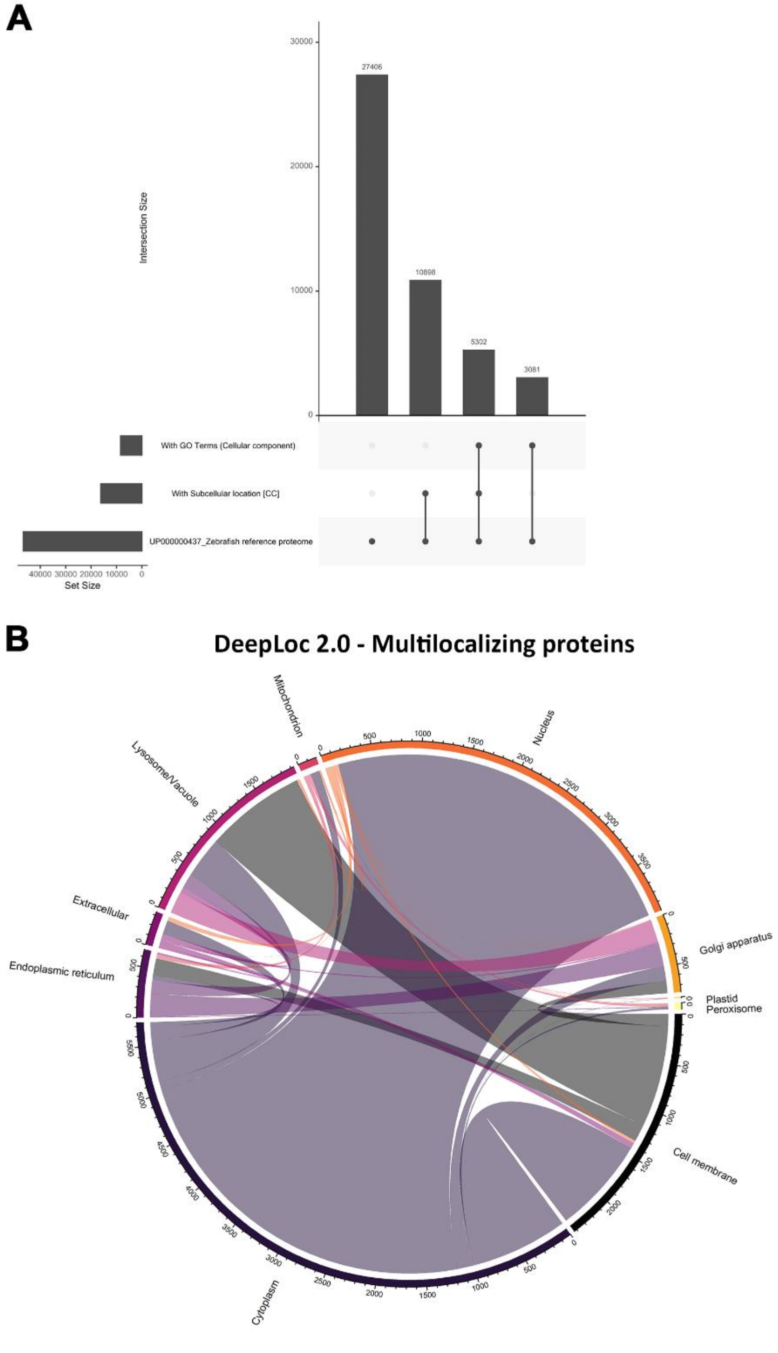
**iScience, Volume *26***

## Supplemental information

# A ligand-receptor interactome atlas of the zebrafish

**Milosz Chodkowski, Andrzej Zielezinski, and Savani Anbalagan**

# Supplementary figures



DeepLoc 2.0 - Multilocalizing proteins

Fig. S1 Chodkowski M. *et al.*,

Figure S1. Annotation status and DeepLoc 2.0 multilocalizing protein localizations; related to Fig. 1.

(**A**) Current status of the cellular localization annotation records of the reference proteome from zebrafish. Upset plot comparing the number of proteins with Gene ontology (GO) term- or UniprotKB-based cellular localization (CC) records.

