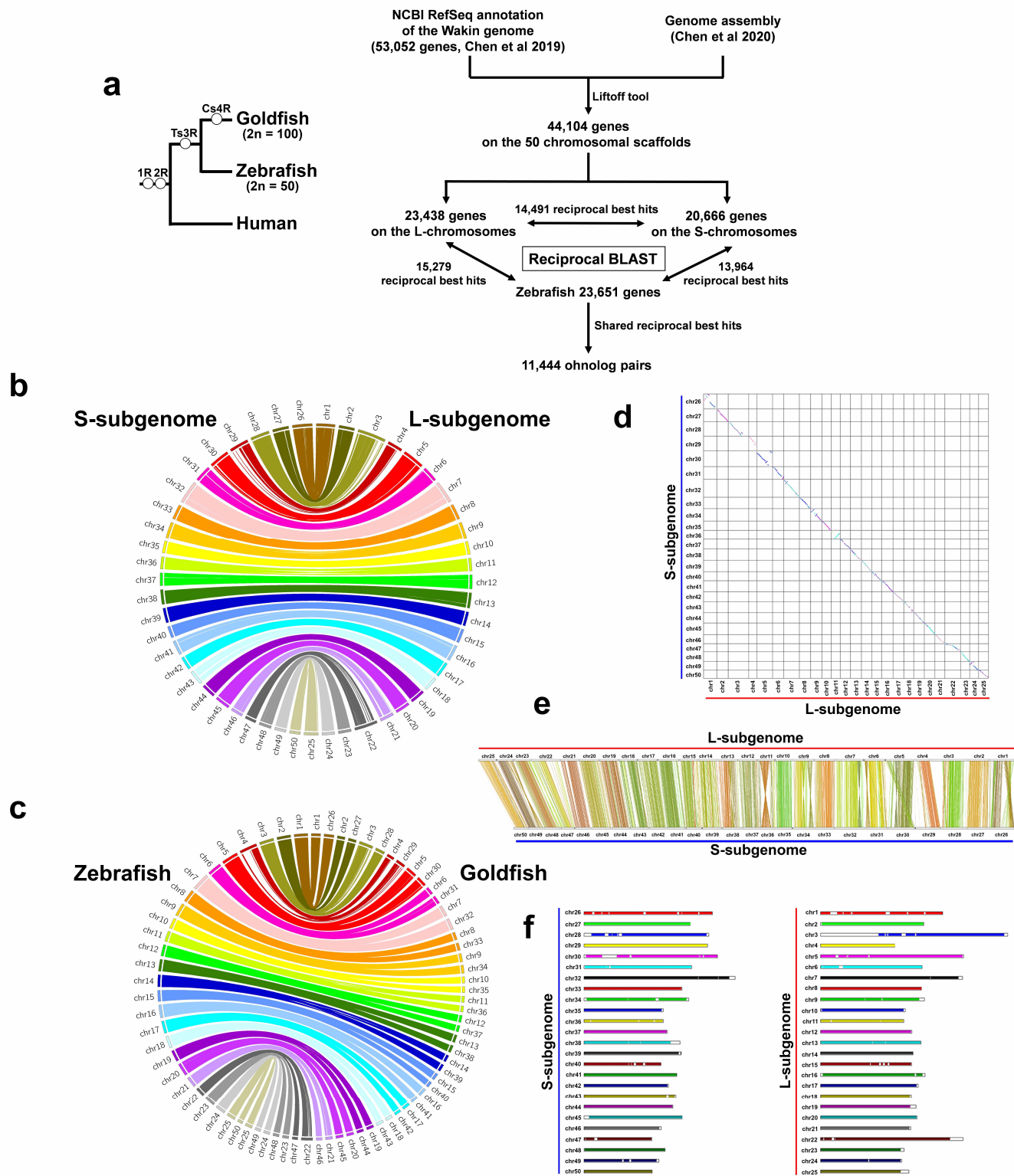


Single-cell transcriptomics of the goldfish retina reveals genetic divergence in the asymmetrically evolved subgenomes after allotetraploidization

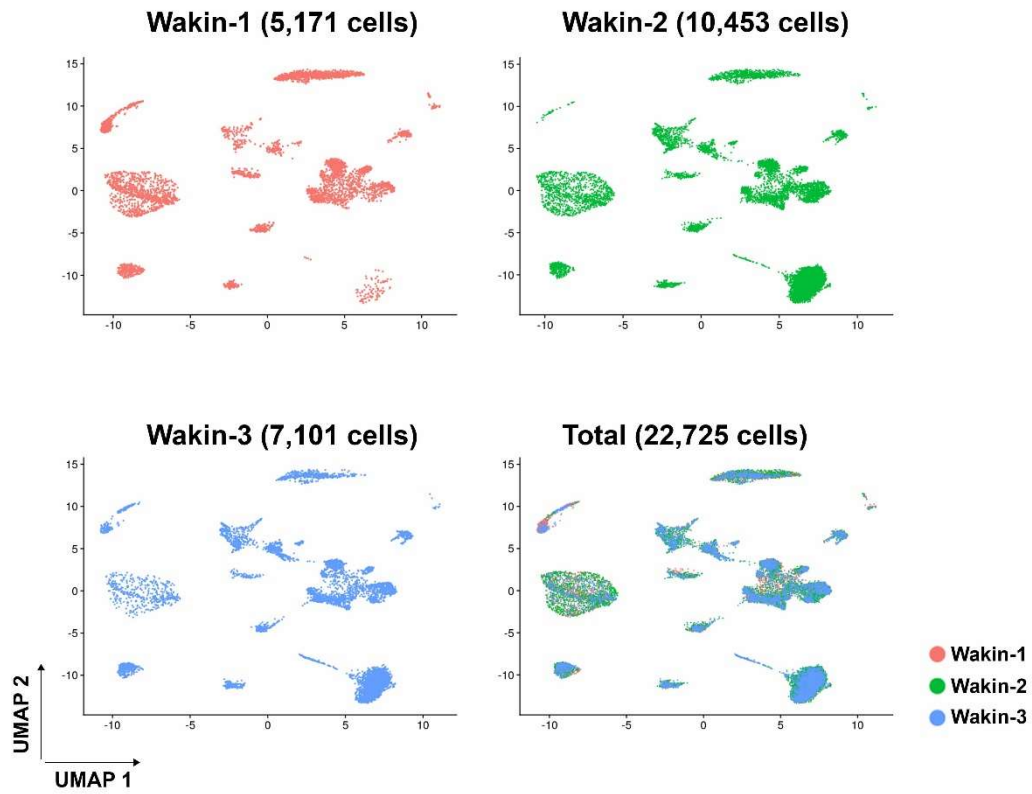
Tetsuo Kon, Kentaro Fukuta, Zelin Chen, Koto Kon-Nanjo, Kota Suzuki, Masakazu Ishikawa, Hikari Tanaka, Shawn M. Burgess, Hideki Noguchi, Atsushi Toyoda, Yoshihiro Omori

