LETTER TO THE EDITOR

Unraveling the stereoscopic gene transcriptional landscape of zebrafish using FishSED, a fish spatial expression database with multispecies scalability

Cheng Guo[1, 2†], Weidong Ye[1, 2†], Danying Cao[1, 2], Mijuan Shi[1, 3, *], Wanting Zhang[1,3], Yingyin Cheng[1,3], Yaping Wang[1, 2, 3], Xiao-Qin Xia[1, 2, 3, *]

SUPPORTING INFORMATION

Figure S1. The proportions of datasets in relation to tissue types and sequencing technologies, the home page of FishSED, and the comparison of the datasets from five databases.

Figure S2. Implementation of FishSED.

Figure S3. Browse page of FishSED.

Figure S4. The advanced search page.

Figure S5. BLAST page of FishSED.

Figure S6. Project detailed page of GSE159709.

Figure S7. Dataset detailed page of GSE159709 (Visium-sample C).

Figure S8. Gene detailed page of *camk2g1*.

MATERIALS AND METHODS

Data collation

From three online platforms, CNGB, GEO, and Cardiovascular Research, we manually compiled zebrafish spatial transcriptome-related articles published before December 2022 and retrieved 56 datasets covered by 10 projects (Figure S2A, S2B). These datasets use 5 different sequencing technologies, including 10x Genomics Visium (Hunter et al., 2021), Stereo-seq (Liu et al., 2022), ST (Baron et al., 2020),Geo-seq (Xue et al., 2019), and Tomo-seq (Burkhard and Bakkers, 2018; Derrick et al., 2022; Holler et al., 2021; Junker et al., 2014; Wu et al., 2016; Yvernogeau et al., 2020). Based on the type of experiment and sample status from the published articles, we divided the datasets into three biological study directions, including baseline, cancer, and regeneration.

Database construction

Database architecture

The Python programming language was used to create the FishSED website, which is hosted on a CentOS server and built on the Django web framework (Figure S2C). The data is stored in a MySQL relational database. The Bootstrap framework, which offers dynamic page layouts dependent on the resolution of the user's monitor, serves as the foundation for the database front-end. Highcharts, echarts, amcharts, and plotly packages are used to develop data visualization and produce interactive charts that make the user experience more pleasant.

Database function realization

The text search function is implemented using Django's ORM model, while comprehensive search, advanced search of gene/project/dataset, and so on are implemented by in-house Python scripts. The homology search function for sequences relies on BLAST software. JBrowse2 implements the visualization of results based on genomic location and gene structure. The expression profile visualization is generated dynamically using in-house JavaScript scripts. The data download service is built by Django static file system.

Data Availability

All data used in this study were derived from public sources mentioned in the method section. All generated data and analyzed results were uploaded to FishSED (http://bioinfo.ihb.ac.cn/fishsed) with corresponding download options. Related script files are available upon reasonable requests.

Supplemental References

Baron, M., Tagore, M., Hunter, M.V., Kim, I.S., Moncada, R., Yan, Y., Campbell, N.R., White, R.M., and Yanai, I. (2020). The stress-like cancer cell state is a consistent component of tumorigenesis. Cell Syst 11, 536-546 e537.

Burkhard, S.B., and Bakkers, J. (2018). Spatially resolved RNA-sequencing of the embryonic heart identifies a role for Wnt/beta-catenin signaling in autonomic control of heart rate. Elife 7.

Derrick, C.J., Sanchez-Posada, J., Hussein, F., Tessadori, F., Pollitt, E.J.G., Savage, A.M., Wilkinson, R.N., Chico, T.J., van Eeden, F.J., Bakkers, J., and Noel, E.S. (2022). Asymmetric Hapln1a drives regionalized cardiac ECM expansion and promotes heart morphogenesis in zebrafish development.

Cardiovasc Res 118, 226-240.

