

调控小胶质细胞代谢的关键基因 XXX

2024-05-17

LiChuang Huang



@ 立效研究院

Contents

1 摘要	1
1.1 需求	1
1.2 结果	1
2 前言	1
3 材料和方法	1
3.1 材料	1
3.2 方法	1
4 分析结果	2
5 结论	2
6 附：分析流程	2
6.1 单细胞数据分析	2
6.1.1 数据来源	2
6.1.2 细胞聚类与鉴定	3
6.1.3 小胶质细胞分析	4
6.1.3.1 差异分析	4
6.1.3.2 拟时分析	4
6.1.3.3 代谢通量预测	6
6.1.3.4 代谢通量差异分析	7
6.1.3.5 差异代谢相关的基因	9
6.1.3.6 PPI 网络	10
6.1.3.7 上游的转录因子	12
6.1.4 星形细胞胶质瘢痕相关基因	13
6.1.4.1 差异分析	13
6.1.4.2 胶质瘢痕	14
6.1.5 调控小胶质细胞代谢且与星形细胞胶质瘢痕相关基因	16
6.1.5.1 交集基因	16
6.1.5.2 在小胶质细胞中的表达	17
Reference	18

List of Figures

1 UMAP Clustering	3
2 The cellType group	3
3 MI principal points	5
4 MI pseudotime	6

5	Convergency of the loss terms during training	7
6	SCF Model vs Control	8
7	TOPFLUX GO enrichment	9
8	MI Pseudotime heatmap of genes	10
9	TOPFLUX MCC score	11
10	TOPFLUX Top MCC score	12
11	Intersection of DB GlialScar with Astrocyte DEGs	16
12	UpSet plot of Genes sources	17
13	Dimension plot of expression level of the genes	18

List of Tables

1	DEGs of the contrasts Microglial	4
2	SCF data Model vs Control	7
3	Transcription Factor binding sites	12
4	DEGs of the contrasts Astrocyte	13
5	Mapped genes Astrocyte DEGs	14
6	GLIALSCAR disease related targets from GeneCards	15
7	Xylt1 related metabolic flux	18

1 摘要

1.1 需求

- 自发性脑出血 (ICH) (人或动物) 后调控小胶质细胞代谢的关键基因 XXX，并且该基因可能也调控星形细胞胶质瘢痕生成。
- 客户前期发表的文章有做过盐诱导激酶 2 (SIK-2) (PMID:29018127)，XXX 是否可以是 SIK-2，或者 XXX 能否富集在 SIK-2 相关通路或其他分子机制上

1.2 结果

- 以小鼠丘脑出血模型 (GSE227033) 的单细胞数据集分析 Microglial cell, 见 Fig. 2。(该模型大体上应该是合适的, 以 collagenase IV 造模 (像 PMID:38433011 也是这种 ICH 造模)。ICH 的单细胞数据很少, 基本没有其他合适的数据了)
- 以 Microglial Cell 预测代谢通量, 并差异分析, 见 Fig. 6
- 与差异代谢相关的基因, 在拟时轨迹 (Control -> Model) 中的表达见 Fig. 8
- 将这些基因映射到人类的基因 (hgnc symbol) 后, 获取上游的转录因子。在这些基因和转录因子中, 尝试寻找 SIK-2。无结果。此外, SIK-2 为非差异表达基因 (Model vs Control)。
- 为筛选 Astrocyte 胶质瘢痕相关基因, 首先获取了 Astrocyte 的差异基因 (Model vs Control), 见 Tab. 4。随后, 获取 GeneCards 的胶质瘢痕相关基因, 见 Tab. 6。尝试取交集, 见 Fig. 11, 有 11 个基因。
- 最后, 调控小胶质细胞代谢且与星形细胞胶质瘢痕相关基因, Fig. 12, 由于 Fig. 11 的基因与 Microglial 代谢相关基因及上游转录因子无交集, 因此, 筛选时直接用了 Tab. 6 的基因。获得结果: XYLT1 (即, Xylt1)。
- Xylt1 的表达见 Fig. 13。Xylt1 主要集中表达于模型组的 Microglial 中, 符合条件。Xylt1 相关代谢通路见 Tab. 7。Xylt1 的其余信息可参考 Fig. 8。

2 前言

3 材料和方法

3.1 材料

All used GEO expression data and their design:

- **GSE227033:** we sequenced the transcriptomes of 32332 single brain cells, revealing a total of four major cell types within the four thalamus sample from mice.

