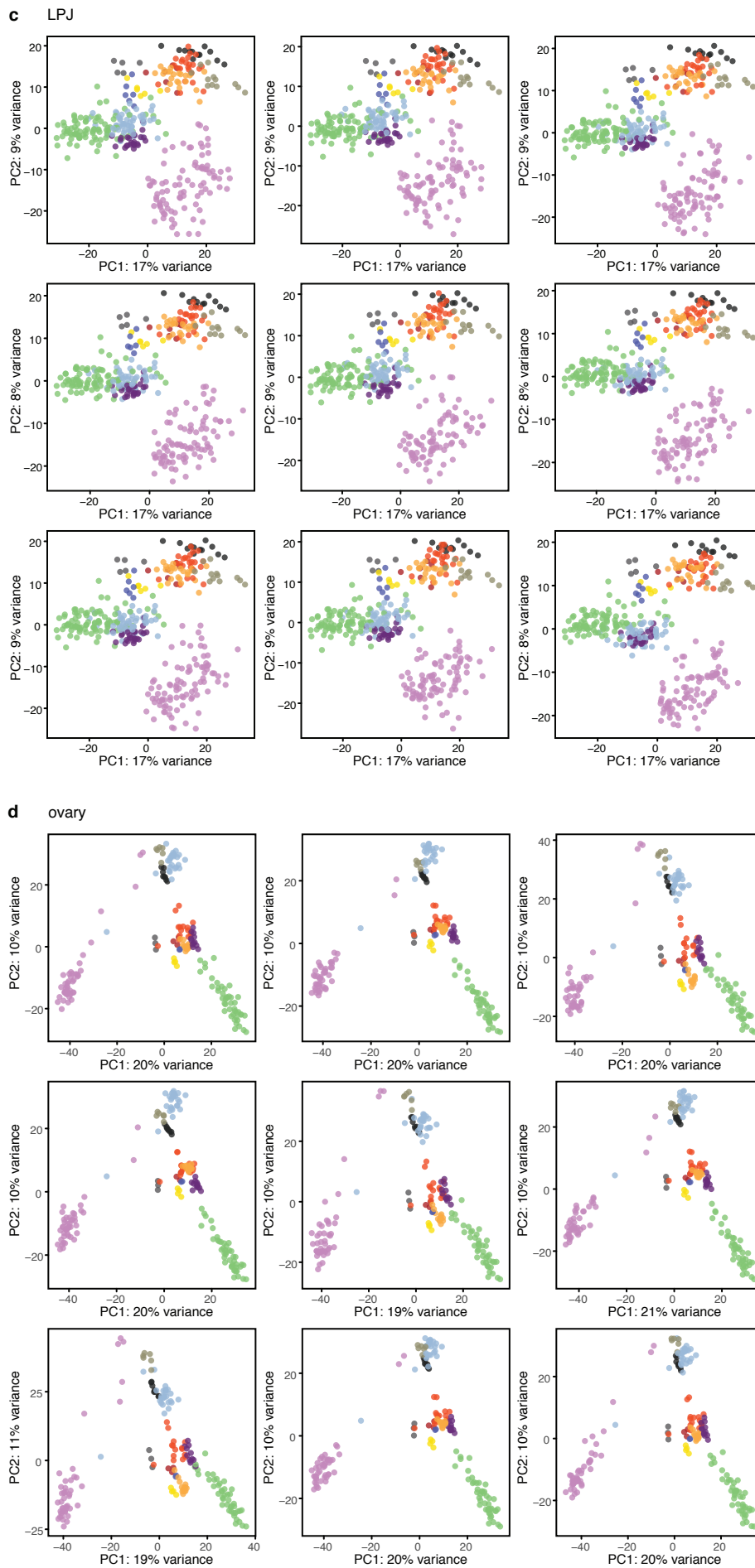


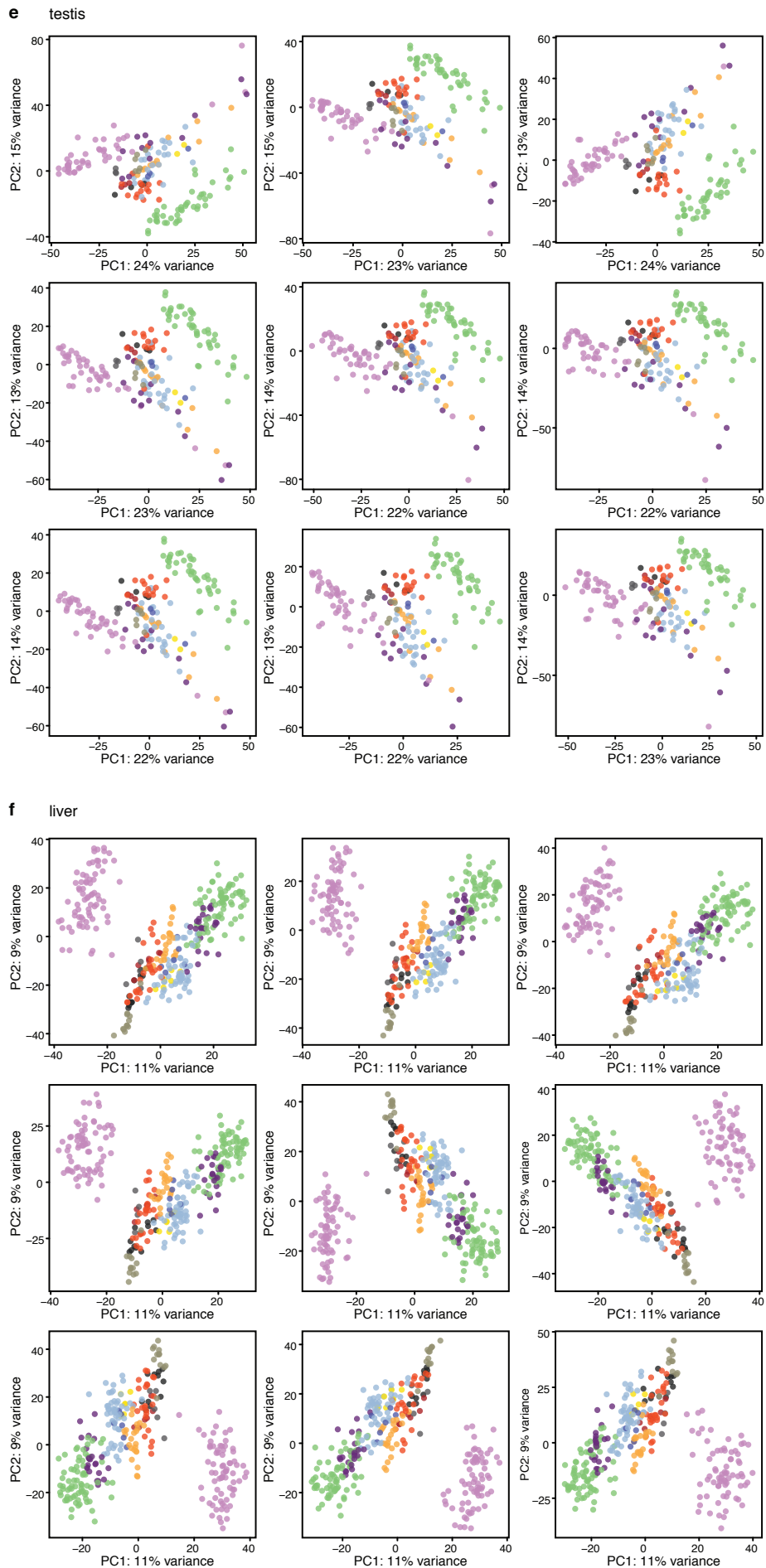
**Supplementary Figure 2 | Within-species and sex variance of gene expression in somatic organs.** Boxplots show the variance of expression observed within species for replicates of the same sex per gene. Males are indicated with “M” and females with “F”. Outliers are not shown. Boxplot centre lines represent the median, box limits the upper and lower quartiles, and whiskers the 1.5x interquartile range. Boxplots are coloured according to tribes as defined in Fig. 1a.