Holler, K., Neuschulz, A., Drewe-Boss, P., Mintcheva, J., Spanjaard, B., Arsie, R., Ohler, U., Landthaler, M., and Junker, J.P. (2021). Spatio-temporal mRNA tracking in the early zebrafish embryo. Nat Commun 12, 3358.

Hunter, M.V., Moncada, R., Weiss, J.M., Yanai, I., and White, R.M. (2021). Spatially resolved transcriptomics reveals the architecture of the tumor-microenvironment interface. Nat Commun 12, 6278.

Junker, J.P., Noel, E.S., Guryev, V., Peterson, K.A., Shah, G., Huisken, J., McMahon, A.P., Berezikov, E., Bakkers, J., and van Oudenaarden, A. (2014). Genome-wide RNA Tomography in the zebrafish embryo. Cell 159, 662-675.

Liu, C., Li, R., Li, Y., Lin, X., Zhao, K., Liu, Q., Wang, S., Yang, X., Shi, X., Ma, Y., Pei, C., Wang, H., Bao, W., Hui, J., Yang, T., Xu, Z., Lai, T., Berberoglu, M.A., Sahu, S.K., Esteban, M.A., Ma, K., Fan, G., Li, Y., Liu, S., Chen, A., Xu, X., Dong, Z., and Liu, L. (2022). Spatiotemporal mapping of gene expression landscapes and developmental trajectories during zebrafish embryogenesis. Dev Cell 57, 1284-1298 e1285.

Wu, C.C., Kruse, F., Vasudevarao, M.D., Junker, J.P., Zebrowski, D.C., Fischer, K., Noel, E.S., Grun, D., Berezikov, E., Engel, F.B., van Oudenaarden, A., Weidinger, G., and Bakkers, J. (2016). Spatially resolved genome-wide transcriptional profiling identifies BMP signaling as essential regulator of zebrafish cardiomyocyte regeneration. Dev Cell 36, 36-49.

Xue, Y., Liu, D., Cui, G., Ding, Y., Ai, D., Gao, S., Zhang, Y., Suo, S., Wang, X., Lv, P., Zhou, C., Li, Y., Chen, X., Peng, G., Jing, N., Han, J.J., and Liu, F. (2019). A 3D atlas of hematopoietic stem and progenitor cell expansion by multi-dimensional RNA-seq analysis. Cell Rep 27, 1567-1578 e1565.

Yvernogeau, L., Klaus, A., Maas, J., Morin-Poulard, I., Weijts, B., Schulte-Merker, S., Berezikov, E., Junker, J.P., and Robin, C. (2020). Multispecies RNA tomography reveals regulators of hematopoietic stem cell birth in the embryonic aorta. Blood 136, 831-844.