Supplementary Figures



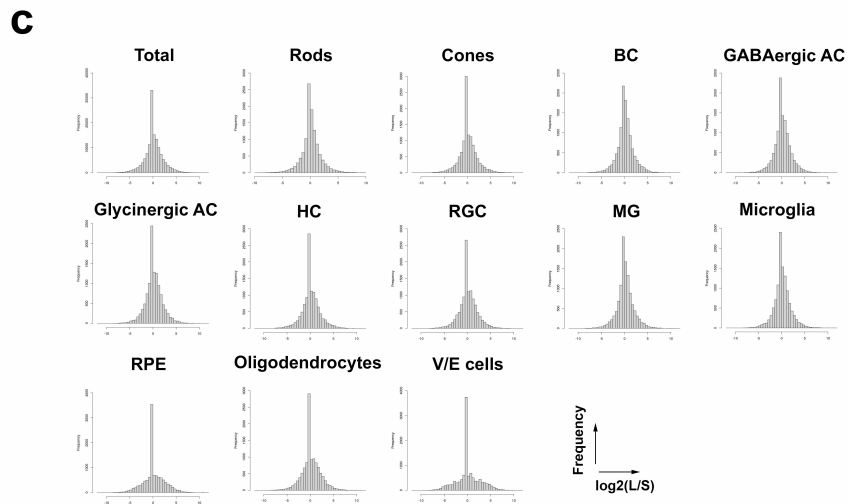
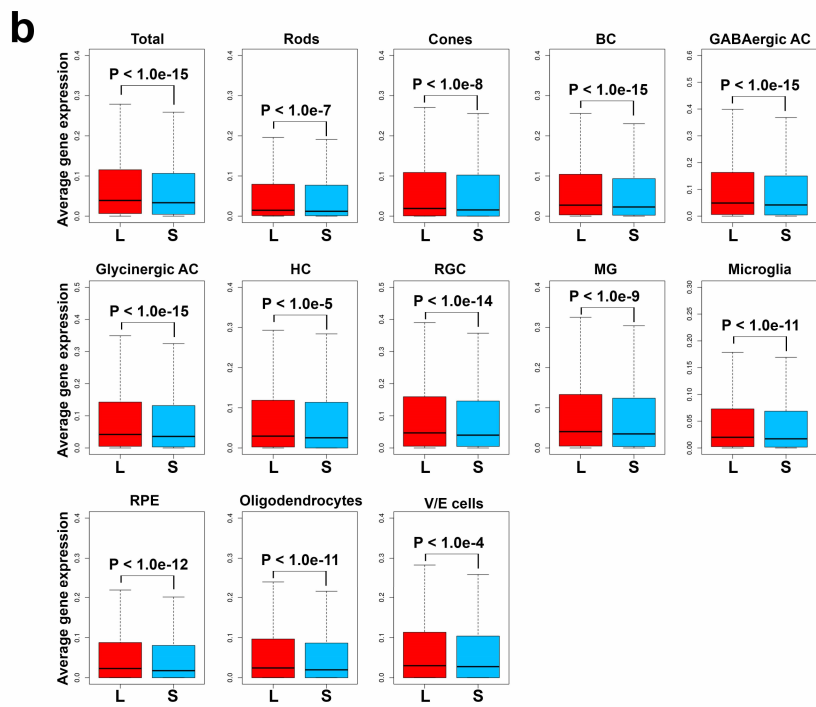
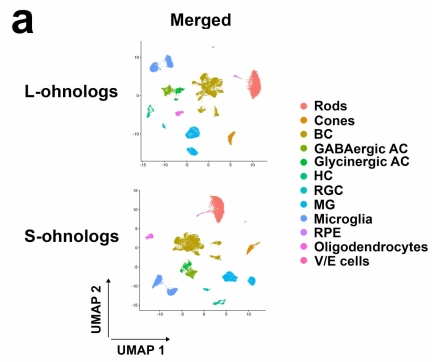
Supplementary Figure 1. Identification of the 11,444 ohnolog pairs in goldfish and zebrafish chromosomes

(a) Left panel; Cladogram showing the evolutionary relationships between goldfish, zebrafish, and humans. The white circles indicate the two rounds of whole genome duplication in the vertebrate ancestor (1R and 2R), Ts3R, and Cs4R. Right panel; Pipeline for identification of 11,444 ohnolog pairs in goldfish genome. (b) Circos plot depicting the 11,444 ohnolog pairs. The L-chromosomes are located on the right and the S-chromosomes are located on the left. (c) Circos plot depicting the ortholog pairs between the 11,444 zebrafish genes and 11,444 goldfish ohnolog pairs. The 50 goldfish chromosomes are located on the right. The 25 zebrafish chromosomes are located on the left. (d) Oxford dot plot of the 11,444 ohnolog pairs. (e) Dual synteny plot of the 11,444 ohnolog pairs. (f) Bar plot showing the chromosomal position of each ohnolog.



