

# High-resolution RNA isoform variation analysis by combining the advantages of single-cell RNA and nanopore sequencing

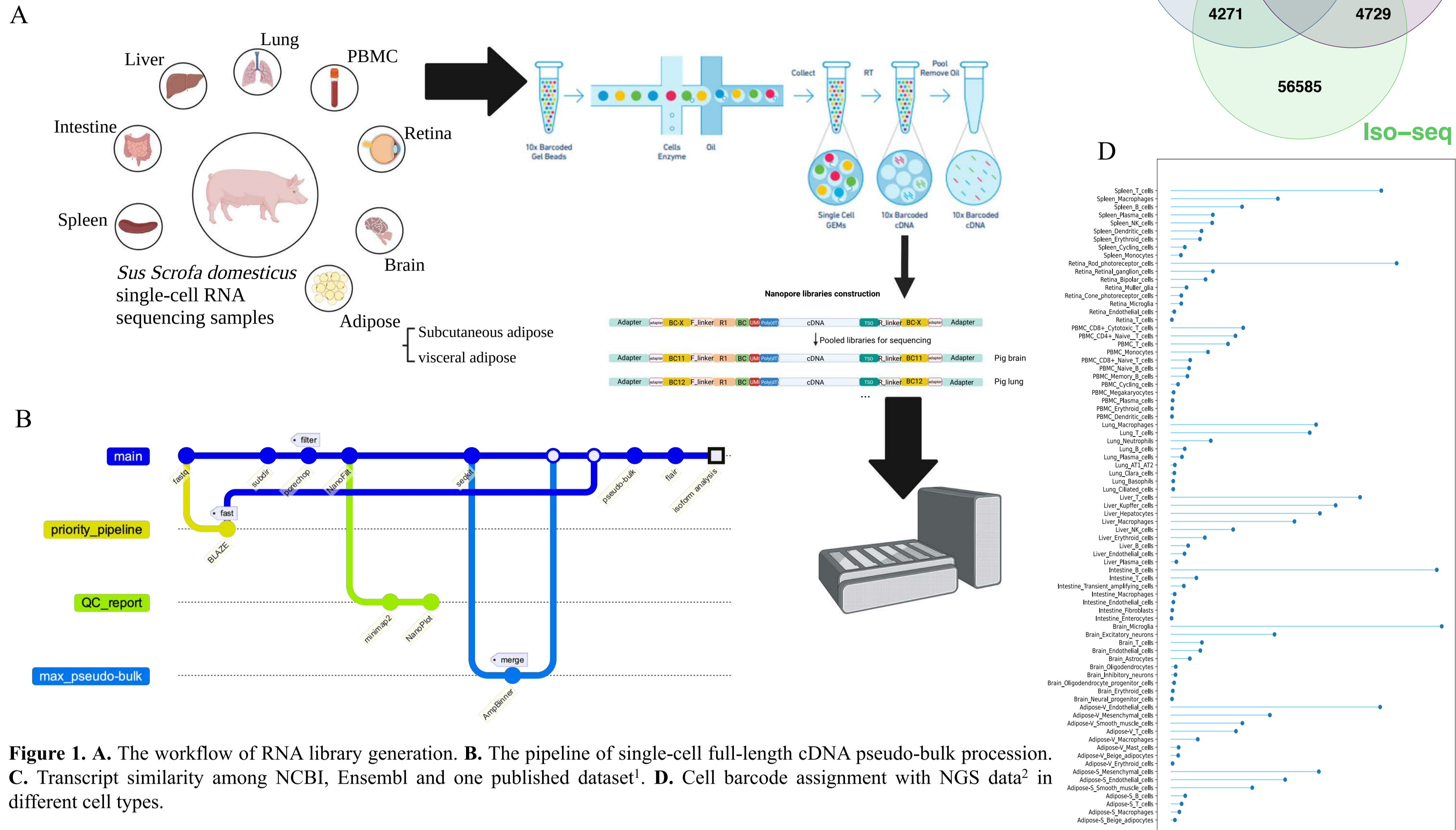
Xiaoyu Zhou<sup>1,3</sup>, Karim Rahimi<sup>2</sup>, Lin Lin<sup>1,3</sup>

1. Department of Biomedicine, AU;
2. Department of Molecular Biology and Genetics - RNA Biology and Innovation, AU;
3. Steno Diabetes Center Aarhus (SDCA), AUH.