Supplemental Figures

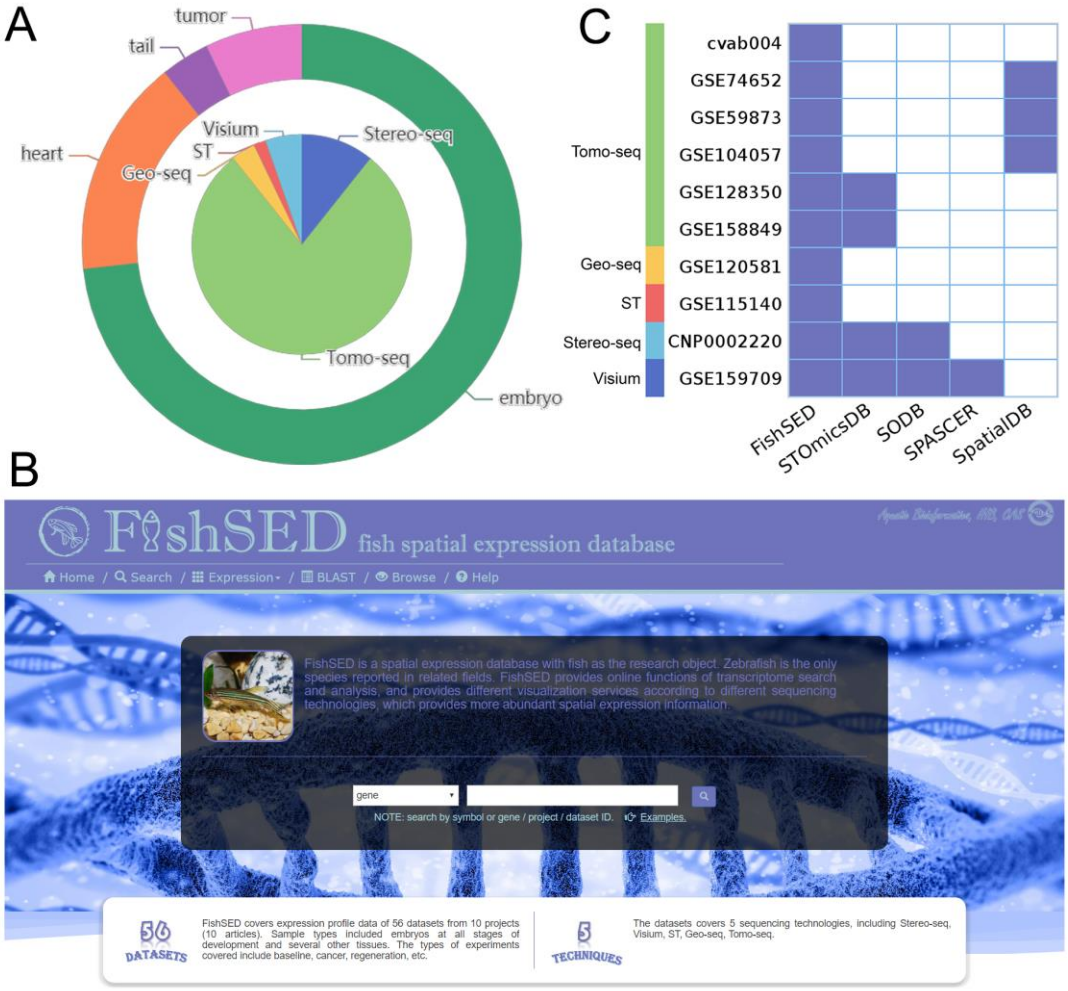


Figure S1. A, The proportions of datasets in relation to tissue types and sequencing technologies. B, The home page of FishSED. C, Comparison of the datasets from five databases.