3.2 方法

Mainly used method:

- R package `biomaRt` used for gene annotation¹.
- The `biomart` was used for mapping genes between organism (e.g., `mgi_symbol` to `hgnc_symbol`)¹.
- The `scFEA` (python) was used to estimate cell-wise metabolic via single cell RNA-seq data².

- R package `ClusterProfiler` used for gene enrichment analysis³.
- GEO <https://www.ncbi.nlm.nih.gov/geo/> used for expression dataset acquisition.
- The Human Gene Database `GeneCards` used for disease related genes prediction⁴.
- R package `Limma` and `edgeR` used for differential expression analysis^{5,6}.
- R package `Monocle3` used for cell pseudotime analysis^{7,8}.
- R package `STEINGdb` used for PPI network construction^{9,10}.
- The R package `Seurat` used for scRNA-seq processing^{11,12}.
- The Transcription Factor Target Gene Database (<https://tfbsdb.systemsbiology.net/>) was used for discovering relationship between transcription factors and genes..¹³
- SCSA (python) used for cell type annotation¹⁴.
- The MCC score was calculated referring to algorithm of CytoHubba¹⁰.
- R version 4.4.0 (2024-04-24); Other R packages (eg., `dplyr` and `ggplot2`) used for statistic analysis or data visualization.

4 分析结果

5 结论

6 附：分析流程

6.1 单细胞数据分析

6.1.1 数据来源

Data Source ID :

GSE227033

data_processing :

Postprocessing and quality control were performed using a 10× Cell Ranger package (v1.2.0; 10 × Genomics). Reads were aligned to the mm10 reference assembly (v1.2.0; 10 × Genomics). sn-RNA seq data (Cellranger_result) contains 4 samples (C1, C2, M1, M2).

data_processing.1 :

Assembly: mm10

data_processing.2 :

Supplementary files format and content: Tab-separated values files and matrix files

(上述信息框内容已保存至 Figure+Table/GSE227033-content)

6.1.2 细胞聚类与鉴定

Figure 1 (下方图) 为图 UMAP Clustering 概览。

(对应文件为 Figure+Table/UMAP-Clustering.pdf)