(**B**) Circle plot showing the number of proteins predicted to localize to multiple cellular regions and organelles.

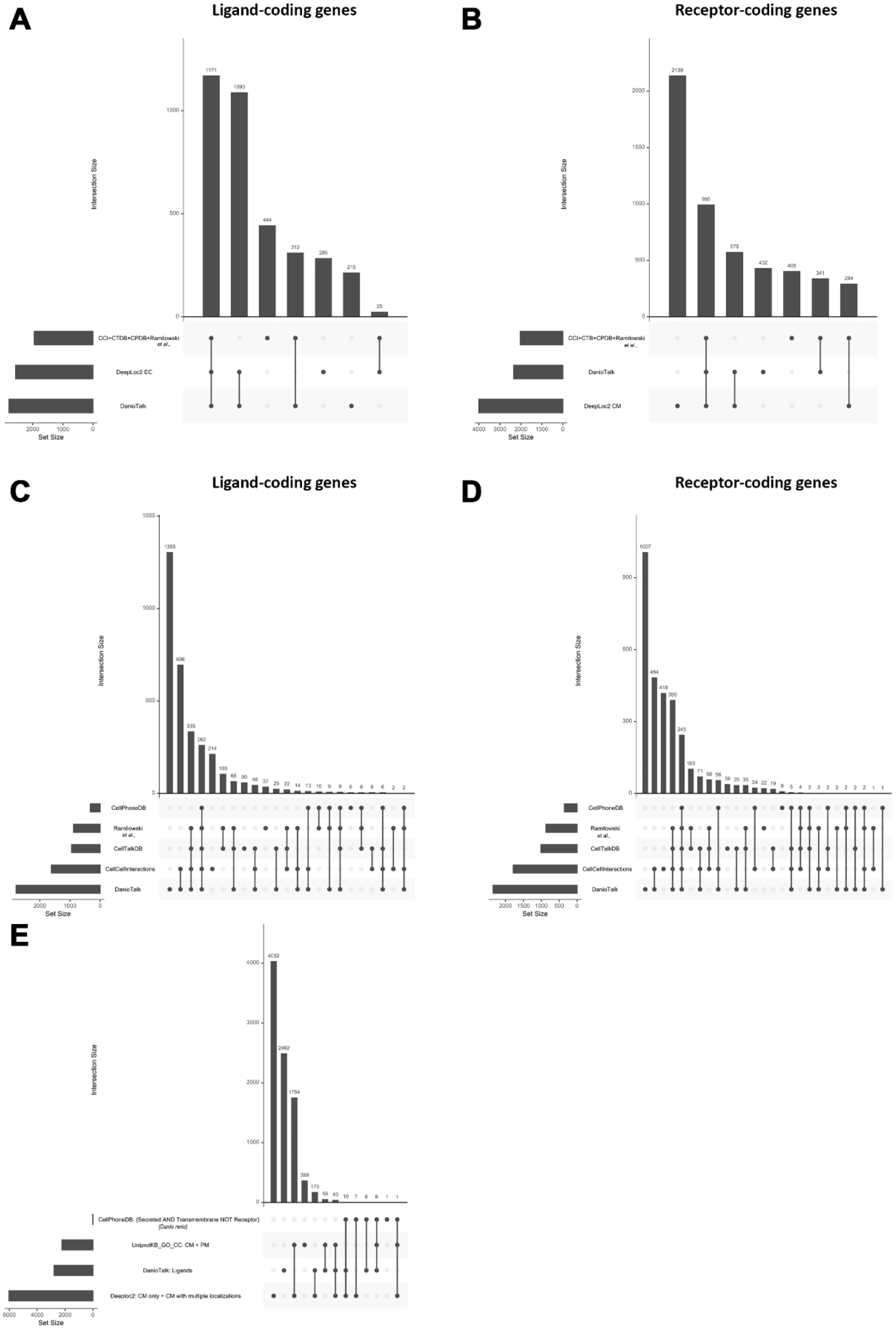


Fig. S2 Chodkowski M. *et al.,*

Figure S2. Comparison of DanioTalk ligand-receptor database with human ligand-receptor database; related to Fig. 2.