Supplementary Figure 2. UMAP plots of scRNA-seq data for three Wakin goldfish

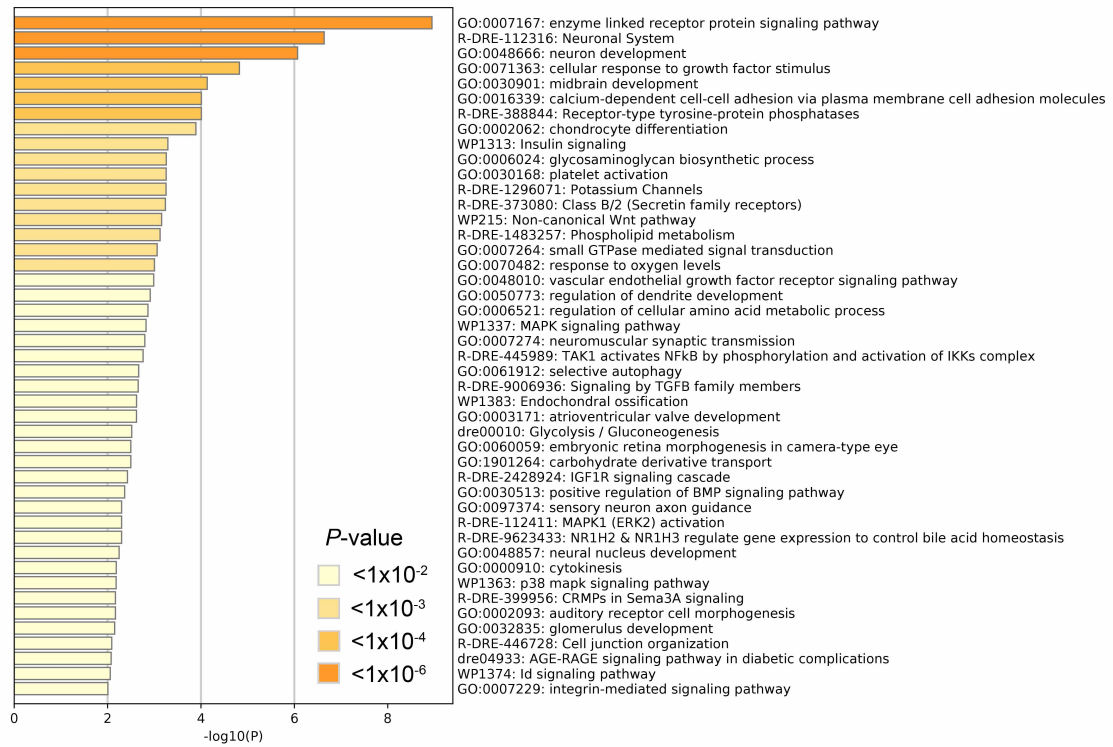
The scRNA-seq data from three Wakin individuals (red, green, and blue) are shown on the UMAP plot.



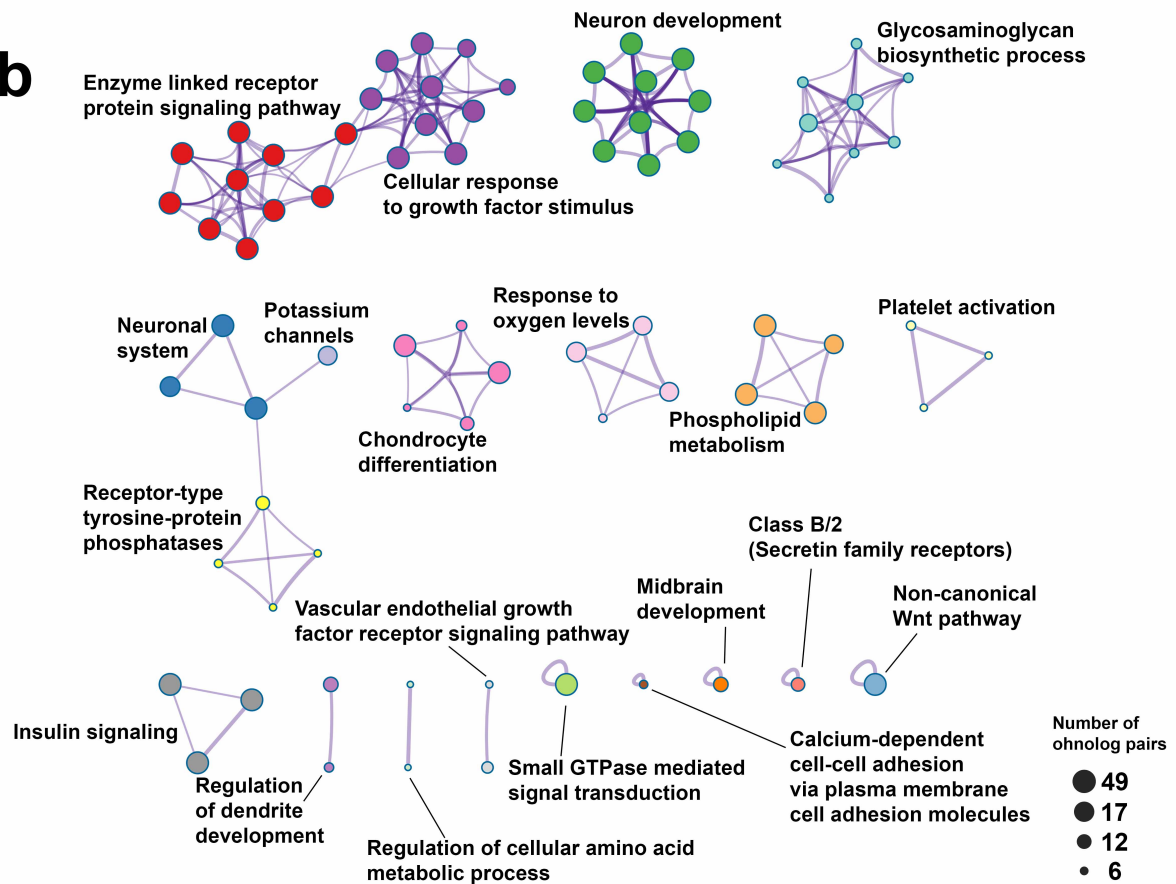
Supplementary Figure 4. Gene expression level of L- and S-ohnologs in the 12 retinal cell types