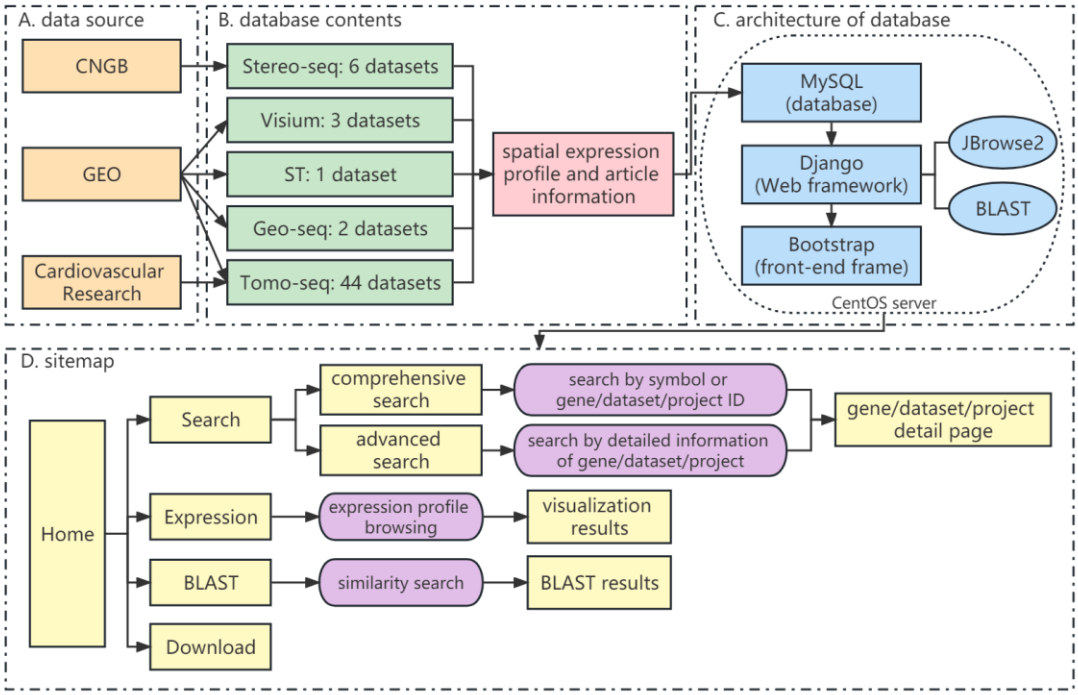


Figure S2. Implementation of FishSED. A, Data sources. B, Database contents. C, Architecture of FishSED. D, Sitemap of FishSED.

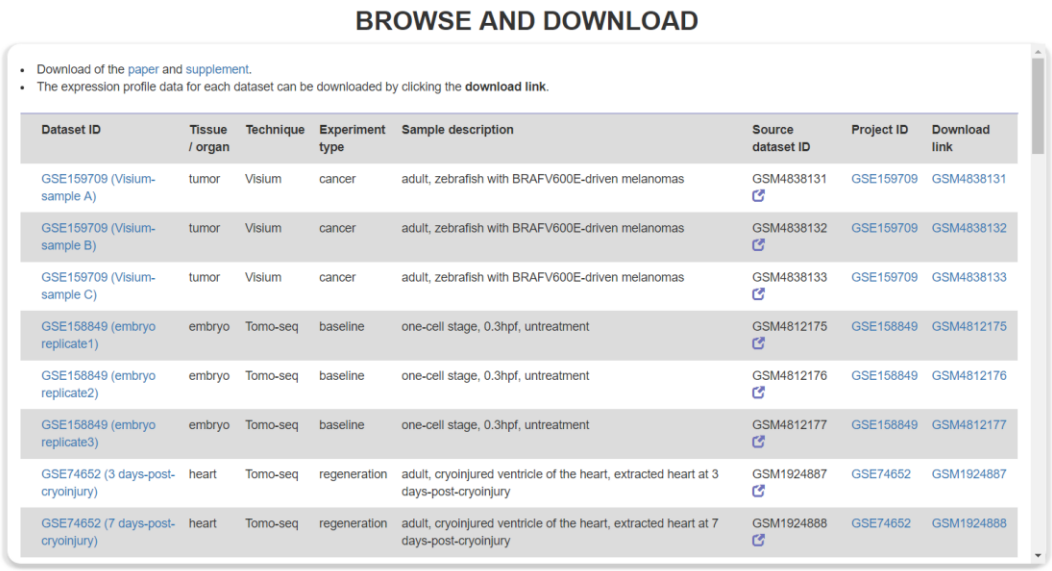


Figure S3. Browse page of FishSED.