**Supplementary Figure 3 | Inter vs. intra-species expression variation.** Boxplots show Euclidean distances of expression calculated between (inter) and within (intra) species. The  $P$ -value of the two-sided t-test between the two groups is represented on the plots. Boxplot centre lines represent the median, box limits the upper and lower quartiles, and whiskers the 1.5x interquartile range.

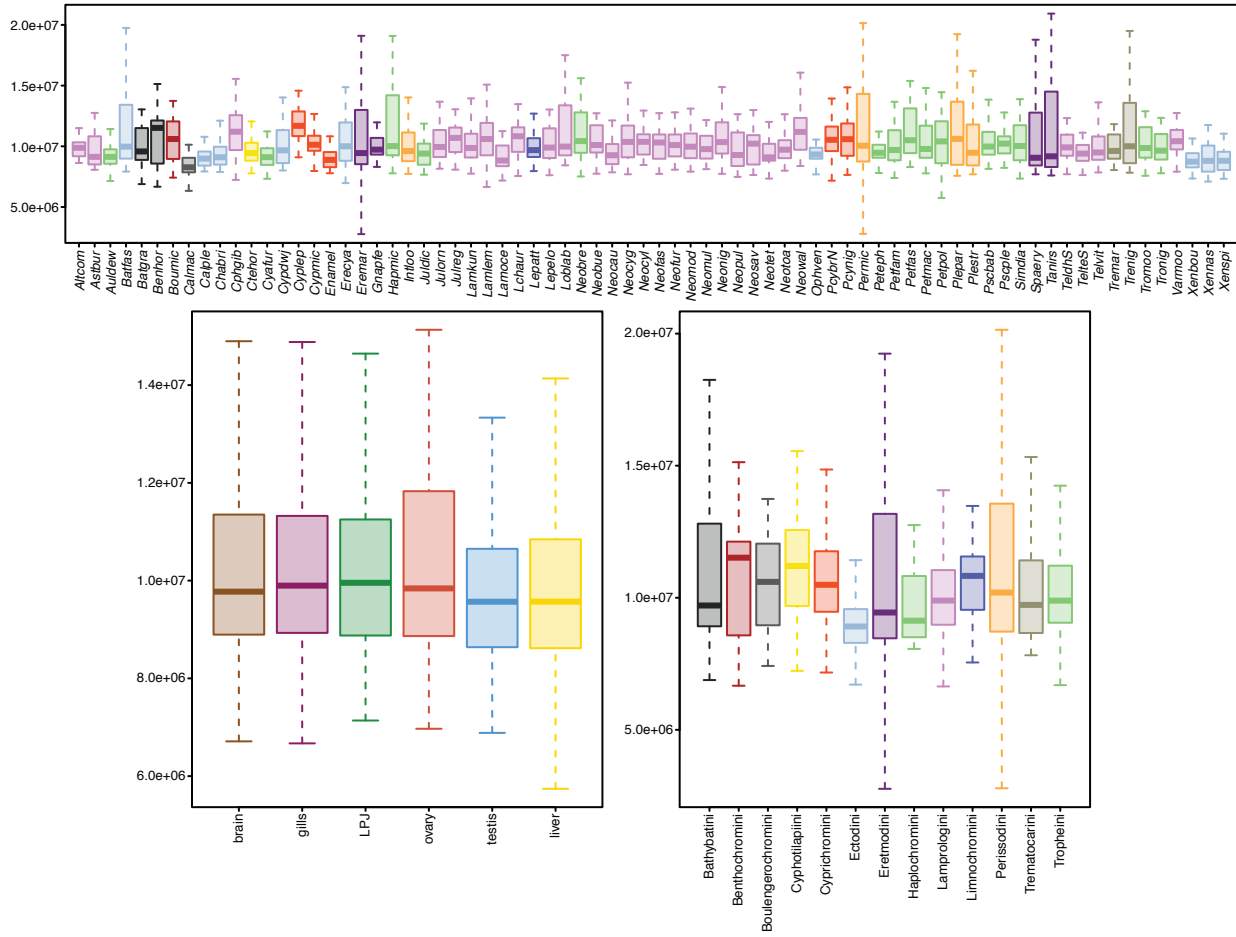




