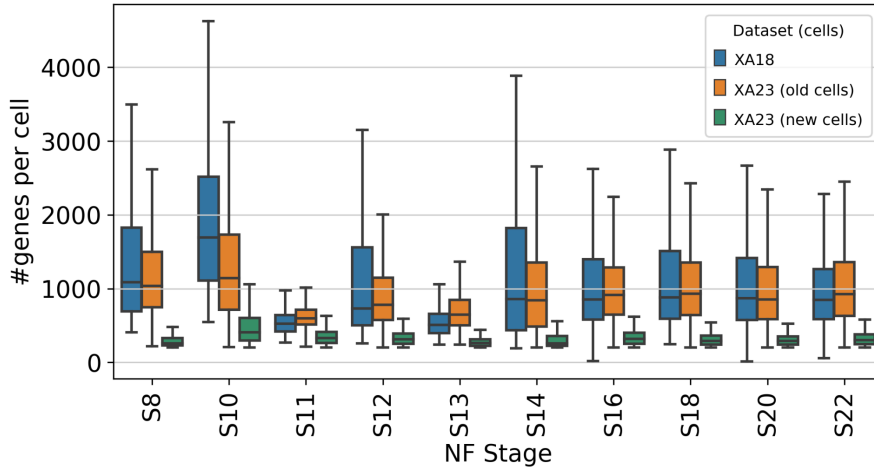
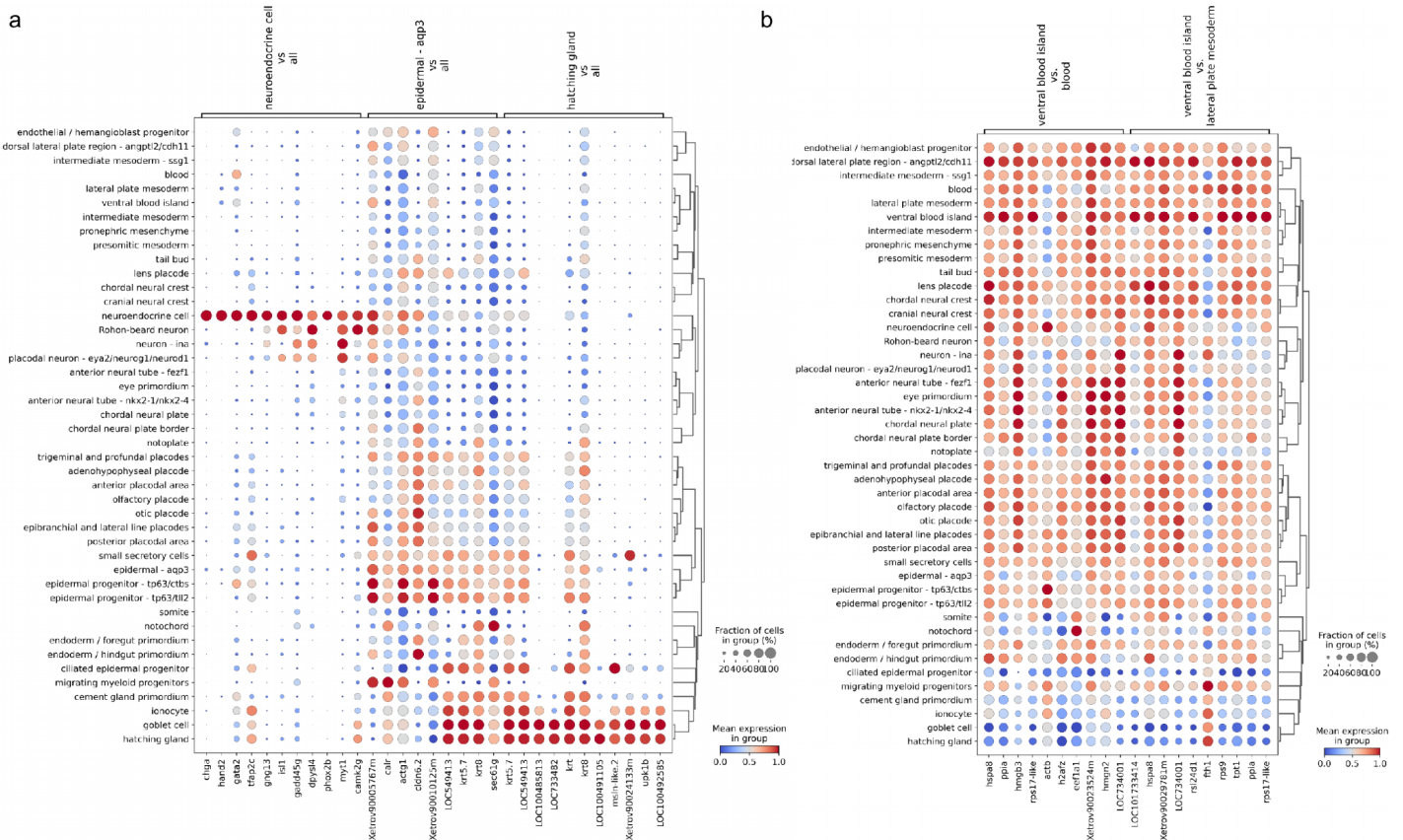