(a) UMAP plots based on the expression data of the 11,444 L-ohnologs (upper panel) and the 11,444 S-ohnologs (lower panel) were shown independently. The 12 cell types identified were shown in different colors. (b) The distribution of the averaged gene expression levels of L-ohnologs (red) and S-ohnologs (blue) in the 12 types of retinal cells was shown. The ends of the box are the 25% and 75% quantiles. The horizontal line in the box indicates the median. The lines extending from the top and bottom of the box represent the minimum and maximum values. (c) Histogram of log₂ transformed L/S ratios. The L/S peak is centered at zero, suggesting that the most ohnolog pairs show no biased expression. This contrasts with the right-shifted peak observed in Fig. 2c (L/S of total gene expression).

a

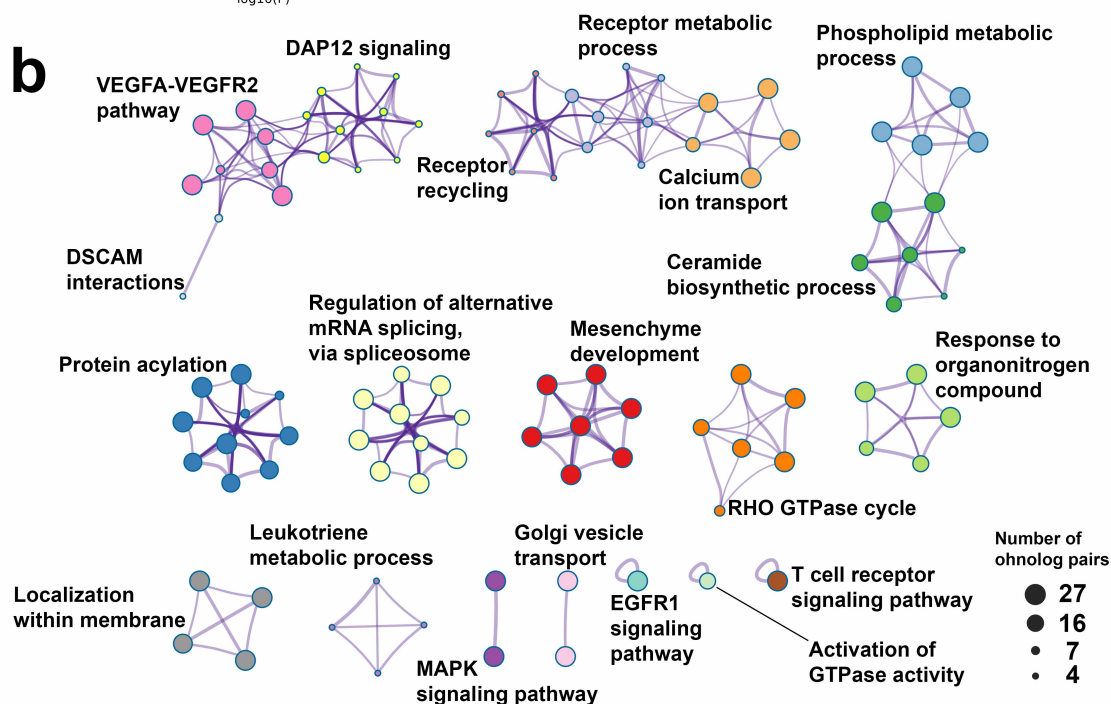
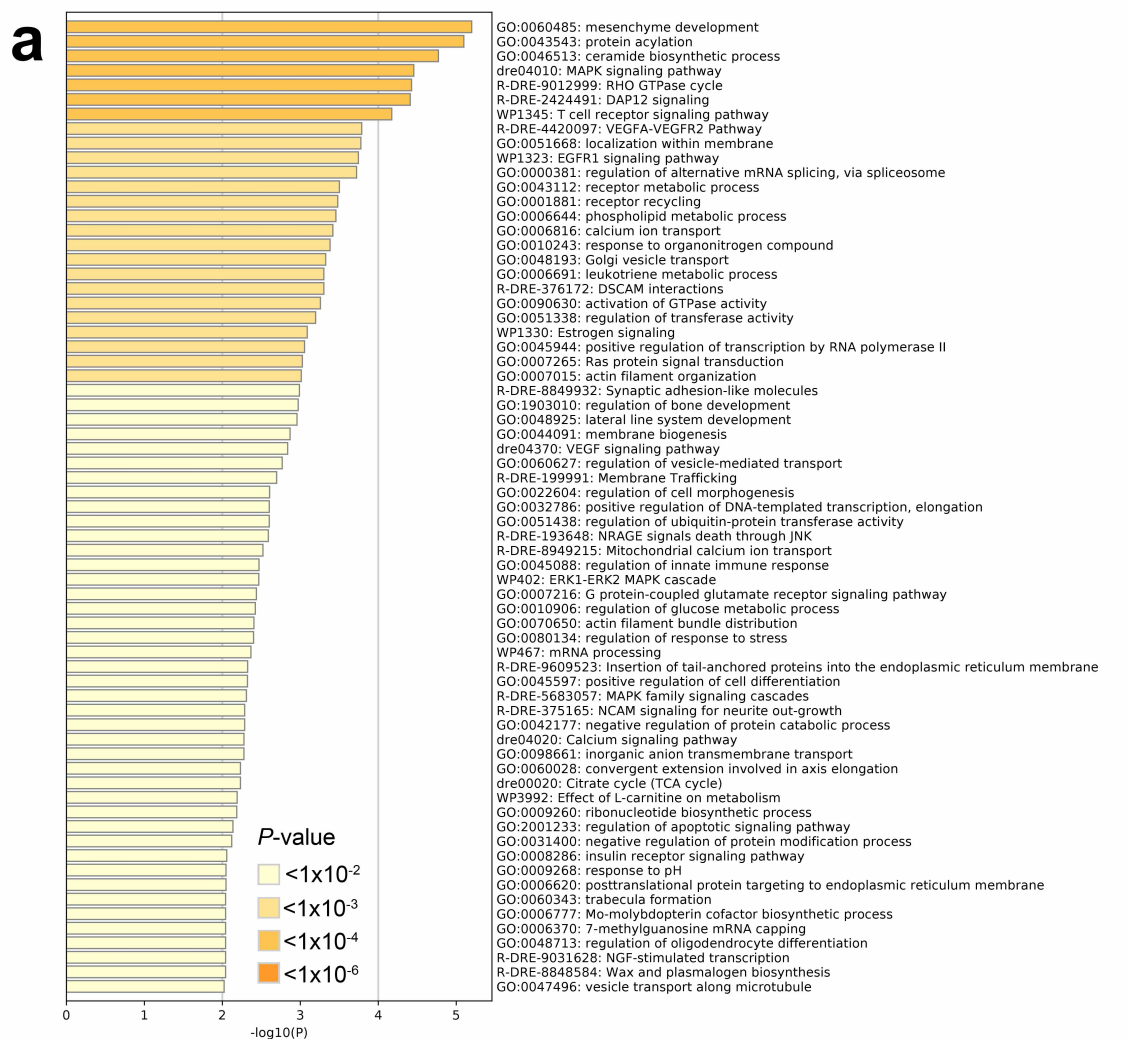