## Introduction

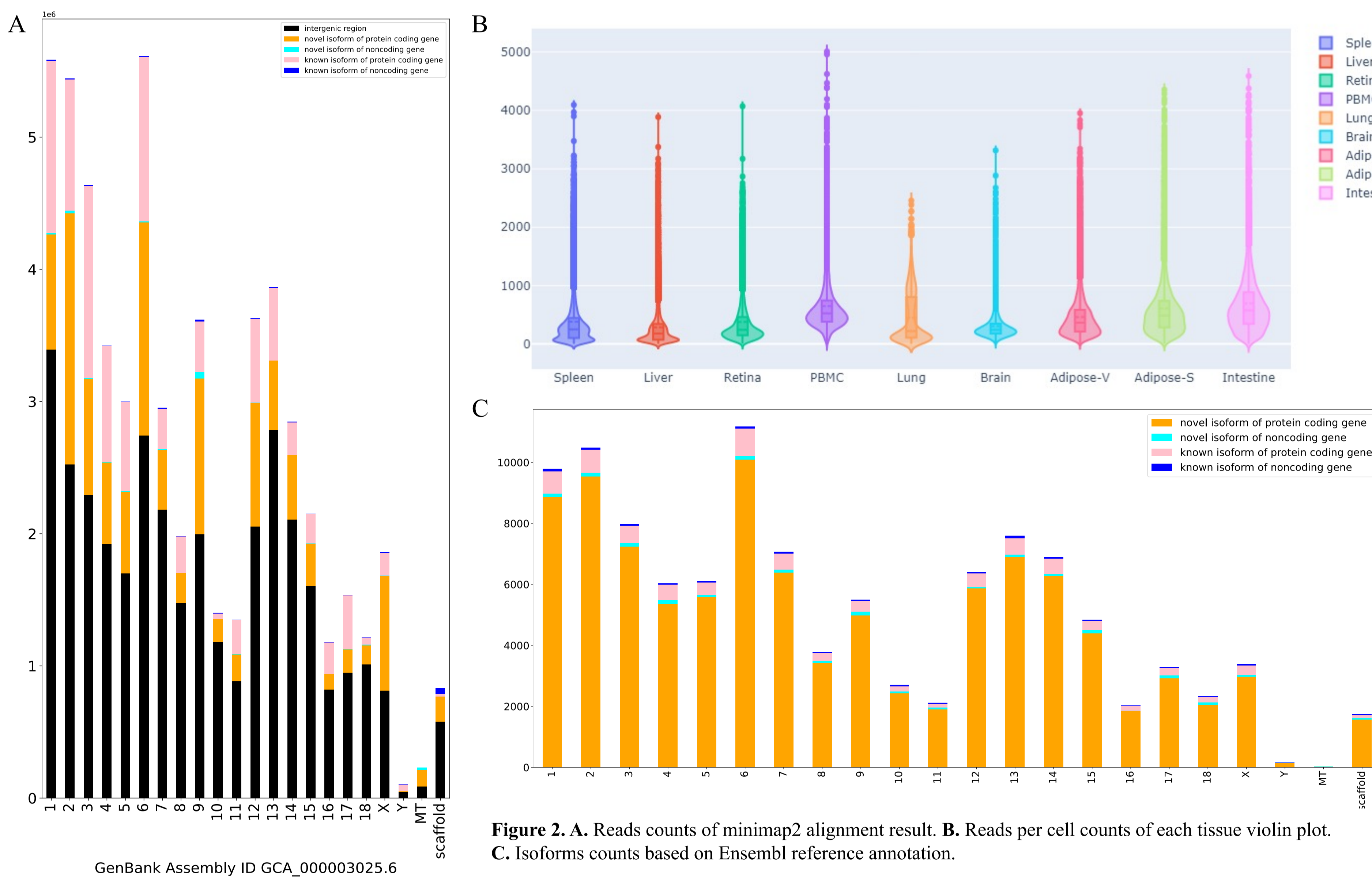
**Model Organism:** Pigs are one of the ideal models for both human diseases and xenotransplantation due to similar genetic backgrounds and organ sizes.  
**Limitation:** However, with the standard single-cell RNA sequencing pipeline, it is only possible to identify the expression level of each gene.



**Figure 1.** A. The workflow of RNA library generation. B. The pipeline of single-cell full-length cDNA pseudo-bulk processing. C. Transcript similarity among NCBI, Ensembl and one published dataset<sup>1</sup>. D. Cell barcode assignment with NGS data<sup>2</sup> in different cell types.