## Supplementary Figures

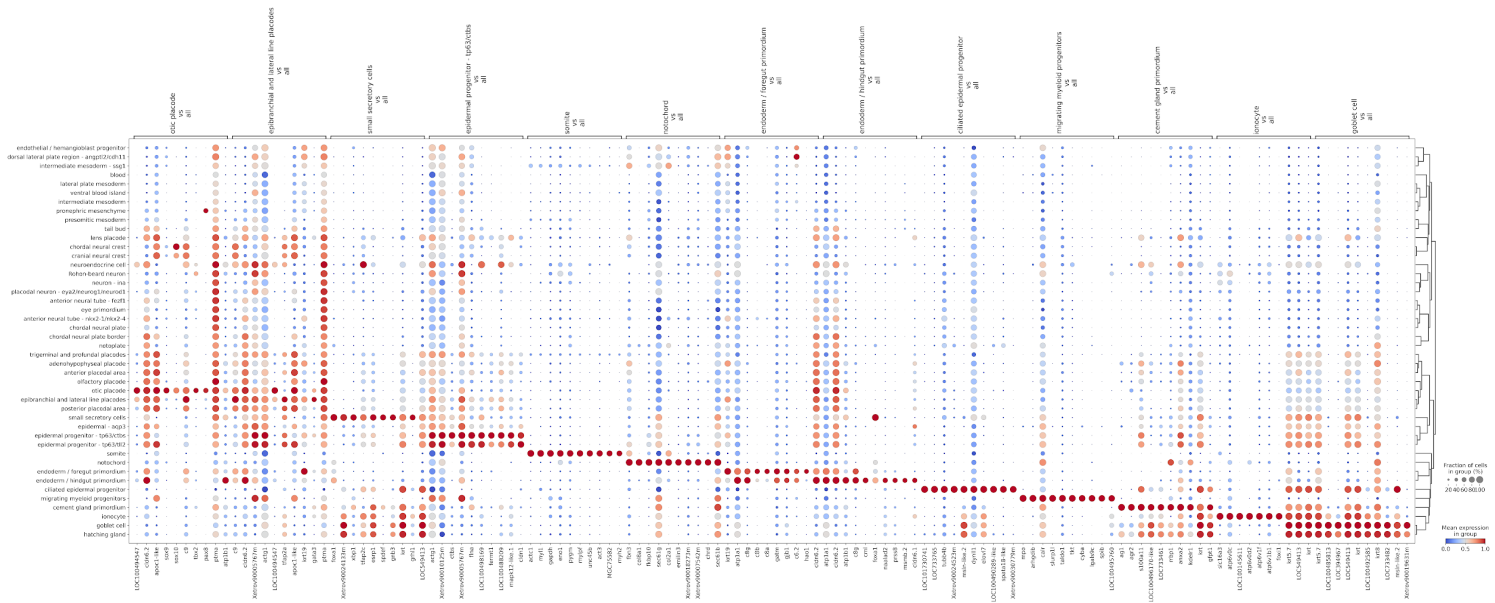


**Figure S1.** Summary statistics for the number of genes per cell per stage for the XA18 and XA23 datasets (old and new cells separately).

