Fig. S1A -B

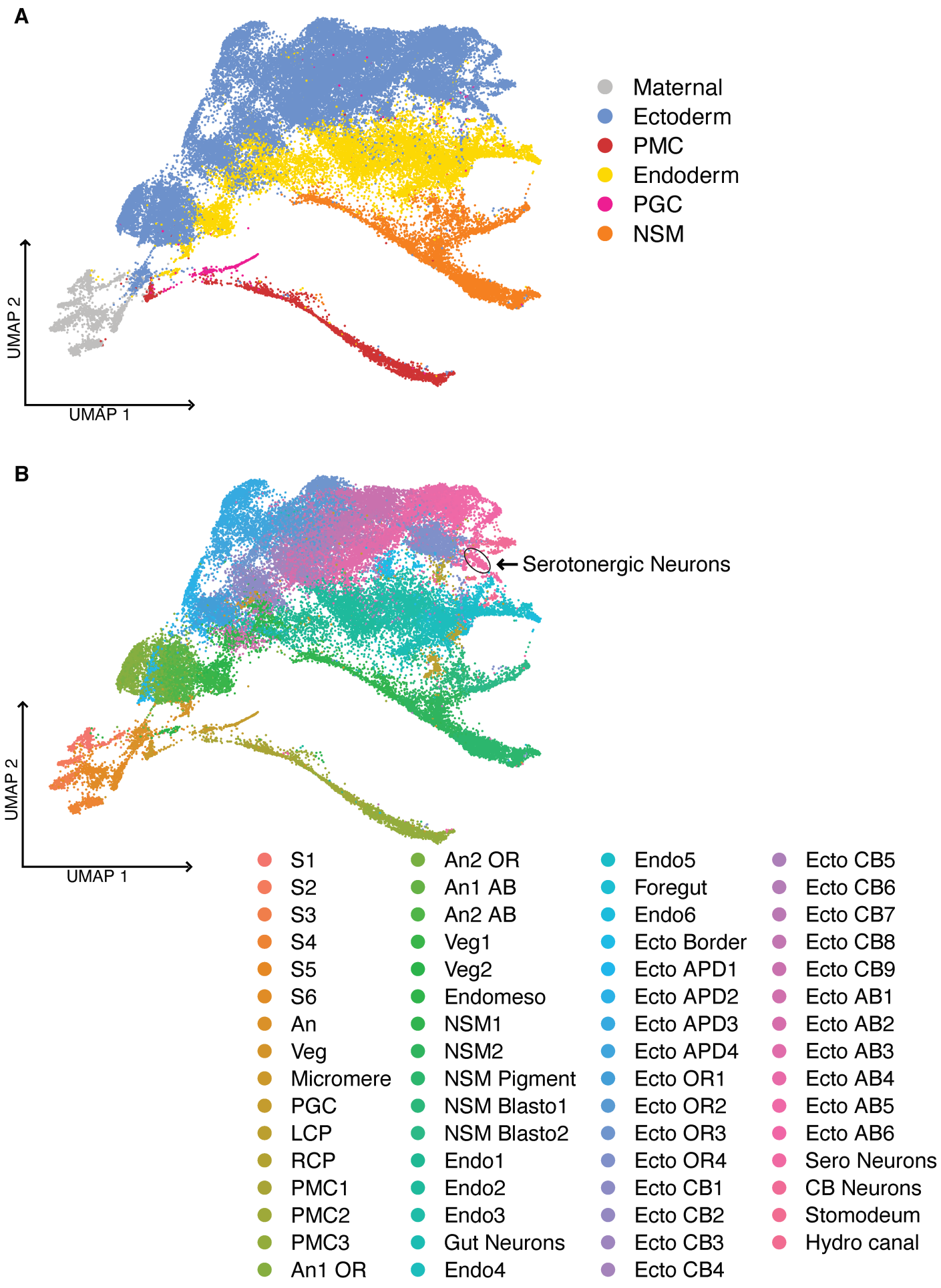


Fig. S1C

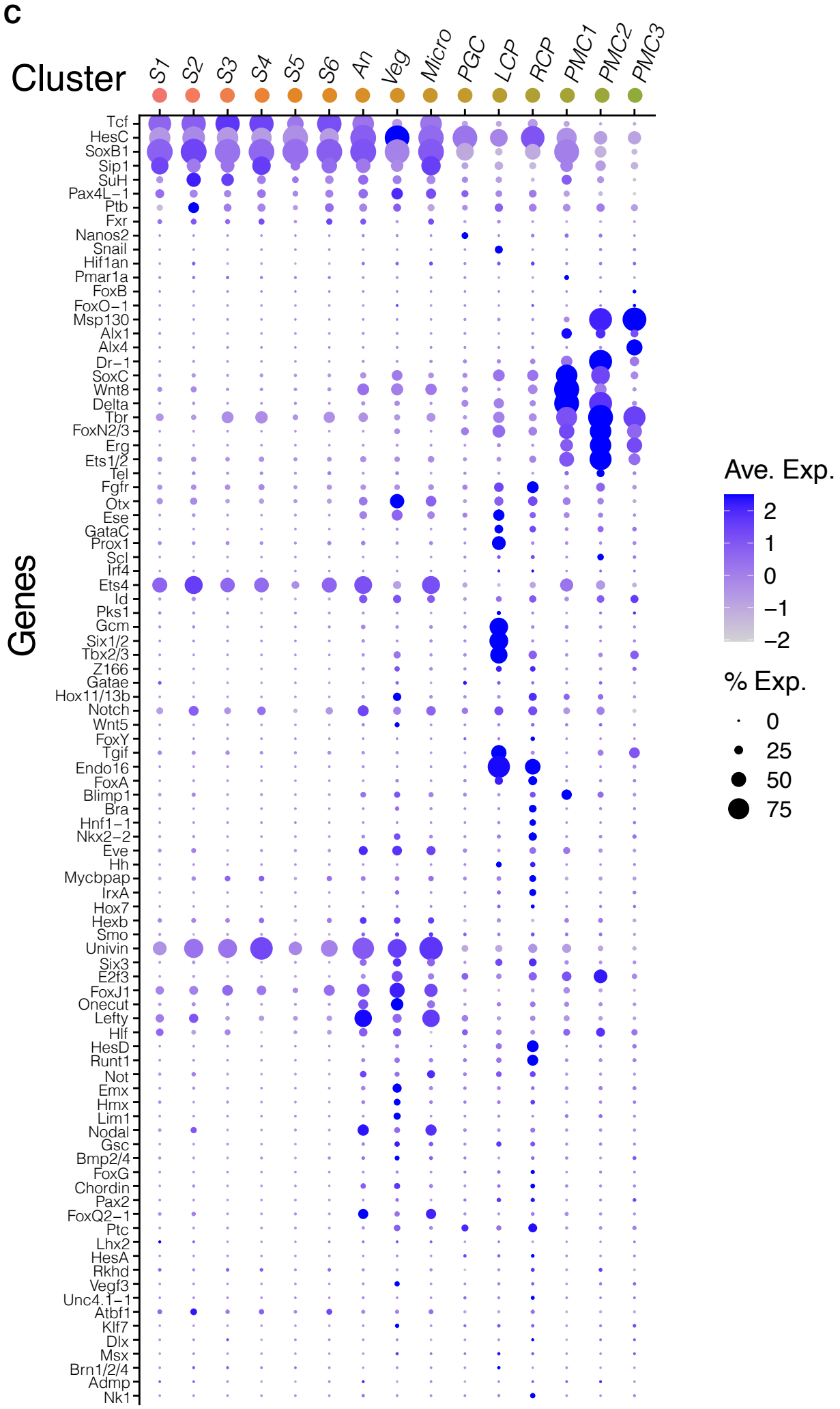


Fig. S1D

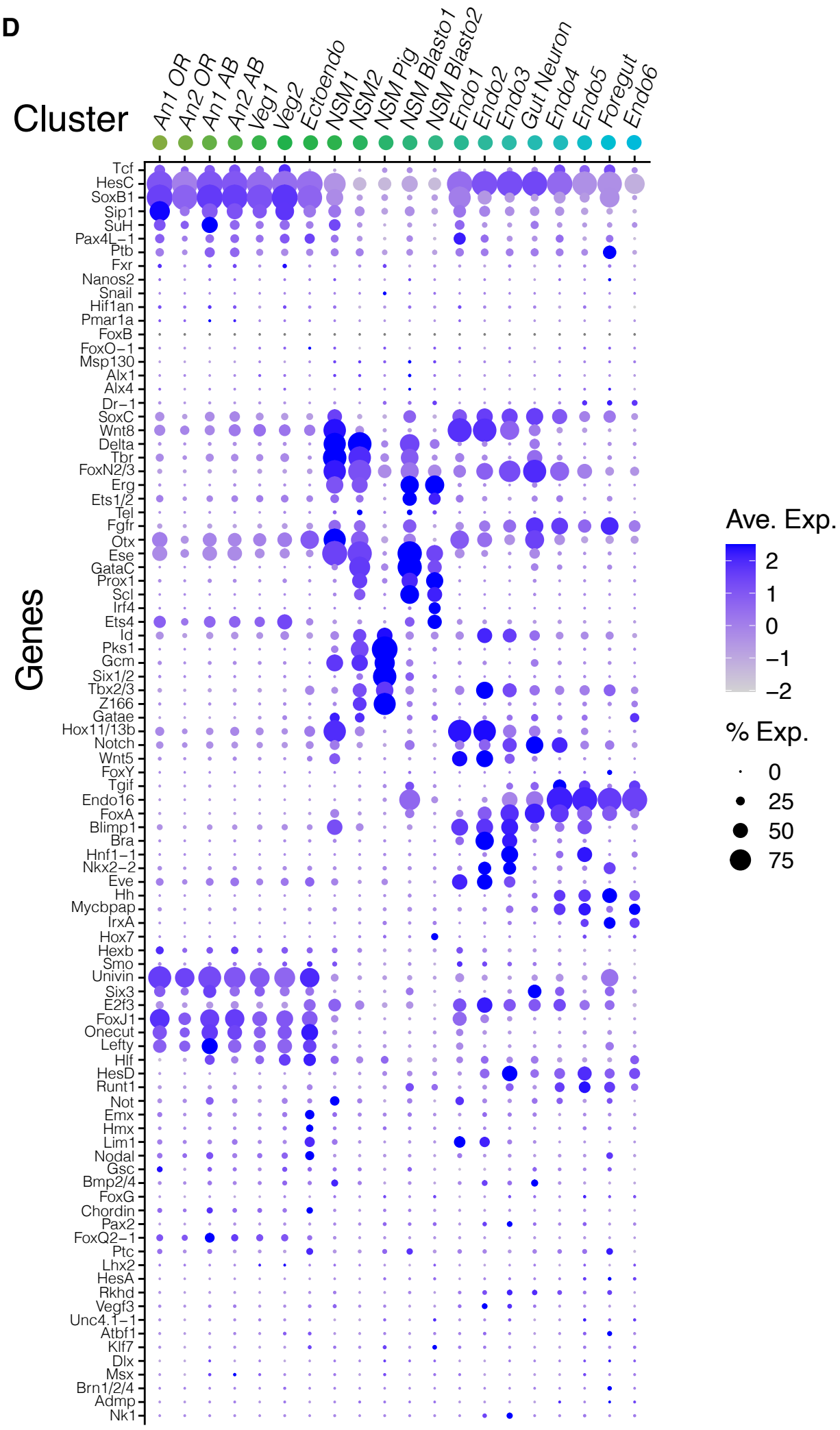


Fig. S1E

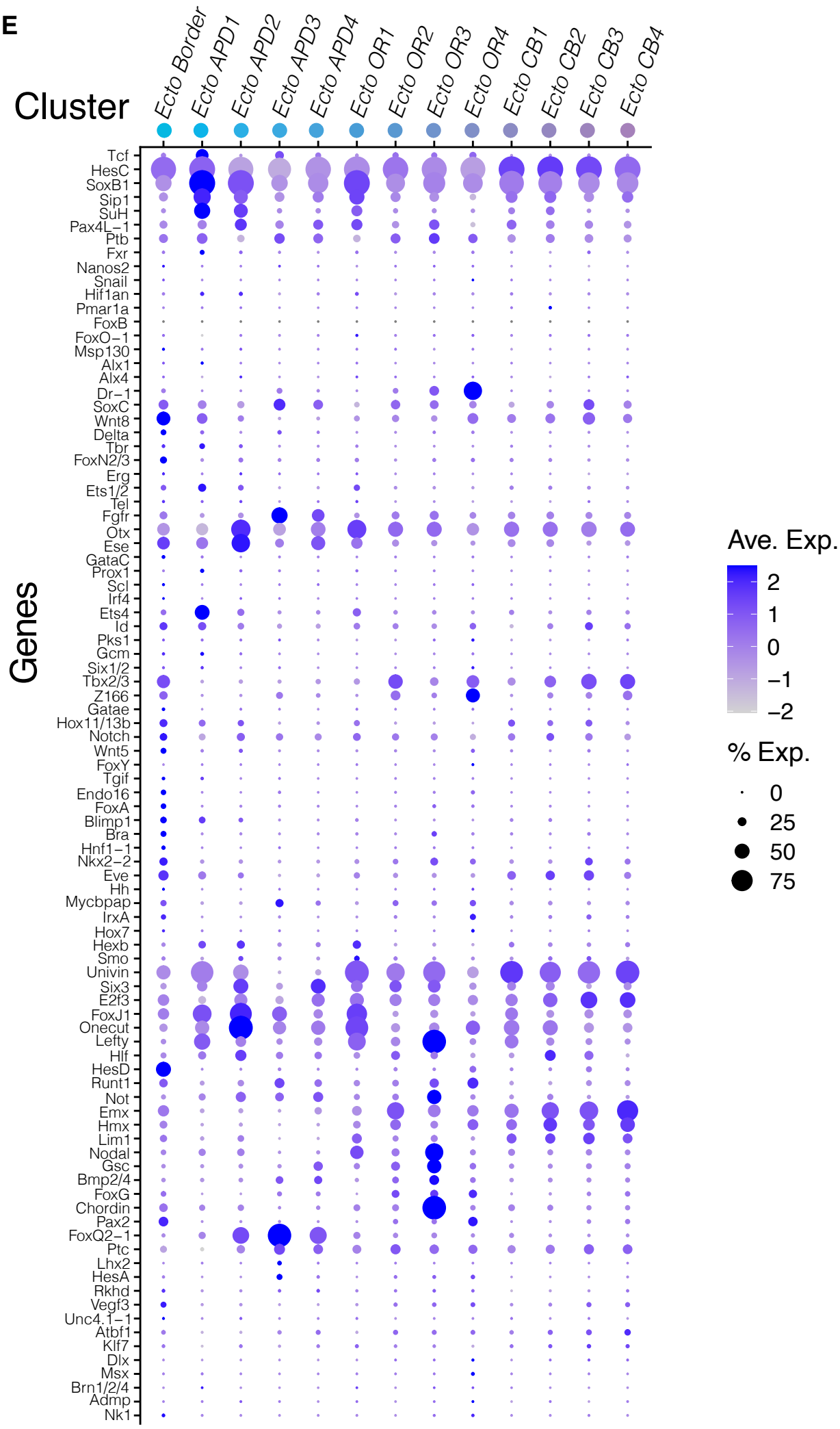


Fig. S1F

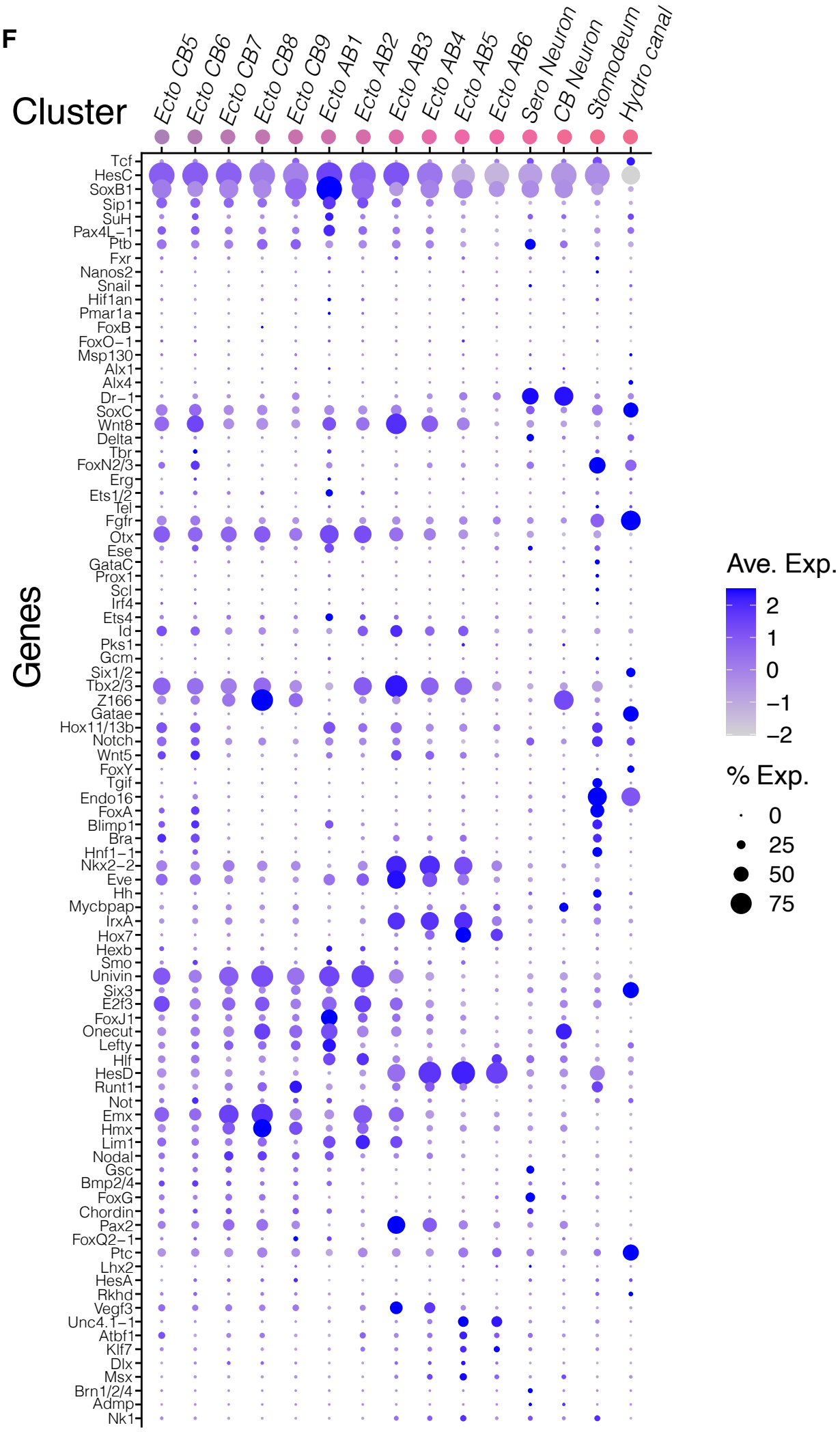


Fig. S1. UMAP plot mapped according to clusters. (A). Sixty-three clusters were produced using signature genes of cell types present at the 24 hpf stage. These spatio-temporal clusters marking trajectories toward cell fates at 24 hpf were collapsed into the six major lineages and colored according to