b



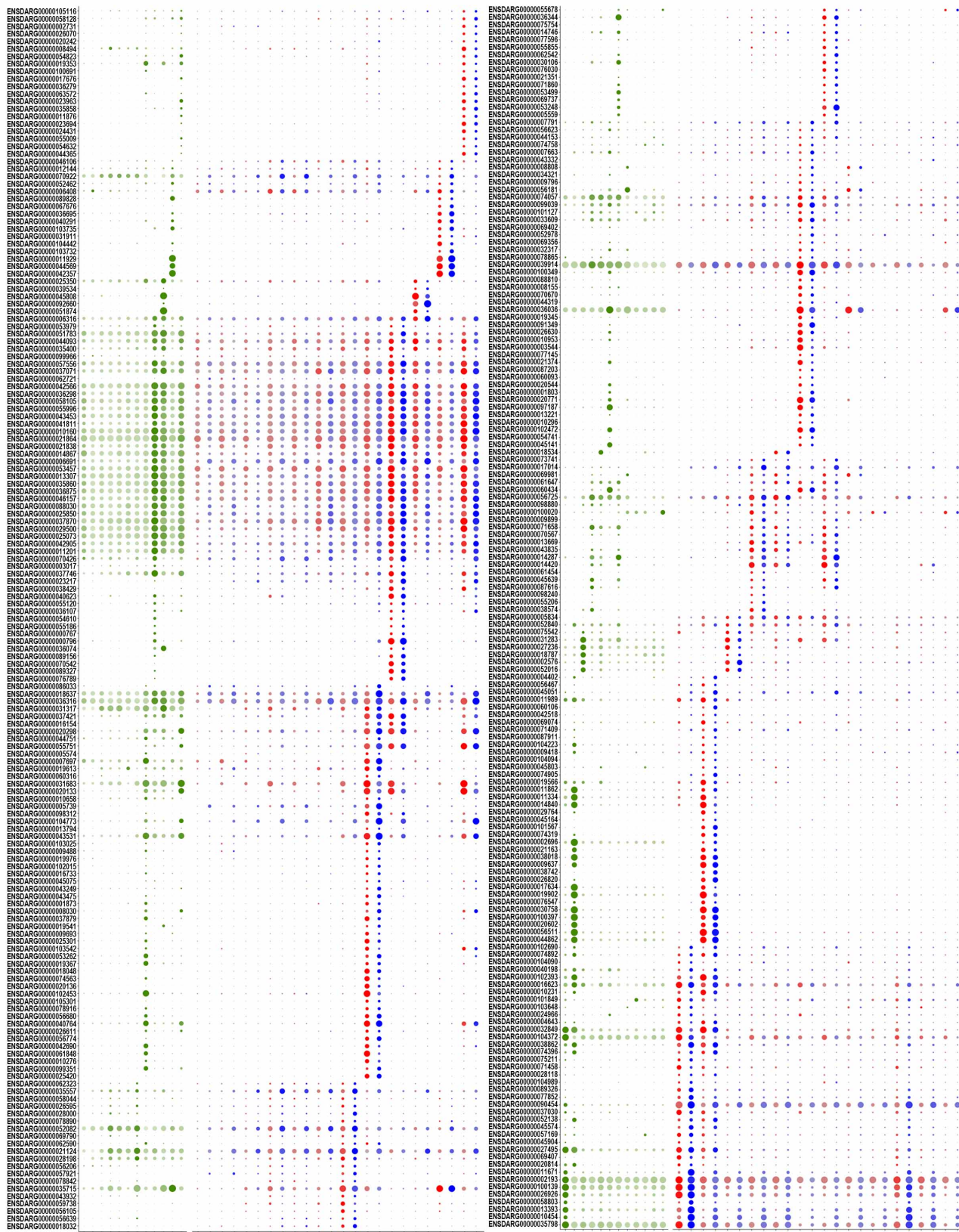
Supplementary Figure 5. Functional enrichment analysis of the 2,430 ohnolog pairs with higher gene expression in the L-ohnolog than in the S-ohnolog

(a) The x-axis represents the negative log₁₀-transformed *P*-value based on the accumulative hypergeometric distribution. The y-axis represents the enriched biological categories. Deeper color of the bar plot means smaller *P*-value. (b) Network representation of the statistically enriched categories in the functional enrichment analysis of the ohnolog pairs with higher gene expression in the L-ohnolog than in the S-ohnolog. The nodes represent the enriched categories and the edges are defined based on the similarities among their gene memberships. The name of the cluster is adopted from the name of the cluster with the smallest *P*-value among the biological categories contained in that cluster. The node size is proportional to the number of input ohnolog pairs grouped into each category.