(A) Upset plot comparing the number of genes coding for curated ligands with DeepLoc2-predicted extracellular-only proteins and ligands in human ligand-receptor databases (pooled).

(B) Upset plot comparing the number of genes coding for curated receptors with DeepLoc2-predicted cell membrane-only proteins and receptors in human ligand-receptor databases (pooled).

(C) Upset plot comparing the number of genes coding for curated ligands with DeepLoc2-predicted extracellular-only proteins and ligands in human ligand-receptor databases.

(D) Upset plot comparing the number of genes coding for curated receptors with DeepLoc2-predicted cell membrane-only proteins and receptors in human ligand-receptor databases.

(E) Membrane-bound ligands in curated ligand database. Upset plot comparing the number of genes coding for curated ligands with cell membrane proteins in UniprotKB or DeepLoc2 predictions and zebrafish orthologs of non-receptor, secreted but transmembrane proteins in CellPhoneDB.

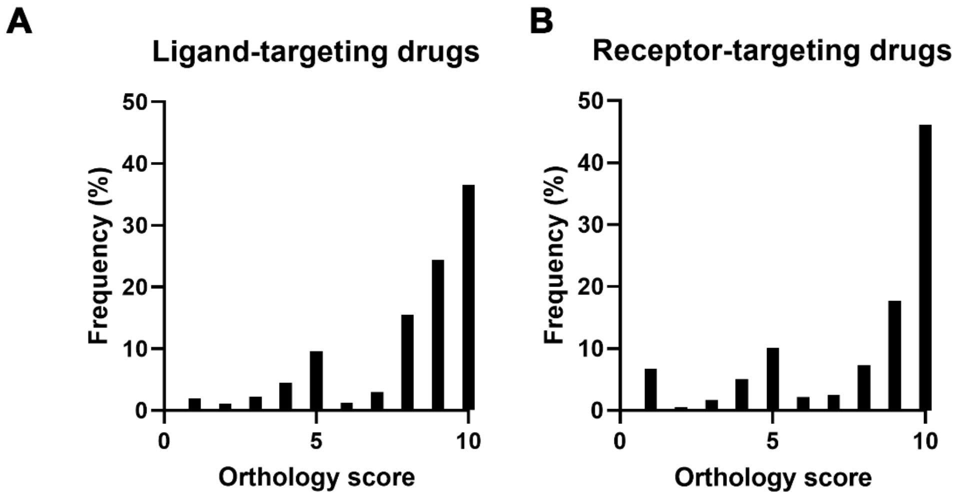


Fig. S3 Chodkowski M. *et al.*,

**Figure S3. Drug targetability and degree of conservation of zebrafish ligand- and receptor- coding genes**; related to Fig. 2.

- **(A)** Histogram showing the orthology score for zebrafish ligand-coding genes and their human orthologs whose coded proteins can be targeted by drugs listed in DrugCentralDB.
- **(B)** Histogram showing the orthology score for zebrafish receptor-coding genes and their human orthologs whose coded proteins can be targeted by drugs listed in DrugCentralDB.