cell type. (B). The sixty-three clusters. Each cluster was probed to determine eventual cell fate. Many of the clusters report state changes of the same lineage at the 2.4 sensitivity level selected. The UMAP plot is the same as in Fig. 1 but this time the clusters are identified by color and by eventual fate. (C-F) Dotplots of the 63 clusters and the 93 genes used to identify each of those clusters. S1-S6 – early blastomeres prior to fate specification. An1, An2, Veg1, Veg2 – blastomeres at 6th cleavage identified by tier from the Animal pole toward the Vegetal pole. Ecto – ectoderm fate: APD – animal pole domain, OR – oral, AB – aboral, Endo – endoderm, PGC – primordial germ cell, PMC– Primary mesenchyme cell (skeletogenic cell), NSM – non-skeletal mesoderm, RCP, LCP – right and left coelomic pouch, N – neural.

Fig. S2

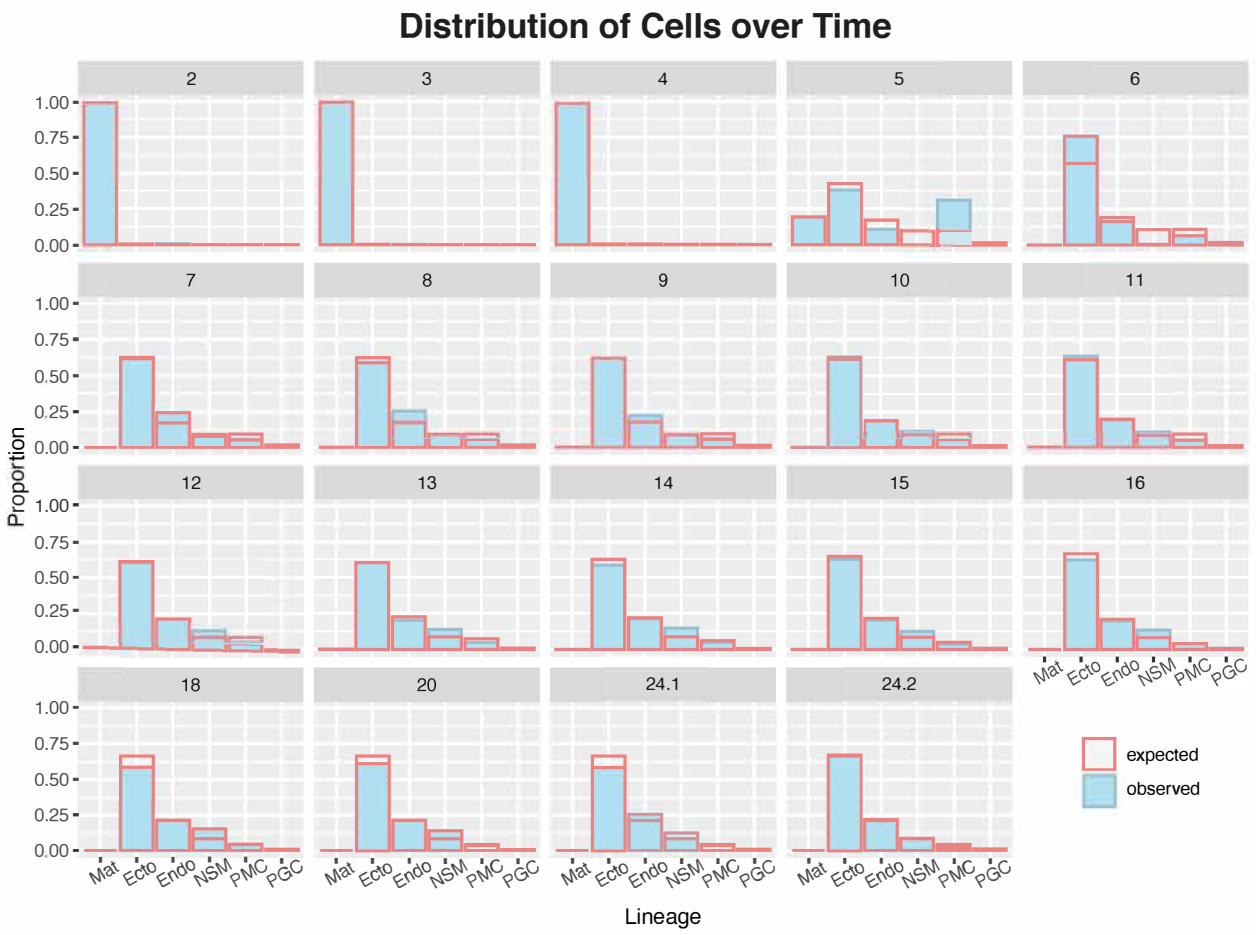


Fig. S2. Cell lineage analysis: Expected vs observed. Data from earlier cell lineage analyses (see Table S1) provided the expected distribution of cells at each time point based on the stereotypic pattern of cleavage. The observed distribution was established by the optimal transport method of Waddington-OT (Schiebinger et al., 2019). That method, using expression of all genes per cell, came into agreement with the expected lineages beginning with the 6 hpf timepoint, with the 5 hpf timepoint expressing suggestive but still imperfect matches with the expected. After that time point there was excellent agreement between the expected vs the observed lineage calls at each time point. This provides evidence that at those time points the cell sampling was not biased toward one lineage, and no major lineage was excluded.