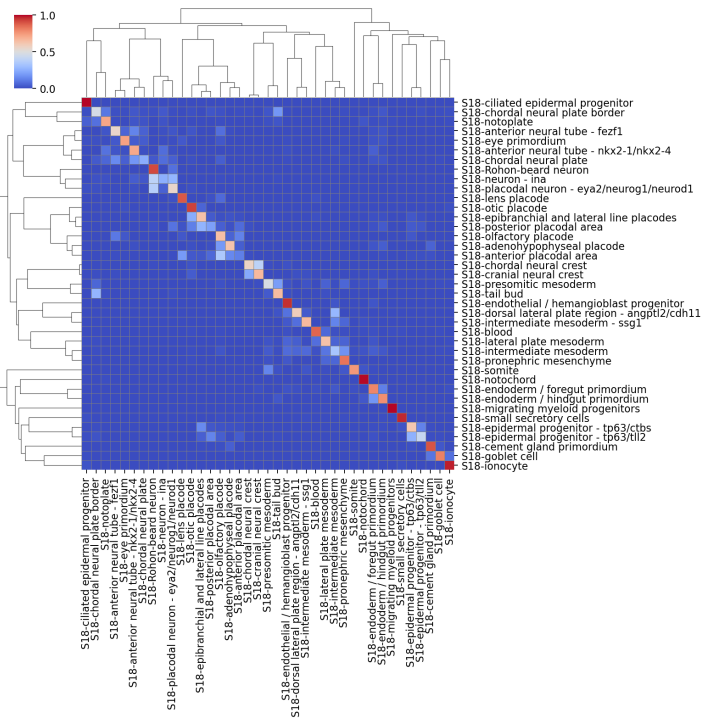


**Figure S2.** Dot plots for top 10 upregulated differentially expressed genes in cell types poorly predictable with mgB approach (a) and top 10 upregulated differentially expressed genes in "ventral blood island" against "blood" and "lateral plate mesoderm" cell types (b).