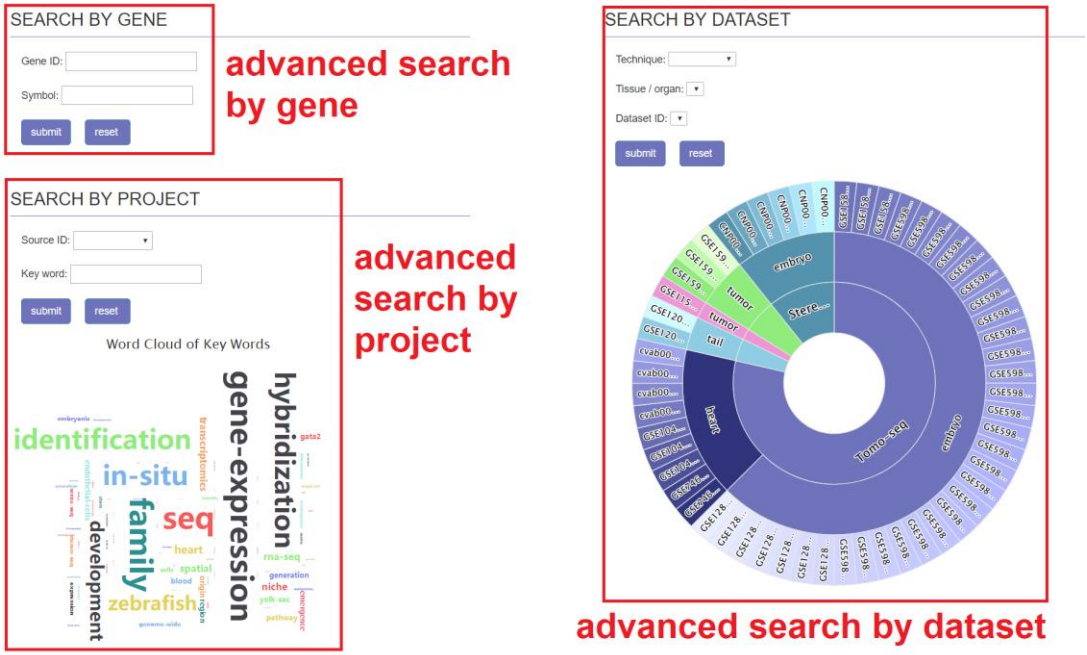


Figure S4. The advanced search page.

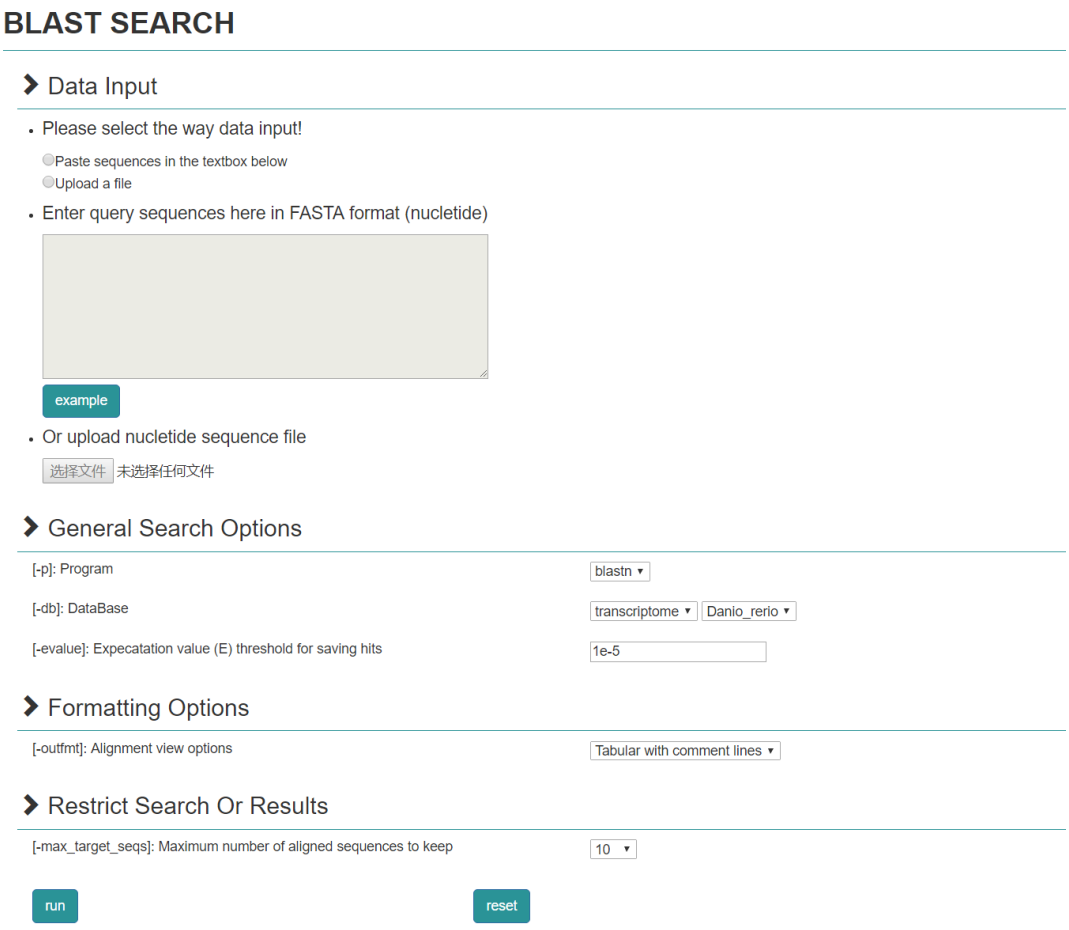
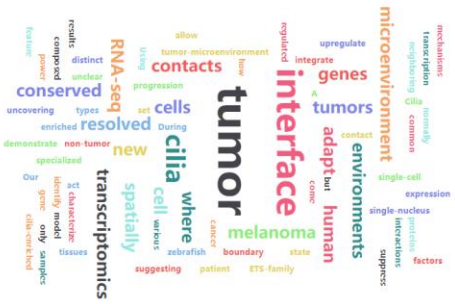