Fig. S3A

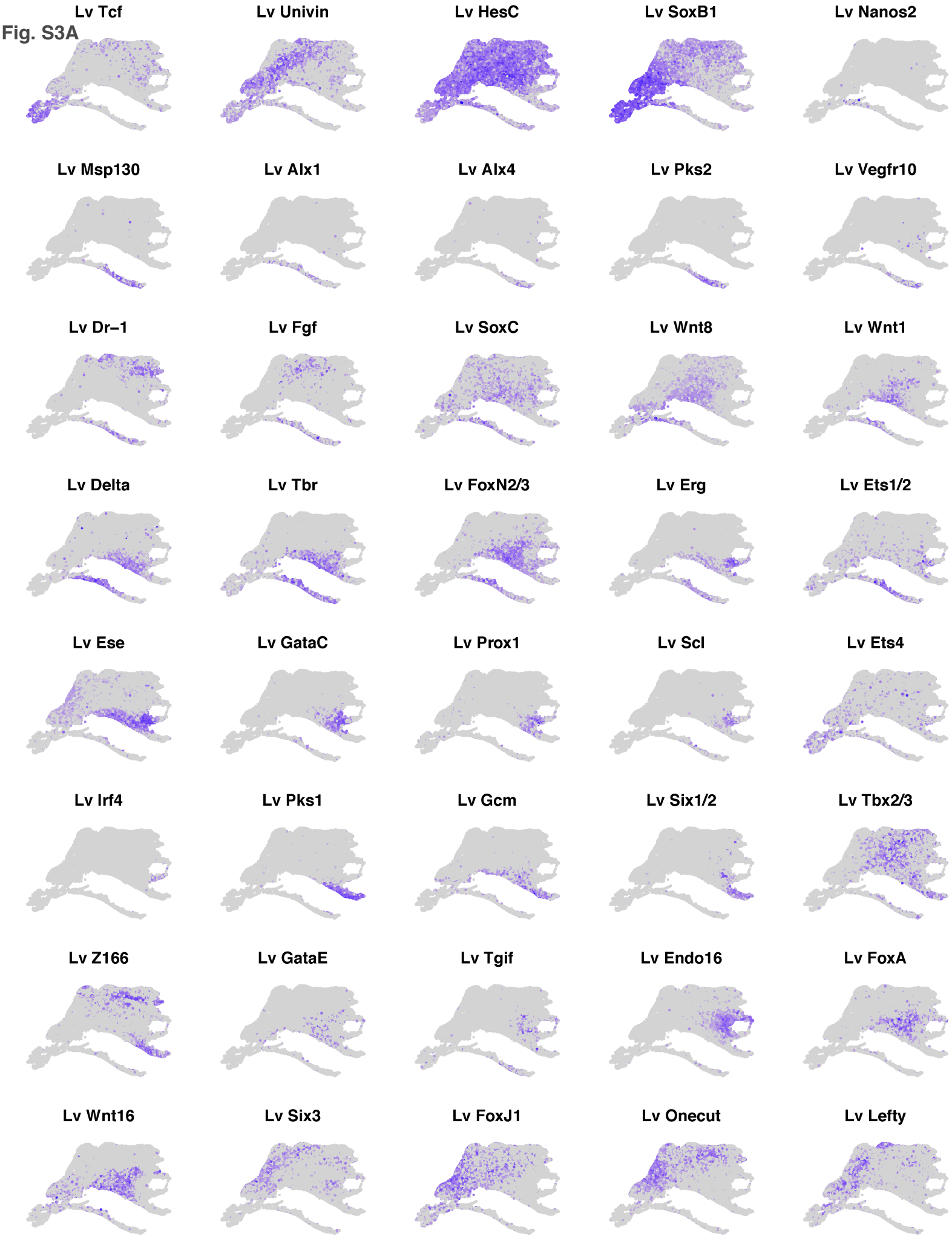


Fig. S3B

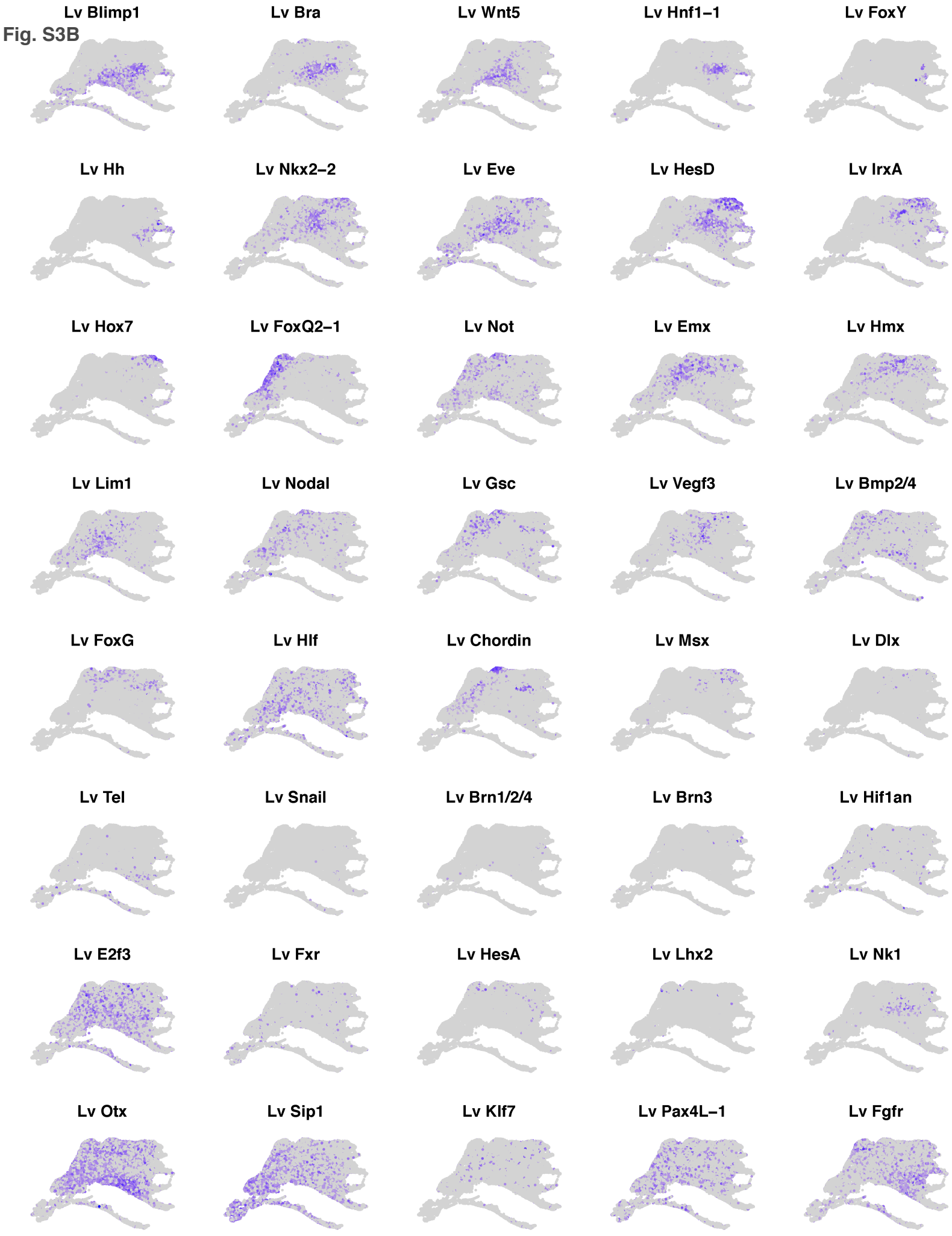


Fig. S3. A and B. dGRN genes mapped on the UMAP. The expression of each of 80 genes included in this analysis mapped onto the UMAP.