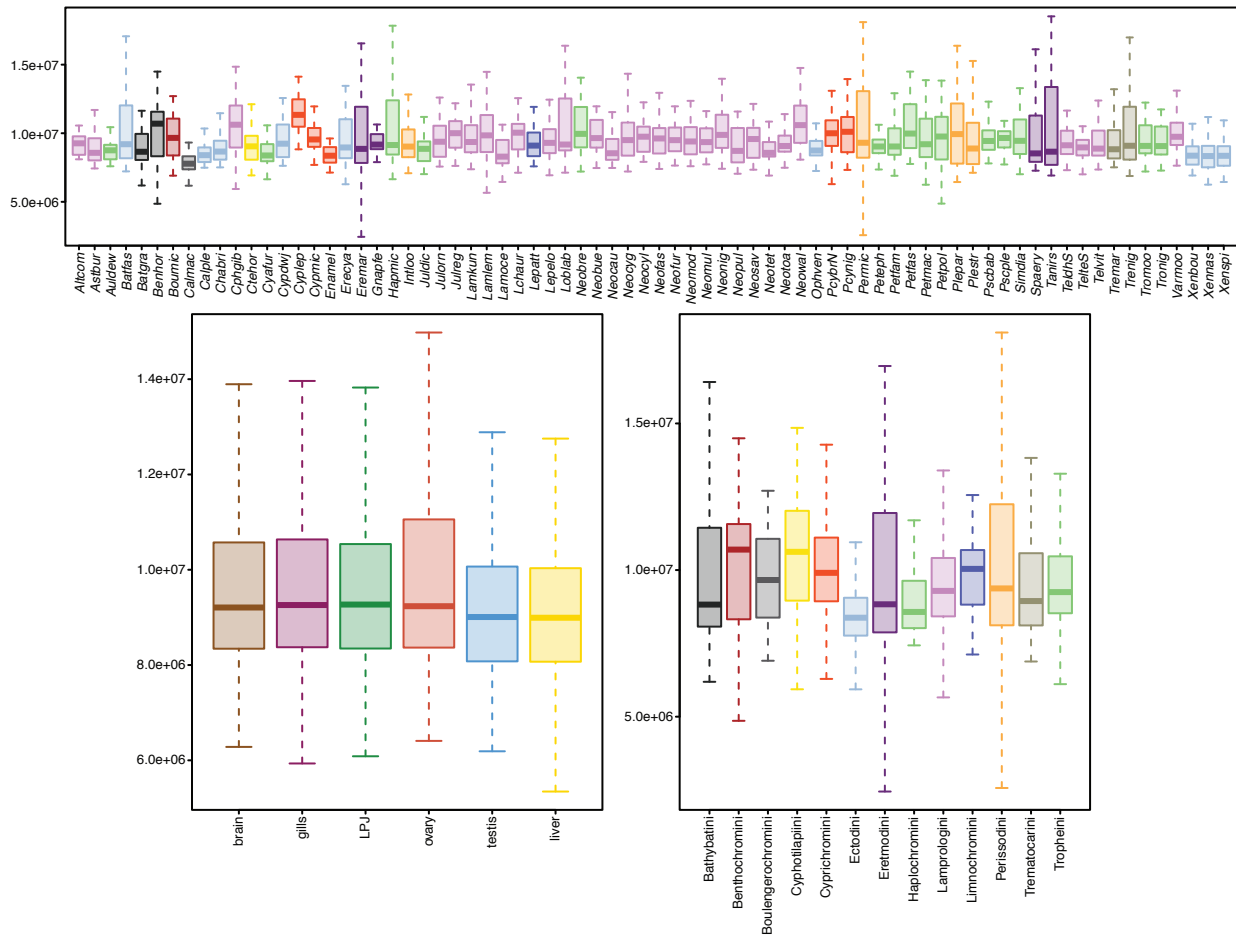


**Supplementary Figure 4 | Principal component analyses of overall gene expression levels within subsets of species.** For this analysis, the most species-rich tribe (Lamprologini) was down-sampled to a random subset of 15 species (instead of 29 species) to match the number of species of the next largest phylogenetic group in the dataset, the Tropheini. Random subsampling has been performed nine times for each organ (a-f).