Figure S5. BLAST page of FishSED.

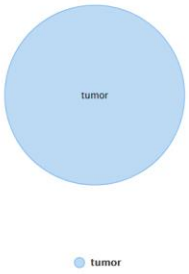
PROJECT ID: GSE159709

Data source	GEO: GSE159709
Description	Cancer cells interact with a wide variety of other cell types, but our understanding of microenvironmental heterogeneity and how it influences tumor phenotypes is limited. While single-cell RNA-seq (scRNA-seq) has helped define these TME cell types, it provides limited information on the mechanisms that define how individual tumor cells interact with TME. Here, we integrate spatial transcriptomics with scRNA-seq to define the architecture and nature of nascent tumor and surrounding microenvironment cells as they come into contact through the process of invasion. Using a well-defined transgenic zebrafish model of BRAFV600E-driven melanoma, we identify a transcriptionally unique â€œinterfaceâ€□ cluster localized at the boundary between tumor cells and surrounding tissues. Using an unbiased, data-driven approach, we identify spatially-patterned gene modules specific to the interface and show that the interface is a distinct transcriptional entity that histologically resembles the microenvironment but transcriptionally resembles the tumor. By complementing ST with scRNA-seq, we demonstrate that the interface is composed of specialized tumor and microenvironment cells. Both cell types in the interface upregulate a common set of cilia genes, and we find enrichment of cilia proteins only where the tumor meets the TME. Cilia gene expression is regulated by ETS-family transcription factors, which normally act to suppress their expression outside of this region. This unique ETS-driven interface transcriptional state is conserved across ten different human patient samples, suggesting this is a conserved feature of human melanoma. Taken together, our results demonstrate the power of spatial and single-cell transcriptomics techniques in uncovering novel biological mechanisms that drive tumor invasion into new tissues.
Key word	genome-wide expression; nucleus rna-seq; gene-expression; primary cilium; single-cell; metastatic melanoma; adult zebrafish; ets family; heterogeneity; progression
Publication	Hunter MV, Moncada R, Weiss JM, Yanai I et al. Spatially resolved transcriptomics reveals the architecture of the tumor-microenvironment interface. Nat Commun 2021 Nov 1;12(1):6278. PMID: 34725363
Abstract	During tumor progression, cancer cells come into contact with various non-tumor cell types, but it is unclear how tumors adapt to these new environments. Here, we integrate spatially resolved transcriptomics, single-cell RNA-seq, and single-nucleus RNA-seq to characterize tumor-microenvironment interactions at the tumor boundary. Using a zebrafish model of melanoma, we identify a distinct "interface" cell state where the tumor contacts neighboring tissues. This interface is composed of specialized tumor and microenvironment cells that upregulate a common set of cilia genes, and cilia proteins are enriched only where the tumor contacts the microenvironment. Cilia gene expression is regulated by ETS-family transcription factors, which normally act to suppress cilia genes outside of the interface. A cilia-enriched interface is conserved in human patient samples, suggesting it is a conserved feature of human melanoma. Our results demonstrate the power of spatially resolved transcriptomics in uncovering mechanisms that