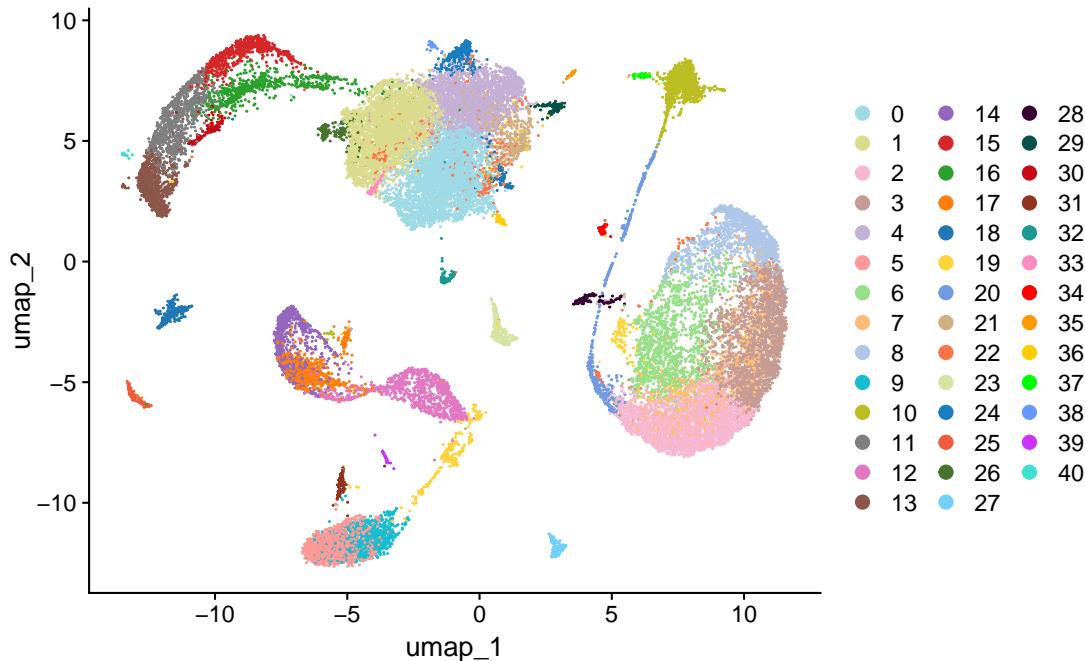


Figure 1: UMAP Clustering

Figure 2 (下方图) 为图 The cellType group 概览。

(对应文件为 Figure+Table/The-cellType-group.pdf)

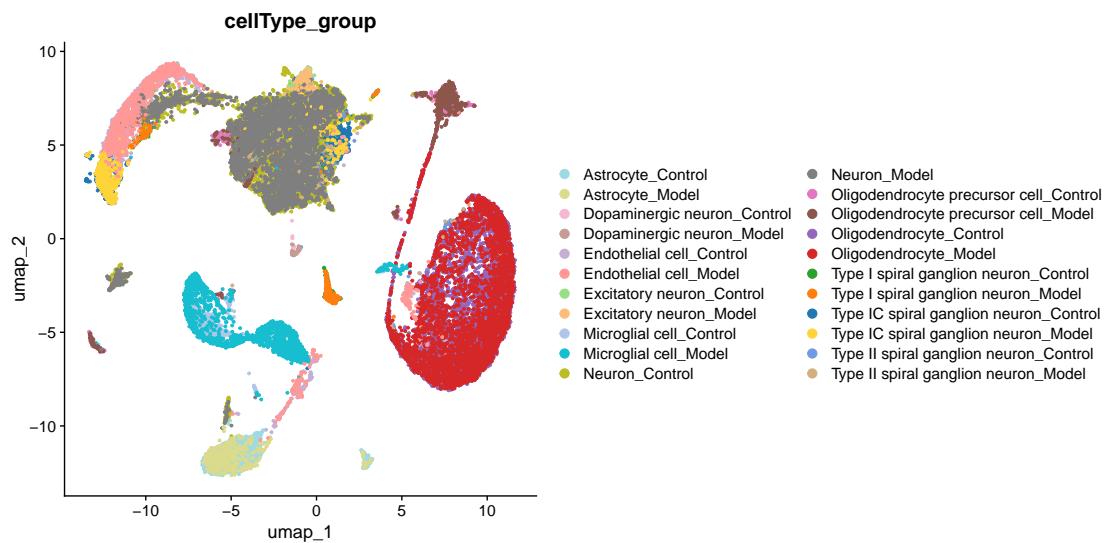


Figure 2: The cellType group

6.1.3 小胶质细胞分析

6.1.3.1 差异分析

Table 1 (下方表格) 为表格 DEGs of the contrasts Microglial 概览。

(对应文件为 [Figure+Table/DEGs-of-the-contrasts-Microglial.csv](#))