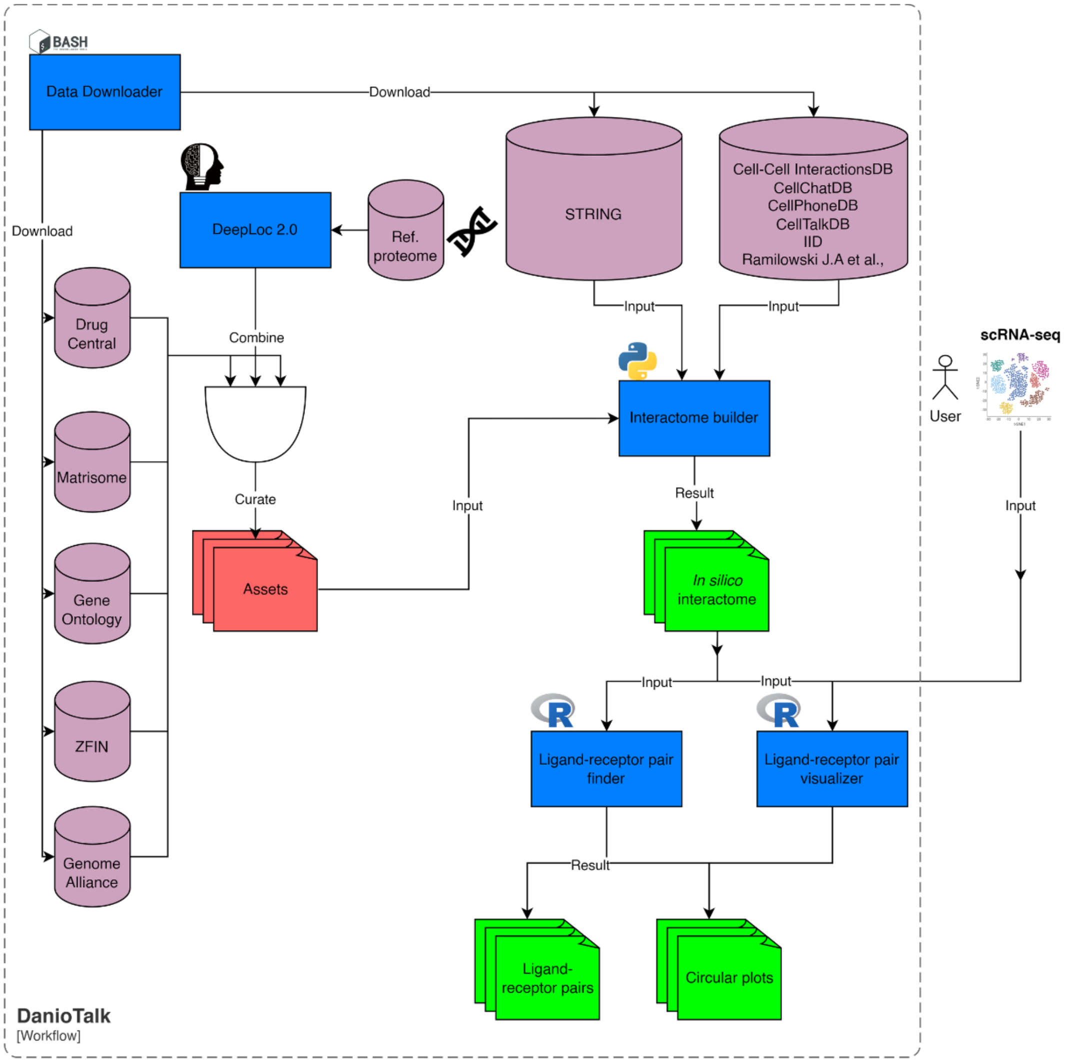


Fig. S4 Chodkowski M. *et al.*,

**Figure S4. DanioTalk workflow**; related to Fig. 3.

Chart depicting the workflow of DanioTalk scripts. DanioTalk scripts are written in bash, python and R. DanioTalk is built based on upon multiple human databases for LR pairs, protein annotations and to identify drugs that can target ligands and receptors. For databases that have a permanent url (zebrafish gene aliases, orthologs information, GO term annotation and drug list from DrugCentralDB), DanioTalk scripts are designed to download the latest database versions when rebuilding the LR databases or annotations.