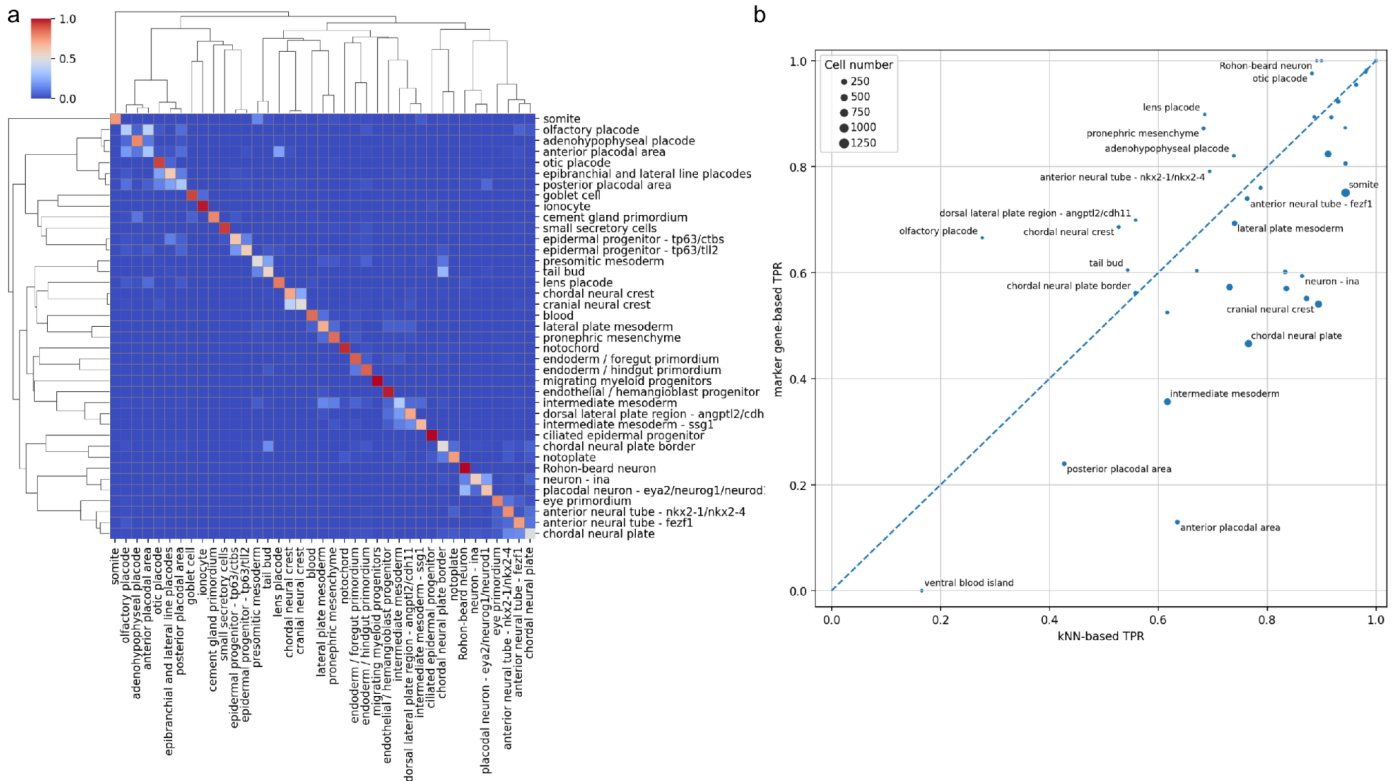




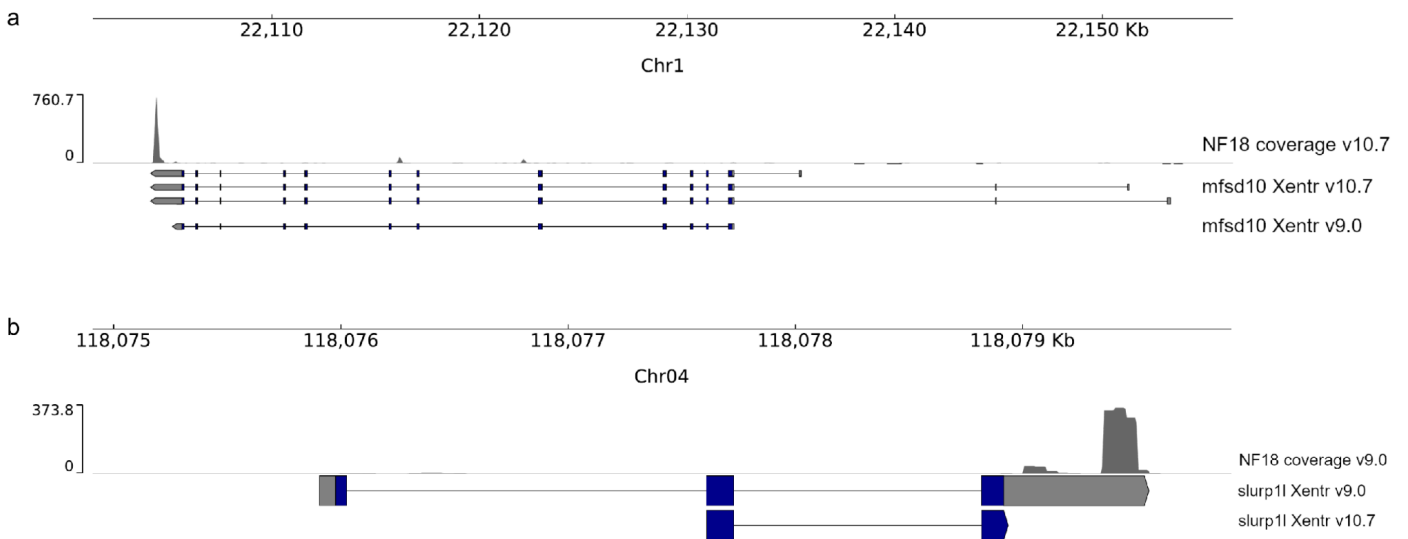
**Figure S3.** Dot plots of normalized counts of top 10 differentially expressed upregulated genes for “well predictable” cell types.