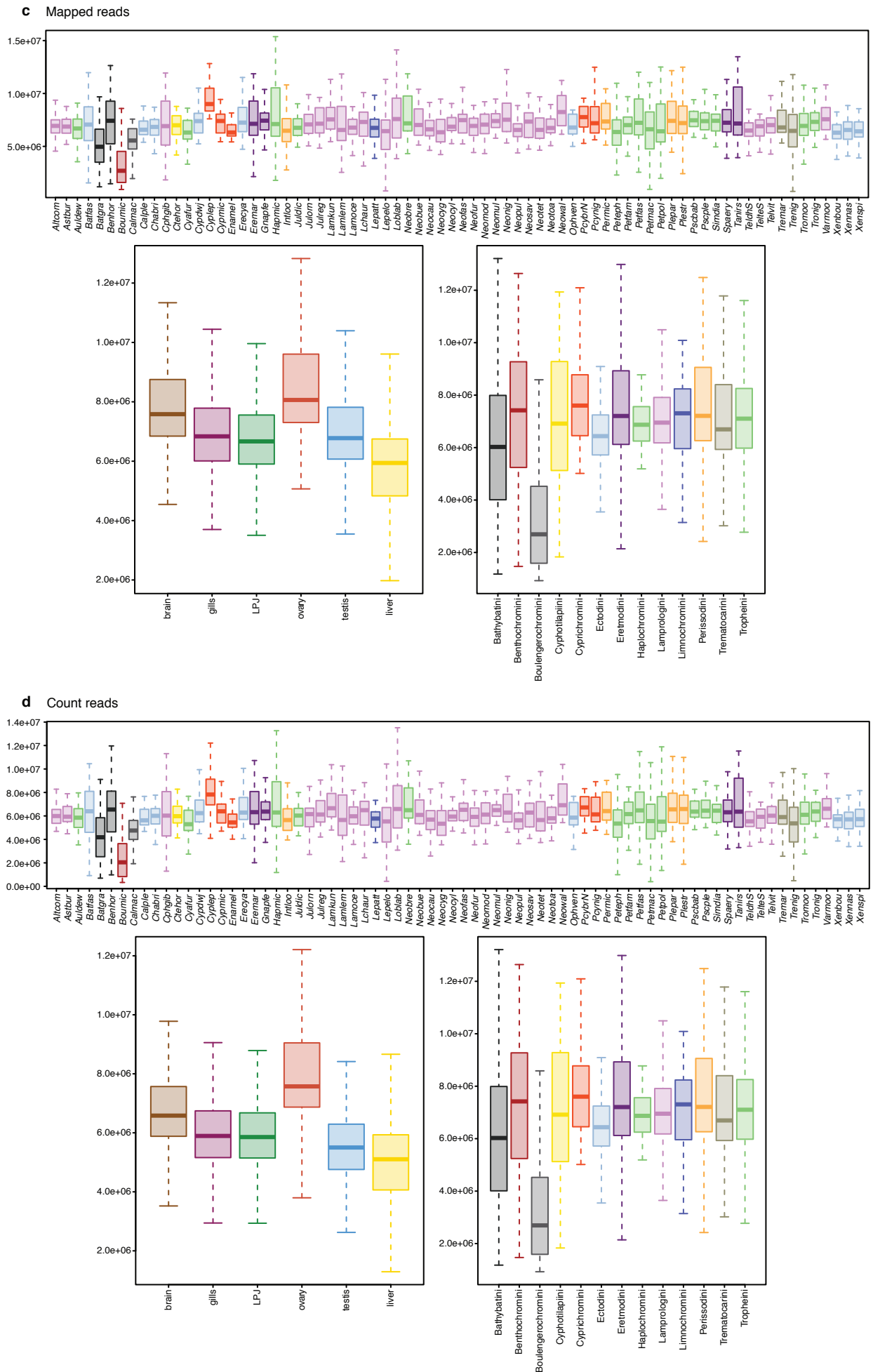
**a** Raw reads

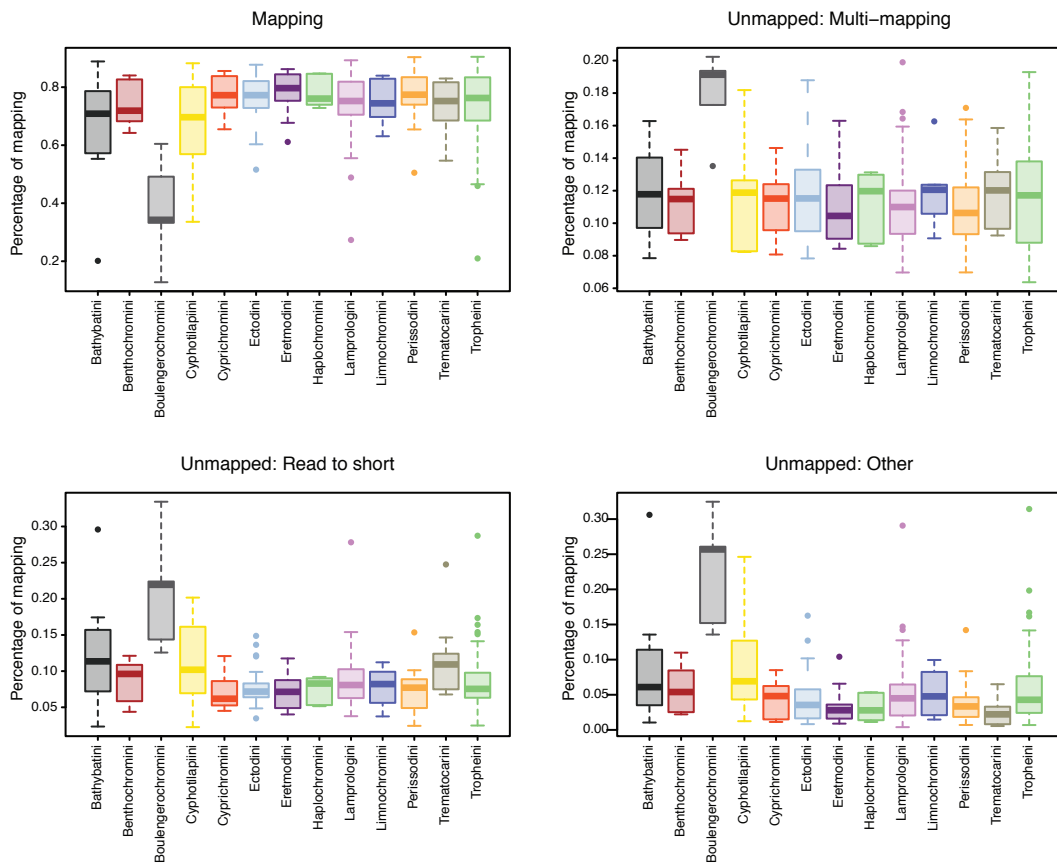


**b** Trimmed reads









**Supplementary Figure 6 | Mapping statistics.** Fractions of reads considered as mapped (uniquely) and unmapped (due to multi-mapping, too short read length or other reasons) per tribe. Boxplots are coloured according to tribe, as defined in Fig. 1a. Boxplot centre lines represent the median, box limits the upper and lower quartiles, and whiskers the 1.5x interquartile range.