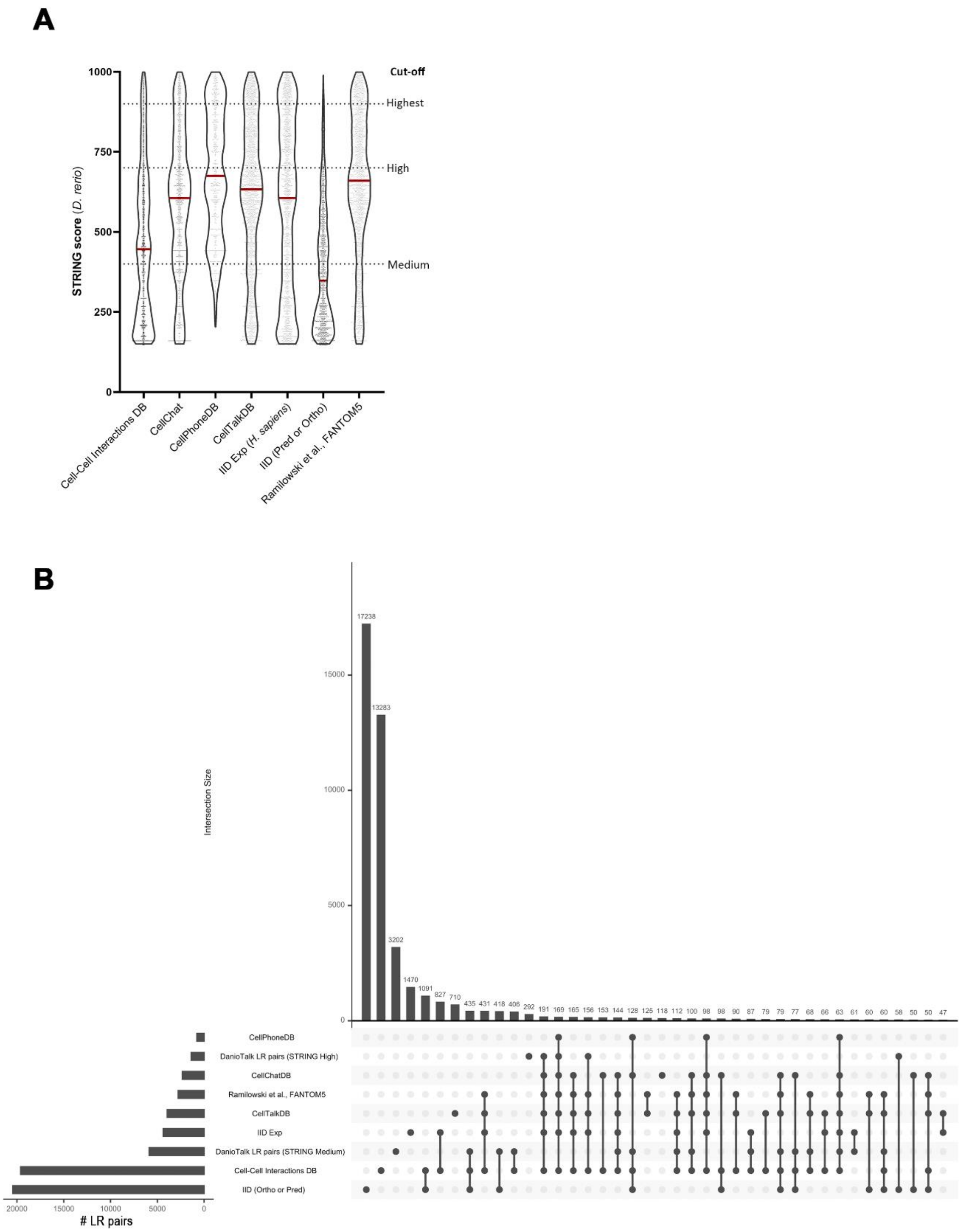


Fig. S5 Chodkowski M. *et al.,*

**Figure S5. Confidence and source of DanioTalk LR pairs**; related to Fig. 3.

- **(A)** Violin plot comparing the zebrafish STRING PPI score for orthologous LR pairs in human LR databases and human IID database.
- **(B)** Upset plot comparing the subset of DanioTalk LR pairs (STRING High and Medium confidence zebrafish LR pairs) with orthologs of human LR pair and PPI databases.