Fig. S4

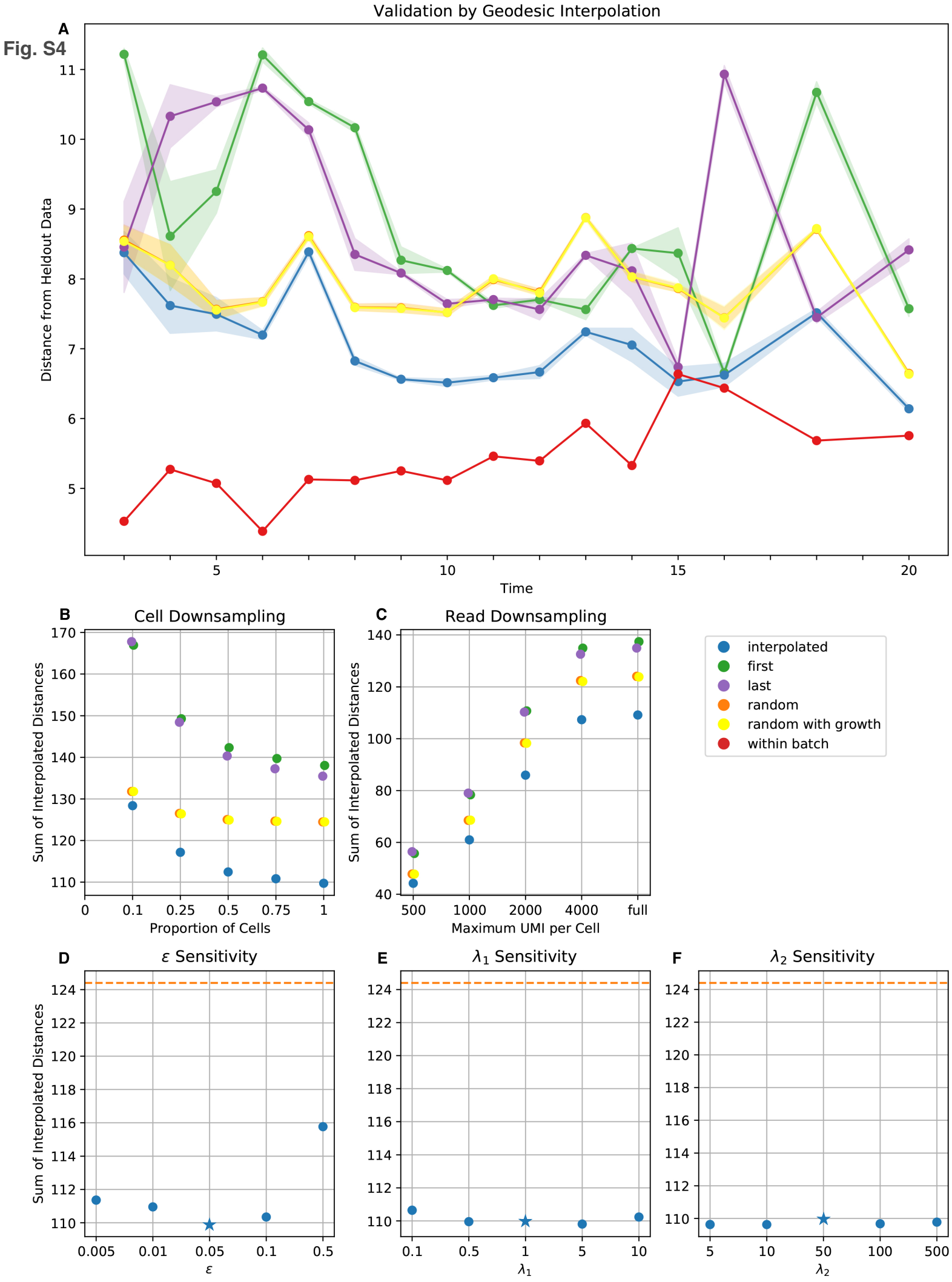


Fig. S4. Geodesic Validation of the Waddington-OT approach. (A). Validation by geodesic interpolation. We validated our results by demonstrating that we can interpolate the distribution of cells at held-out timepoints. For each triplet of consecutive time-points (e.g. 5,6,7 or 6,7,8, etc), we held out the data from the middle time-point and attempted to reconstruct it by connecting the first to the third. We then quantified our performance by comparing to the held-out midpoint. The blue curve shows the results from optimal transport, which is lower than various noise models (yellow, green, purple). The red curve shows the distance between random sub-divisions of the held-out midpoint as a base-line to compare against. (**B-D).