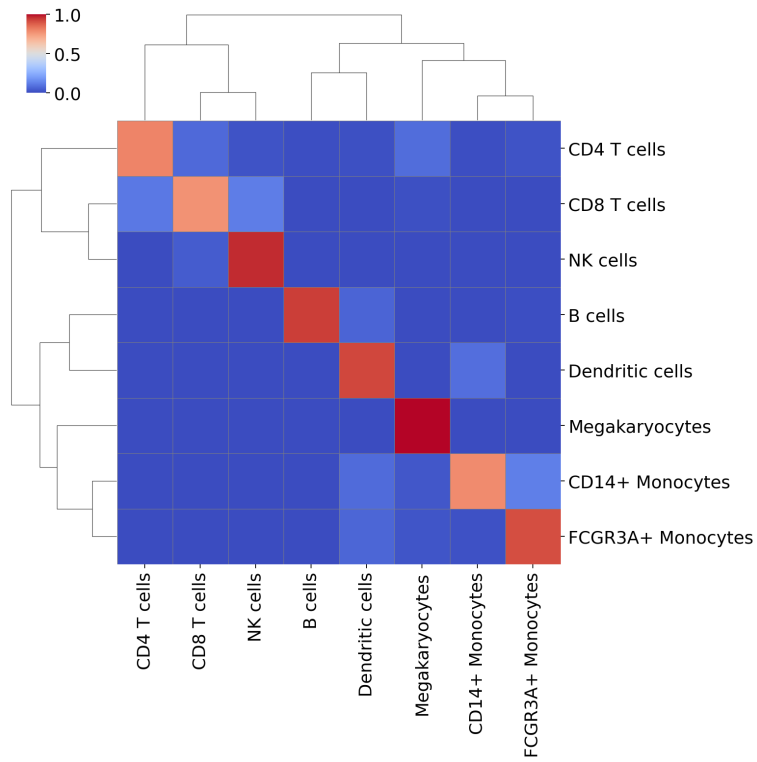
**Figure S4.** A confusion matrix for NF18 cell type prediction with mgB after discarding poorly predictable cell types.



**Figure S5.** Original cell types prediction accuracy in the new atlas, TPR for mgB vs. TPR for kNN **(a)** and mgB confusion matrix **(b)**.



**Figure S6.** Re-annotation of gene models across genome releases leads to variable results. **(a)** Bam coverage of one of NF18 samples aligned to transcriptome annotation v10.7 (“NF18 coverage v10.7”), *mfsd10* gene coordinates in genome annotation v10.7 (“mfsd10 Xentr v10.7”) and liftover to v10.7 of *mfsd10* gene coordinates in genome annotation v9.0 (“mfsd10 Xentr v9.0”). **(b)** Bam coverage of one of NF18 samples aligned to transcriptome annotation v9.0 (“NF18 coverage v9.0”), *slurp1l* gene coordinates in genome annotation v9.0 (“slurp1l Xentr v9.0”) and liftover to v9.0 of *slurp1l* gene coordinates in genome annotation v10.7 (“slurp1l Xentr v10.7”).



**Figure S7.** A confusion matrix for the 10x PBMC 3K dataset cell-type prediction via marker gene overrepresentation analysis.

Preliminary PDF with figures is [here](#).