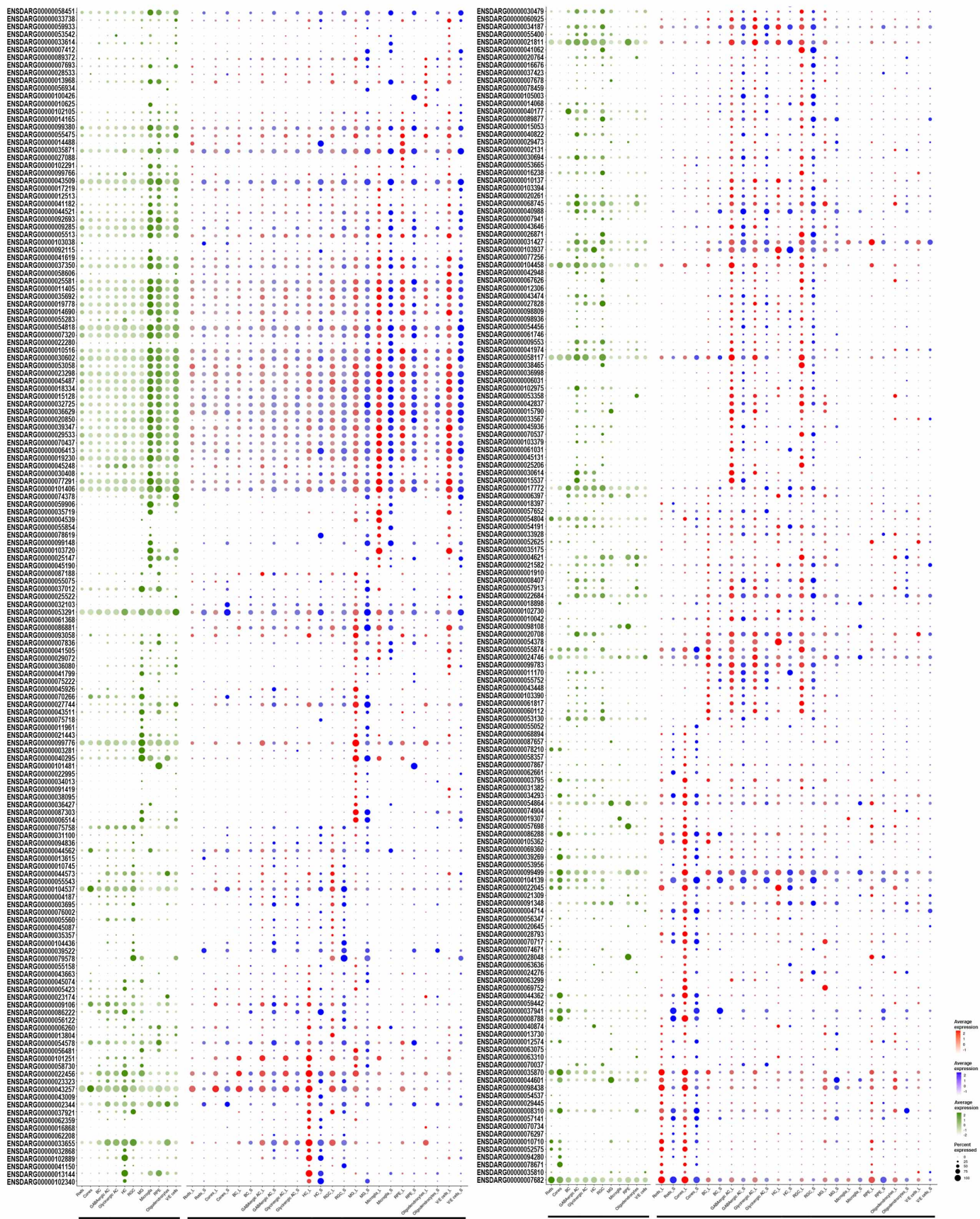
Supplementary Figure 6. Functional enrichment analysis of the 2,188 ohnolog pairs with higher gene expression in the S-ohnolog than in the L-ohnolog

(a) The x-axis represents the negative log₁₀-transformed *P*-value based on the accumulative hypergeometric distribution. The y-axis represents the enriched biological categories. Deeper color of the bar plot means smaller *P*-value. (b) Network representation of the statistically enriched categories in the functional enrichment analysis of the ohnolog pairs with higher gene expression in the S-ohnolog than in the L-ohnolog. The nodes represent the enriched categories and the edges are defined based on the similarities among their gene memberships. The name of the cluster is adopted from the name of the cluster with the smallest *P*-value among the biological categories contained in that cluster. The node size is proportional to the number of input ohnolog pairs grouped into each category.



Supplementary Figure 7. Gene expression profiles of ohnolog pairs with a similar expression pattern

Gene expression profile of the 326 ohnolog pairs in which both the L-ohnologs (red circles) and S-ohnologs (blue circles) exhibited the similar expression patterns. The expression patterns of zebrafish orthologs are also shown in green circles.