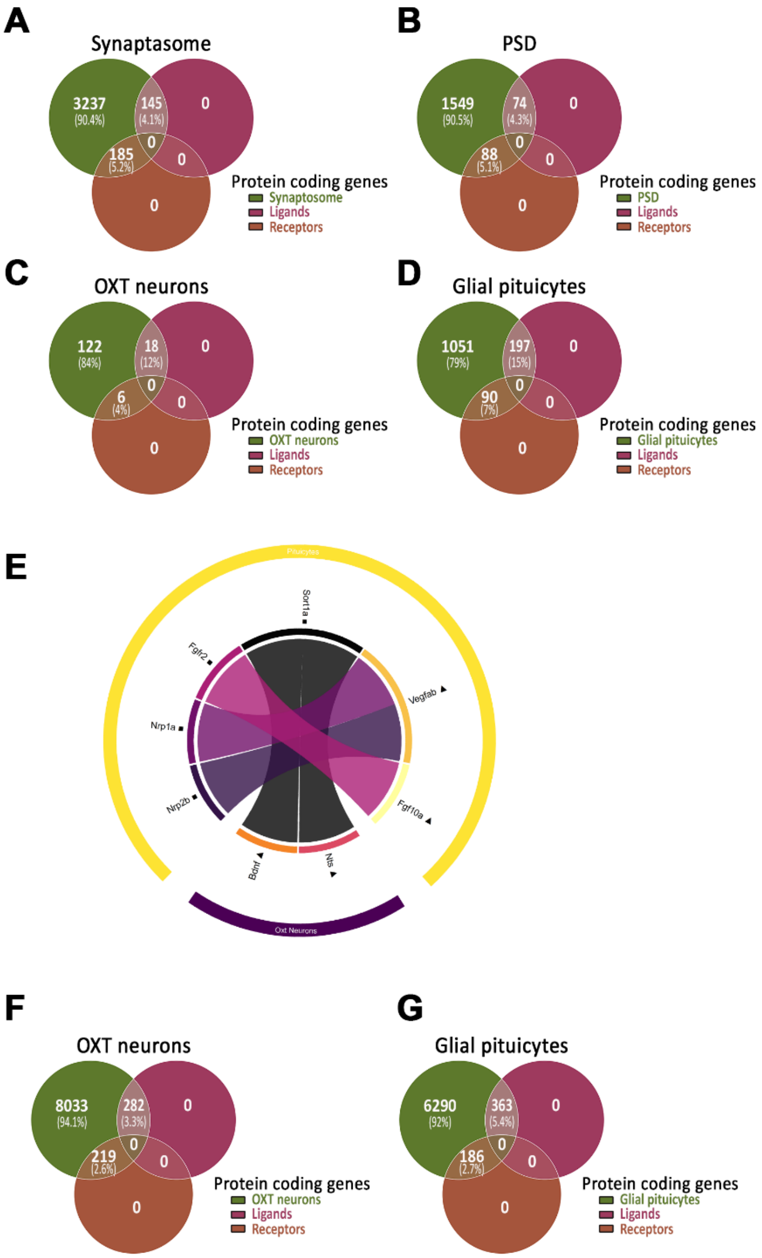


Fig. S6 Chodkowski M. *et al.,*

**Figure S6. Application of DanioTalk**; related to Fig. 4.

- **(A)** Venn diagram showing the number and percentage of genes coding for synaptosome-enriched ligands and receptors.
- **(B)** Venn diagram showing the number and percentage of genes coding for PSD-enriched ligands and receptors.
- **(C)** Venn diagram showing the number and percentage of ligand- and receptor-coding genes identified in oxytocin scRNA-seq differentially expressed dataset ($\log_2$FC>2, p-adj<0.05).
- **(D)** Venn diagram showing the number and percentage of ligand- and receptor-coding genes identified in Glial pituicytes differentially expressed dataset ($\log_2$FC>2, p-adj<0.05, average AMCA+ read count >50).
- **(E)** Circle plot showing top-ranked differentially expressed, ligand-receptor interactions between oxytocin neurons and glial pituicytes. Interactions were ranked based on gene expression and STRING physical interaction score (>900).
- **(F)** Venn diagram showing the number and percentage of ligand- and receptor-coding genes identified in oxytocin pseudo-bulk RNA seq dataset (Expression >0.05).
- **(G)** Venn diagram showing the number and percentage of ligand- and receptor-coding genes identified in Glial pituicytes differentially expressed dataset (Average AMCA+ read count >50).