## Methods

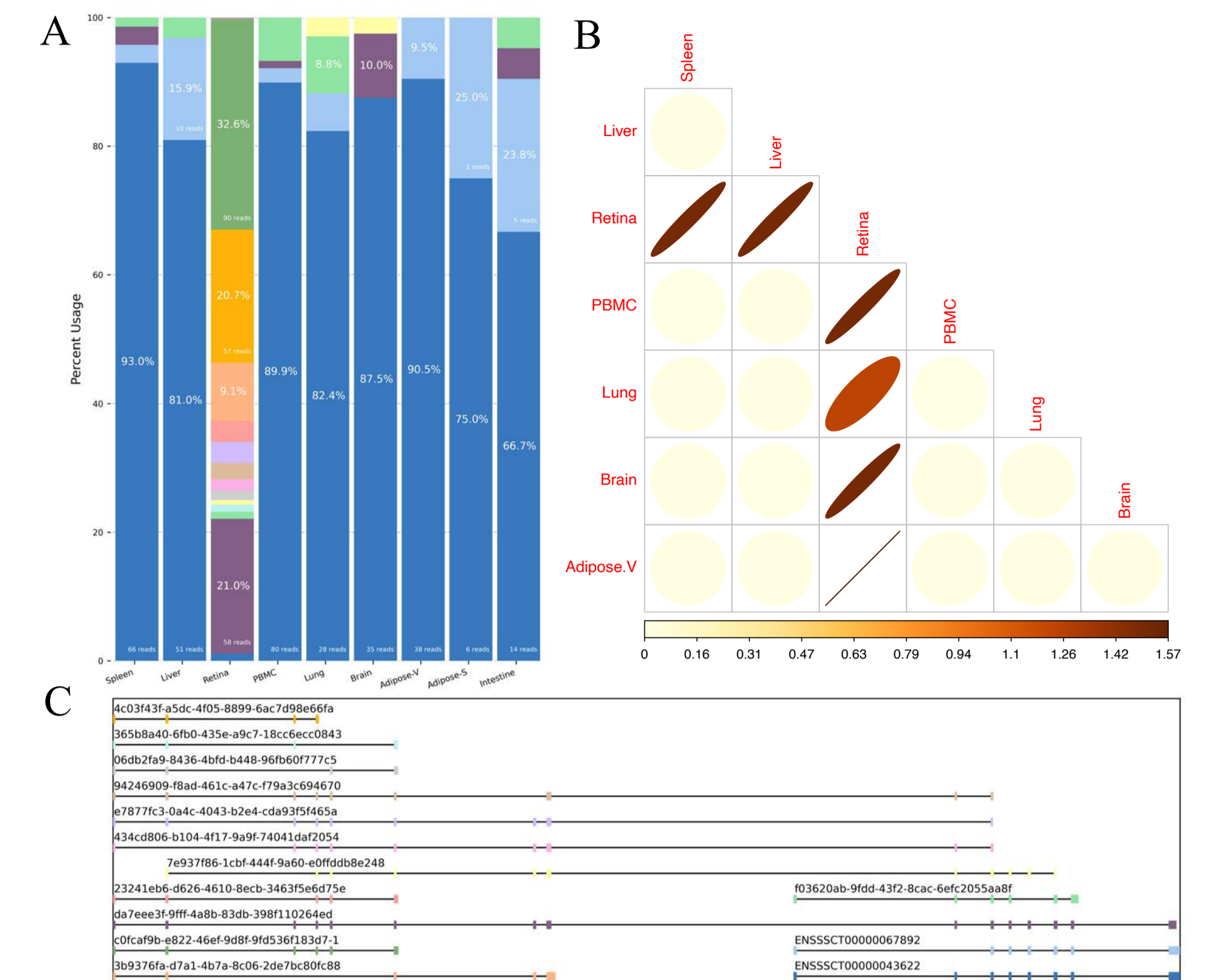
We have established an RNA isoform atlas of pigs to analyse the isoform heterogeneity. We aim to find the preferred dominant isoforms and reveal differential transcript usage (DTU) to detect splicing events at high resolution.



**Figure 2.** A. Reads counts of minimap2 alignment result. B. Reads per cell counts of each tissue violin plot. C. Isoforms counts based on Ensembl reference annotation.