** Sensitivity to parameter selection. Blue dots indicate performance of optimal transport for various parameter settings. Performance is quantified by the area under the blue curve from **(A).** The blue star indicates the parameters we selected. The horizontal dashed line indicates the performance of the null-model (area under the orange line in **(A)**). (**B).** Sensitivity to choice of entropic regularization parameter. (**C-D).** Sensitivity to choice of regularization parameters for unbalanced transport. (**E-F).** Sensitivity to downsampling cells and reads. Similar to the plots (**B- D)**, the y-axis quantifies performance in terms of the area under the curve shown in **(A)**, but for different numbers of cells **(E)** or reads **(F)**.

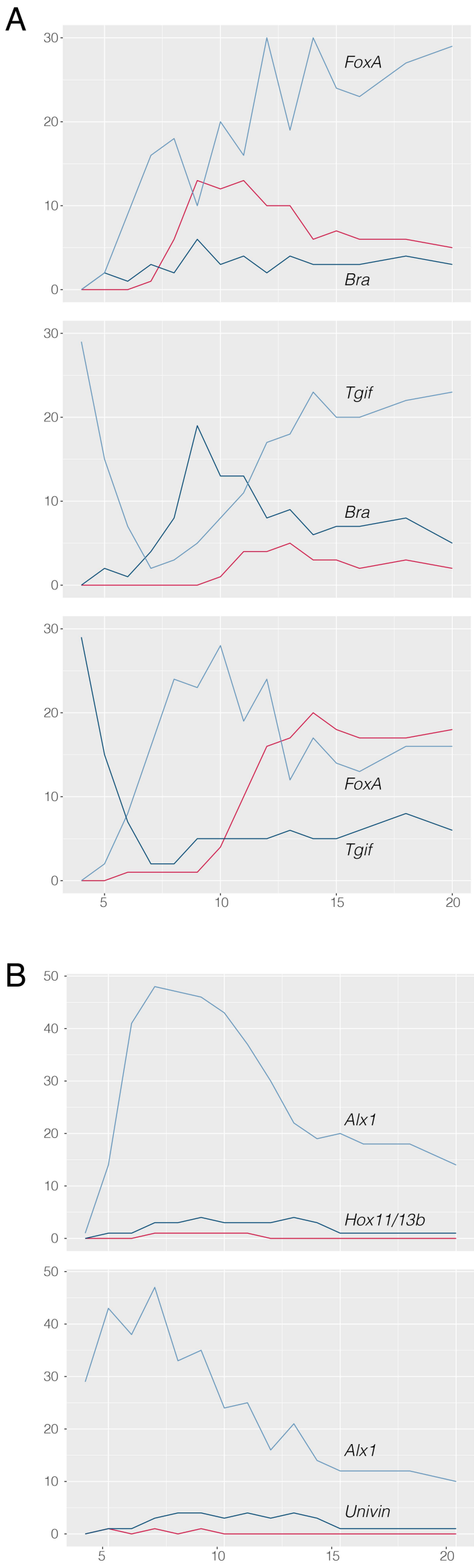
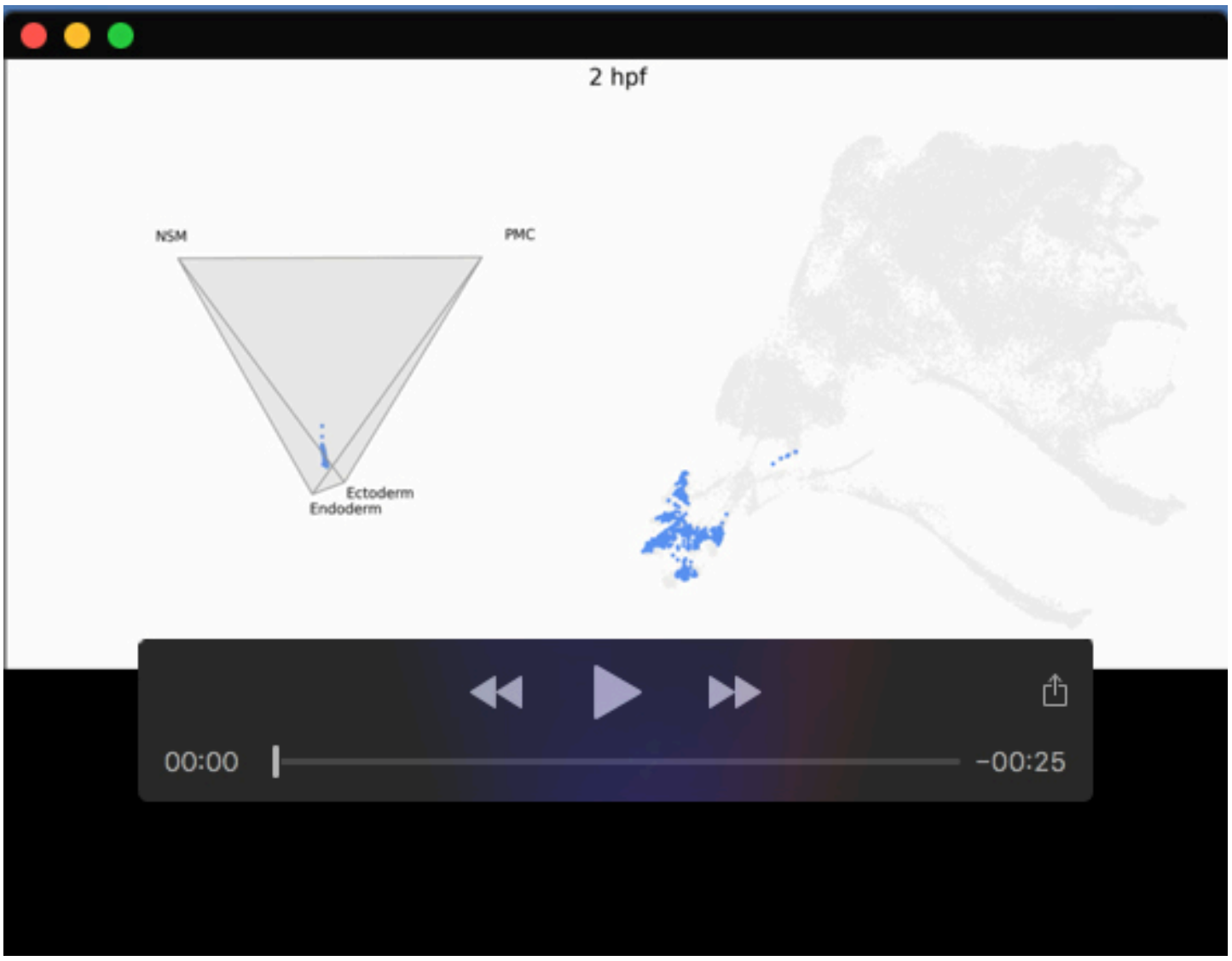