Word Cloud of Abstract



Distribution of Tissue/Organ in the Project



Dataset Information

	Dataset ID	Tissue / Organ	Technique	Experiment type	Sample description	Source dataset ID
1.	GSE159709 (Visium-sample A)	tumor	Visium	cancer	adult, zebrafish with BRAFV600E-driven melanomas	GEO: GSM4838131
2.	GSE159709 (Visium-sample B)	tumor	Visium	cancer	adult, zebrafish with BRAFV600E-driven melanomas	GEO: GSM4838132
3.	GSE159709 (Visium-sample C)	tumor	Visium	cancer	adult, zebrafish with BRAFV600E-driven melanomas	GEO: GSM4838133

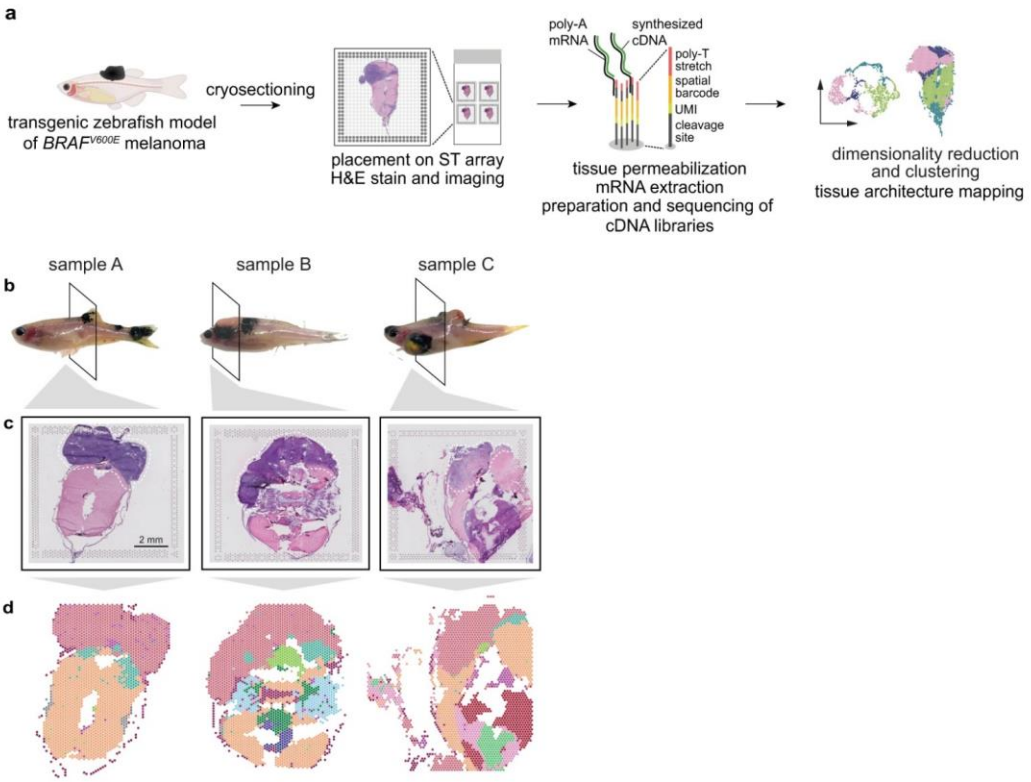


Figure S6. Project detailed page of GSE159709.