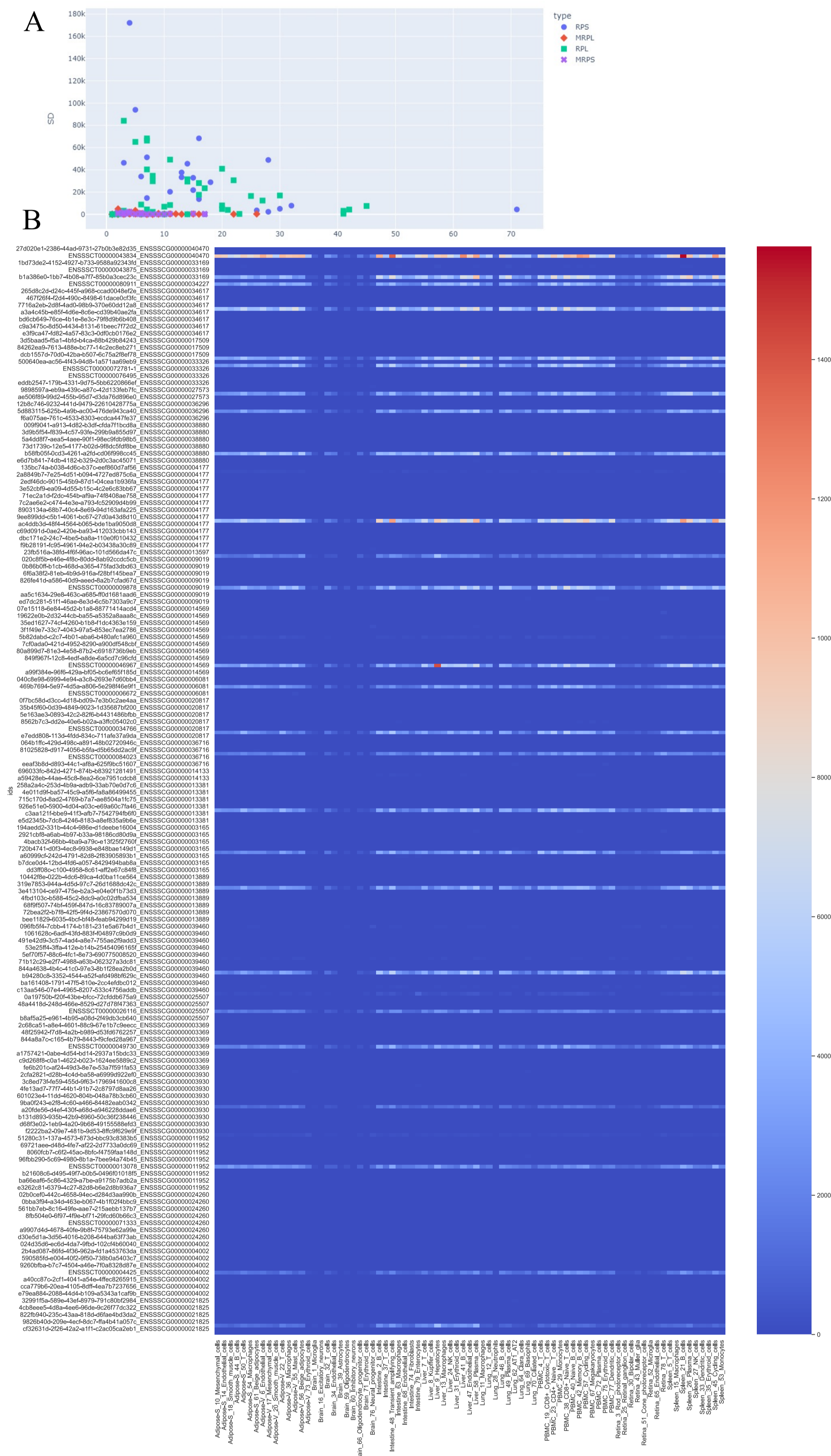
**Figure 3.** A. post-hoc test result of isoform percentage of different endothelial cells. B. FLAIR<sup>3</sup> pipeline isoform result of gene ENSSCG0000001229 (SLA-3). C. Differential transcript usage results of 79 cell types.



**Figure 4.** A. DTU results of 9 tissues. B. post-hoc test result of gene ENSSCG00000038062 isoform percentage of different tissue. C. FLAIR<sup>3</sup> pipeline isoform type result.

## Results

The more convincing isoforms of the target gene feature, the better for cell-type heterogeneity analysis. The NCBI database has more mRNA isoforms compared with Ensembl. According to the Friedman test result of the isoform percentage, we select P-value <0.05 genesets for heterogeneity analysis. The expression level of the selected housekeeping gene in Fig4 showed pig retina has significant DTU.



**Figure 5.** A. 138 pig ribosomal protein subunit genes were calculated by isoform quant and standard deviation (SD) of corresponding cumulative reads. B. Isoform TPM expression heatmap of the selected gene from Fig5A, high SD>20000.

## Conclusion

To explore the isoform usage variation in different cell types and tissues, we supply a newly integrated fast workflow and statistics scripts for establishing the standard operating protocol. The Friedman test and post-hoc test statistics results could help to filter the target genes. According to the isoform analysis on both tissue and cell type levels, we identified the novel isoforms which are dominant in specific groups. The heterogeneity of the MHC gene (SLA-3) isoform preferences were observed in endothelial cells across different tissue. Although the majority cells keep stable DTU of the nucleus and mitochondria-related RPS/RPL genesets, which represented the conserved genes, we could also detect different isoform expression levels, such as some expressed lower in brain and retina.

## References

1. Beiki H & al (2019). BMC Genomics 20: 344.
2. Wang F & al (2022). Nat Commun 13, 3620.
3. Tang A.D & al (2020). Nat Commun 11, 1438.