Fig. S5. Co-expression of genes. Additional comparisons of pairs of genes expressed in a survey of all cells in the database or in a cluster. Lines in blue indicate percent of cells expressing one of the two genes surveyed, and the line in red are percent of cells expressing both genes surveyed. **A.** Endomesoderm to endoderm expression. In the top panel *foxA* is expressed in the endomesoderm and at about 7-8 hpf many *foxA* expressing cells are co-expressed with *bra* which is expressed only in definitive endoderm. This indicates that the first definitive endoderm sells include *foxA*. *Tgif* is a maternal transcript that is expressed in endoderm beginning at 10-11 hpf as shown by co-expression with *bra* at that time. *Bra* is expressed in a subset of endoderm cells and then its expression is inactivated. *FoxA* by contrast, once activated in endoderm, continues to be expressed by that germ layer. The bottom panel of (A) shows the consequence of this. *Tgif,* as above, is activated in endoderm at 10-11 hpf and it is co-expressed with *foxA* from that point forward. **B.** The cluster of PMCs expresses *alx1*. In a comparison, that cluster does not express or co-express endoderm as marked by *hox11/13b,* nor does it express *univin*, a marker for ectoderm. **C.** Endomesoderm co-expression lasts until about 12 hpf. *Delta* and *ese* (mesoderm markers) are co-expressed with *hox11/13b* (endoderm marker) from 5-6 hpf until 12-14 hpf. Definitive mesoderm and Endoderm cells accumulate during this time asynchronously. **D.** *nodal* (oral ectoderm) and *foxQ2* (animal pole domain) are co-expressed in some ectoderm cells early in cleavage but soon are expressed in separate regions of the ectoderm. In the lower panel *lefty* and *nodal* are co-expressed at high levels at two intervals, the first is known to establish the oral-aboral axis (5-9 hpf) and the second is known to establish right-left asymmetry (13-16 hpf, or mid- to late-gastrula) (Duboc, et al., 2004; Duboc, et al., 2005).

Table S1. Cell lineage numbers. Cleavage of sea urchin embryos is stereotypic providing a predictable distribution of cell numbers to each of the five major lineages through the sixth cleavage and approximations based on cell doubling rates for each lineage through gastrulation. (Cameron et al., 1987; Logan and McClay, 1997; Logan and McClay, 1999; Martik and McClay, 2017; McClay et al., 2020). The divisions of ectoderm and endomesoderm cells occurs synchronously until the sixth cleavage. Micromeres (PMC lineage) and PGCs divide less frequently as indicated. After the sixth cleavage ectoderm progenitors slow their rate of cleavage relative to endomesoderm progenitors. The Veg1 tier of endoderm cells appears in this sixth cleavage which is equatorial in the vegetal hemisphere. At eighth cleavage an equatorial cell division augments the endoderm with the addition of upper Veg2 cells, while the lower Veg2 tier cells become non-skeletal mesoderm cells. The final cell count for each lineage is reached somewhat asynchronously with the approximate final number being reached for each lineage by 15hpf. From that hour until larvae begin to feed there are few additional cell divisions. Several variables exist that make these latter cell numbers approximate. The position of the equatorial third cleavage varies. This results in a variable number (though small) of Veg1 cells becoming ectoderm cells. The asymmetric 4th cleavage varies occasionally such that one or two micromeres arise that are larger or smaller than their sibs. This results eventually in a range of 60-68 PMCs with 64 found in the vast majority of embryos. The Waddington OT series used these cell numbers and approximations of the variations observed (see methods).

Lineage	3 hpf	4 hpf	5 hpf	6 hpf	7 hpf	8 hpf	9 hpf	10 hpf	11 hpf	15 hpf	
Ectoderm	8	16	32	64	64	128	256	256	512	1024	
Endmesoderm	4	8	8	16	16						
Endoderm			8	16	16	48	95	192	384	384	
Mesoderm			8			16	32	64	128	128	
PMC	4	4	8	8	8	16	32	64	64	64	
PGC		4	4	4	4	4	8	8	8	8	



Movie 1. Optimal transport of four lineages. To the left, a tetrahedral plot reports transport of ectoderm (blue), endoderm (yellow), NSM (orange), and PMCs (red) toward their respective vertices with the vertices reached by cells expressing cohort of genes for that lineage at 20hpf. To the right the same trajectories are reported on the UMAP. Gray cells are cells that have yet to reach 50% probability toward any of the fates.