Zebrafish

Goldfish

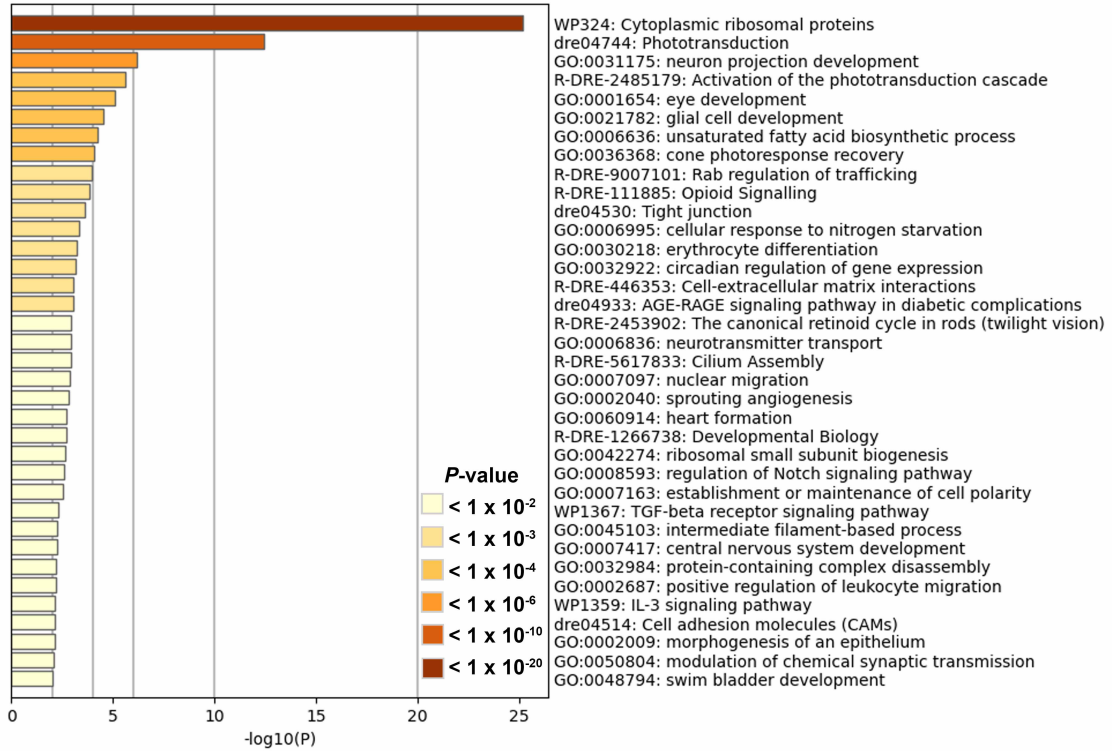
Zebrafish

Goldfish

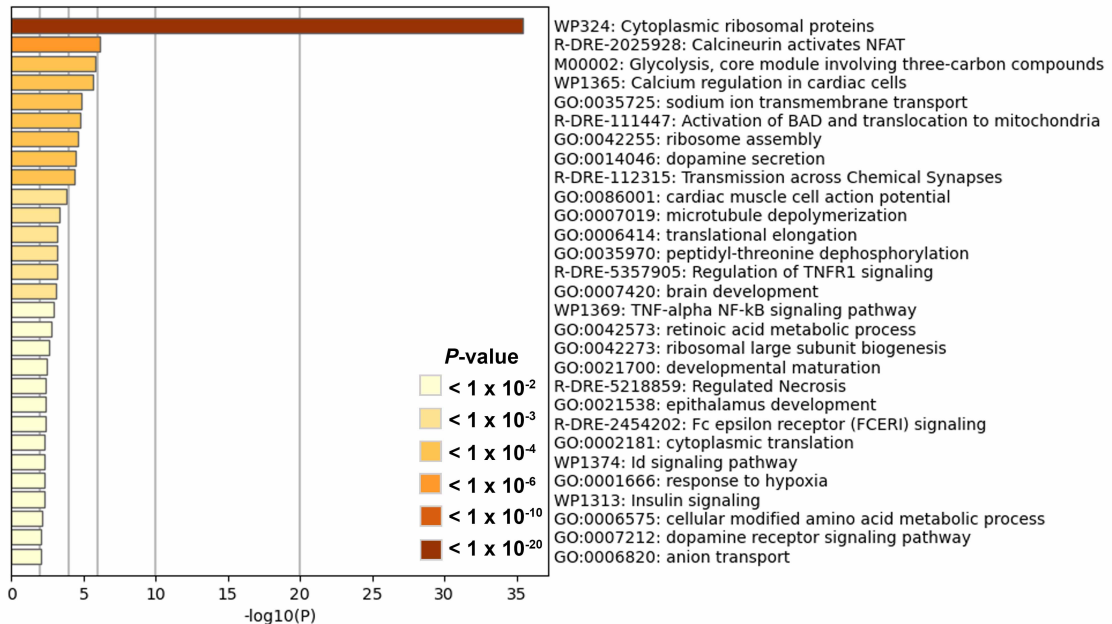
Supplementary Figure 8. Gene expression profiles of ohnolog pairs with a diversified expression pattern

Gene expression profile of the 306 ohnolog pairs showed different expression patterns in the L-ohnologs (red circles) and S-ohnologs (blue circles). The expression patterns of zebrafish orthologs are also shown in green circles.

a



b

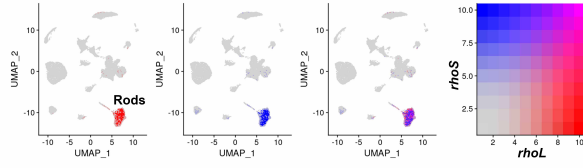


Supplementary Figure 9. Functional enrichment analysis of ohnolog pairs with a diversified expression pattern

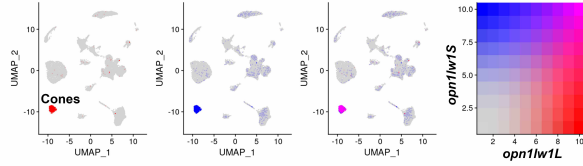
(a) Functional enrichment analysis of the 326 ohnolog pairs in which both the L- and S-ohnologs exhibited the similar expression patterns. (b) Functional enrichment analysis of the 306 ohnolog pairs showed different expression patterns in the L- and S-ohnologs. The x-axis represents the negative log₁₀-transformed *P*-value based on the accumulative hypergeometric distribution. The y-axis represents the enriched biological categories.