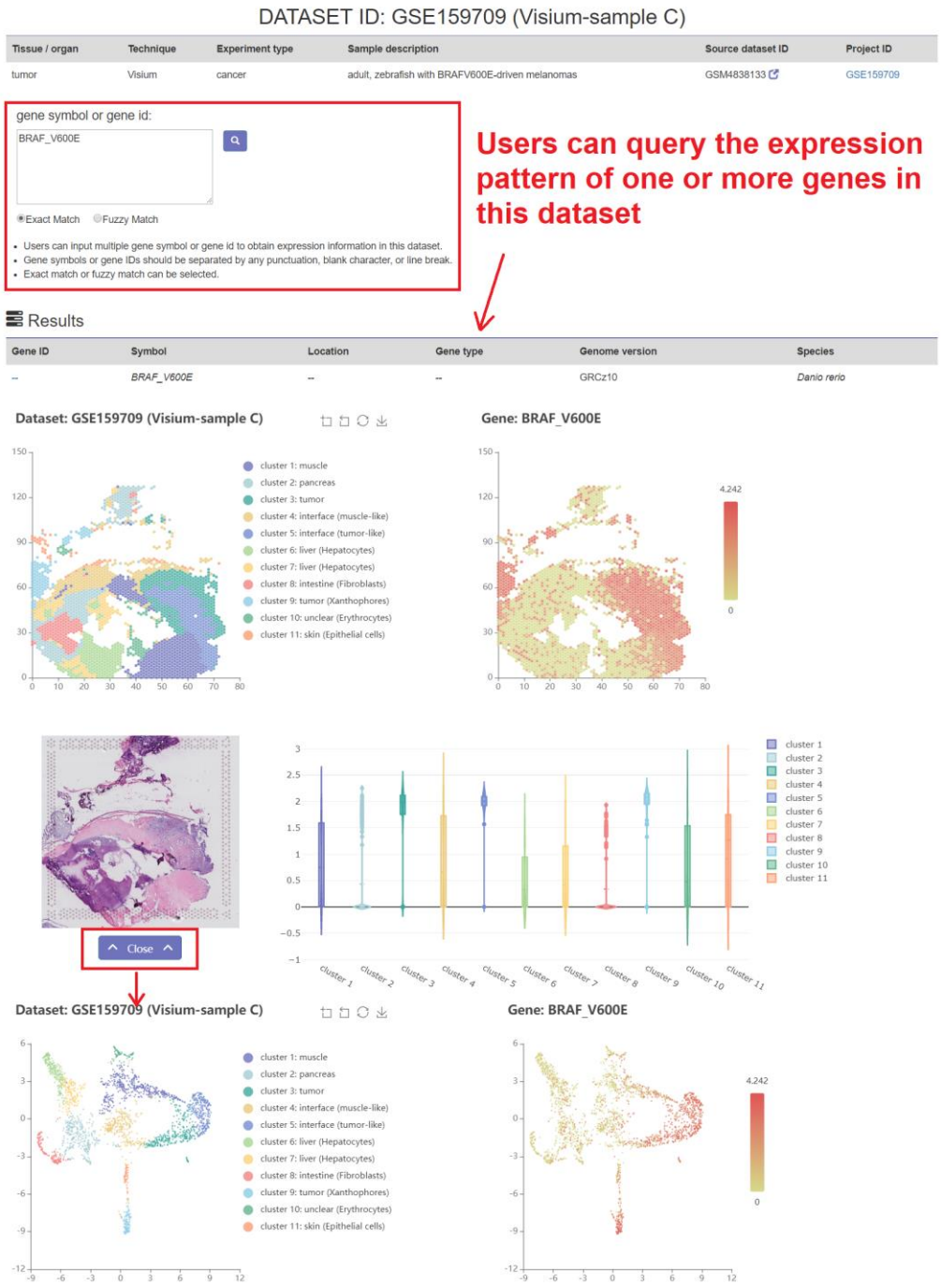


Figure S7. Dataset detailed page of GSE159709 (Visium-sample C).

Figure S8. Gene detailed page of *camk2g1*.