a

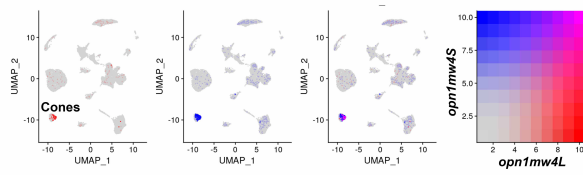
rho



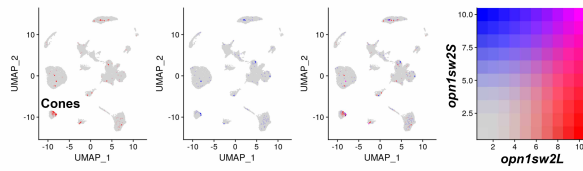
opn1lw1



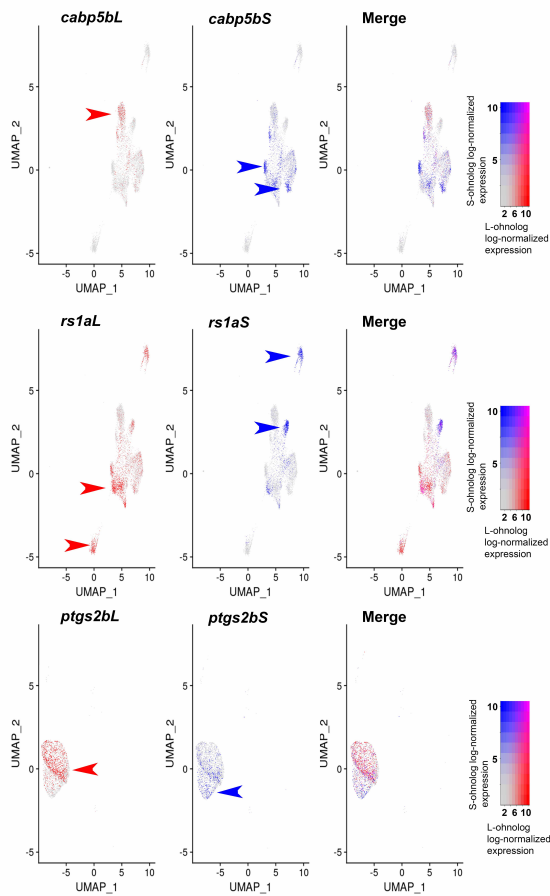
opn1mw4



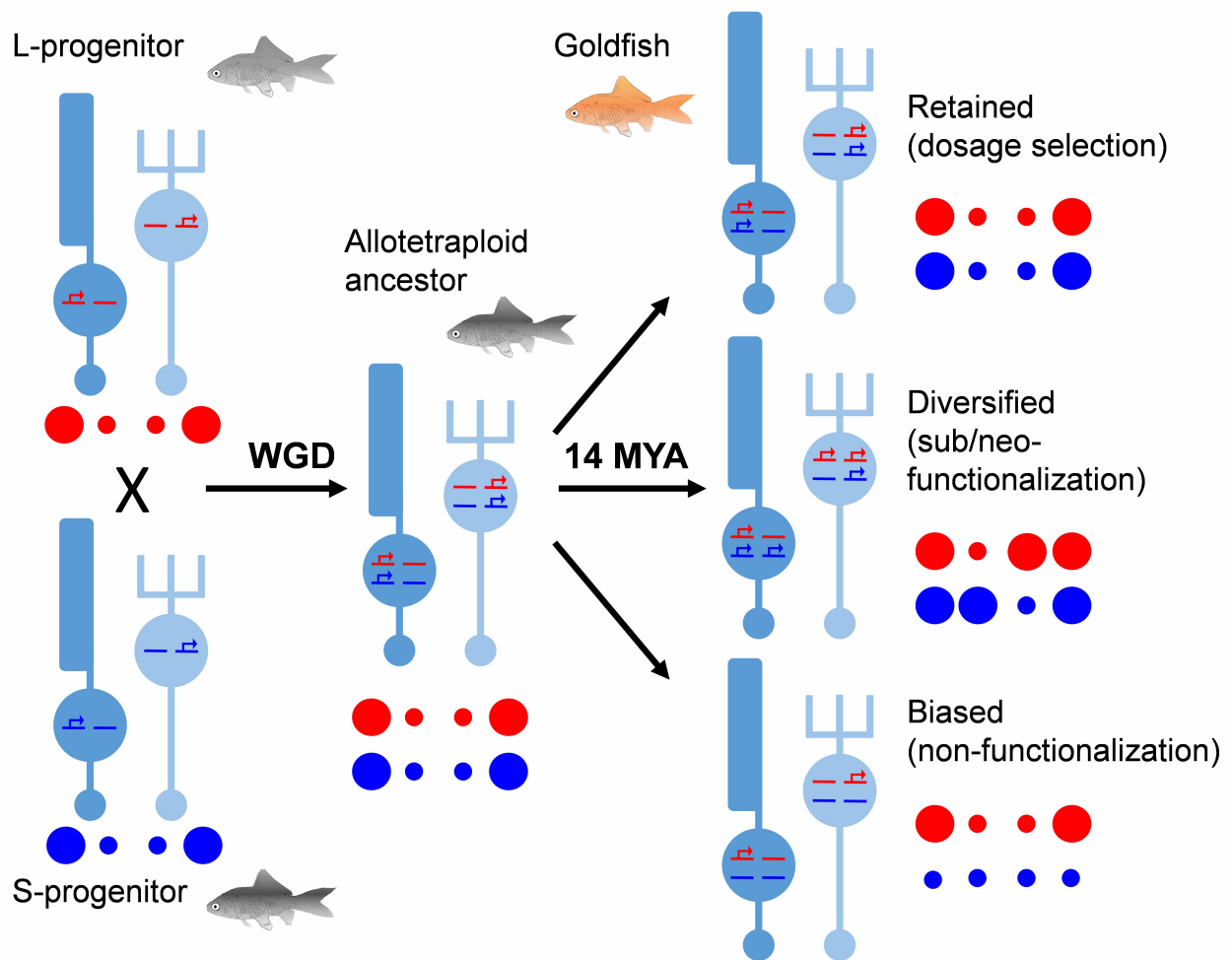
opn1sw2



b



Supplementary Figure 10. Expression of ohnolog pairs in the retina. (a) Expression of opsin L/S ohnolog pairs. The gene expression levels of opsin ohnologs in the goldfish retina are shown. (b) Ohnolog pairs in which the L-ohnolog and S-ohnolog show different expression profiles in retinal cell clusters. The *cabp5b* and *rs1a* ohnologs exhibit diversified expression patterns in the bipolar cells of the UMAP cluster. The *ptgs2b* ohnologs exhibit diversified expression pattern in the Müller glia of the UMAP cluster. The red or blue arrowheads show the enriched cells expressing L- or S-ohnologs, respectively.



Supplementary Figure 11. Diagram showing the transcriptional divergence in the asymmetrically evolved subgenomes

The ancestor of goldfish experienced an allotetraploidization event 14 MYA. Ohnologs have various types of fates after WGD. A group of genes underwent dosage selection, accounting for 5% of the total 11,444 ohnolog pairs. In the goldfish retina, 306 putative sub/neo-functionalized ohnolog pairs were undergo cell-type-specific genetic variation. These diversifications of ohnolog expression patterns are likely to contribute to the functional evolution of the goldfish retina.