注：表格共有 1546 行 7 列，以下预览的表格可能省略部分数据；含有 1 个唯一 ‘contrast’。

Table 1: DEGs of the contrasts Microglial

contrast	p_val	avg_log2FC	pct.1	pct.2	p_val_adj	gene
Microglial...	1.83518861...	2.92401237...	0.09	0.36	5.50556584...	Cenpe
Microglial...	4.48315517...	9.81893833...	0.111	0.374	1.34494655...	Cdkn1a
Microglial...	1.73323788...	-4.0813517...	0.013	0.103	5.19971364...	Qrfpr
Microglial...	9.12205066...	-2.4590397...	0.103	0.217	2.73661519...	Dock6
Microglial...	2.54749154...	0.28706360...	0.077	0.288	7.64247464...	Gm26870
Microglial...	2.93243498...	5.68612318...	0.269	0.059	8.79730495...	Pantr2
Microglial...	8.30783340...	-6.9973674...	0.91	0.207	2.49235002...	Agt
Microglial...	4.69847504...	2.60879746...	0.612	0.092	1.40954251...	Penk
Microglial...	1.07440758...	-1.7746631...	0.104	0.087	3.22322274...	D7Ertd443e
Microglial...	1.67179273...	8.88979035...	0.136	0.499	5.01537820...	Ndc80
Microglial...	6.93761245...	5.36162953...	0.125	0.34	2.08128373...	Tmem123
Microglial...	1.41432503...	12.6809064...	0.124	0.184	4.24297509...	Egln3
Microglial...	1.97677667...	2.11642957...	0.334	0.072	5.93033003...	Car14
Microglial...	1.05458751...	9.33084014...	0.089	0.178	3.16376254...	Gadd45b
Microglial...	3.04118215...	1.46565262...	0.126	0.277	9.12354645...	Parp3
...

6.1.3.2 拟时分析

选择 Control 集中区域作为拟时起点。

Figure 3 (下方图) 为图 MI principal points 概览。

(对应文件为 [Figure+Table/MI-principal-points.pdf](#))

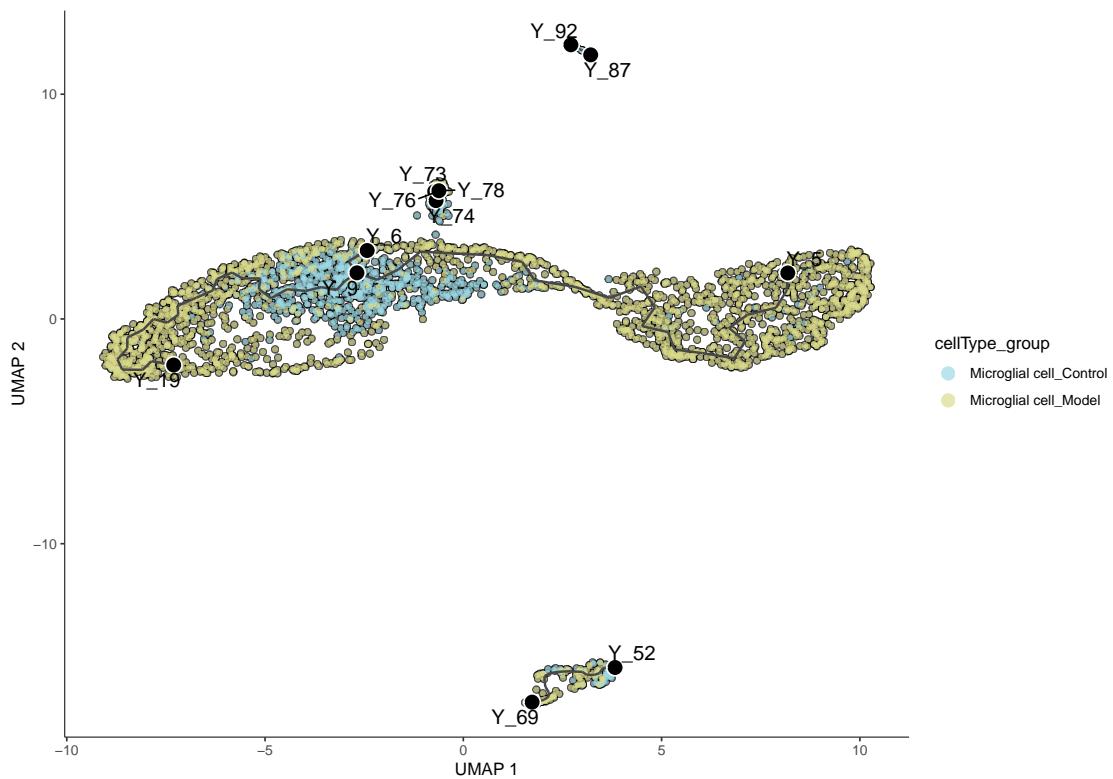


Figure 3: MI principal points

Figure 4 (下方图) 为图 MI pseudotime 概览。

(对应文件为 Figure+Table/MI-pseudotime.pdf)

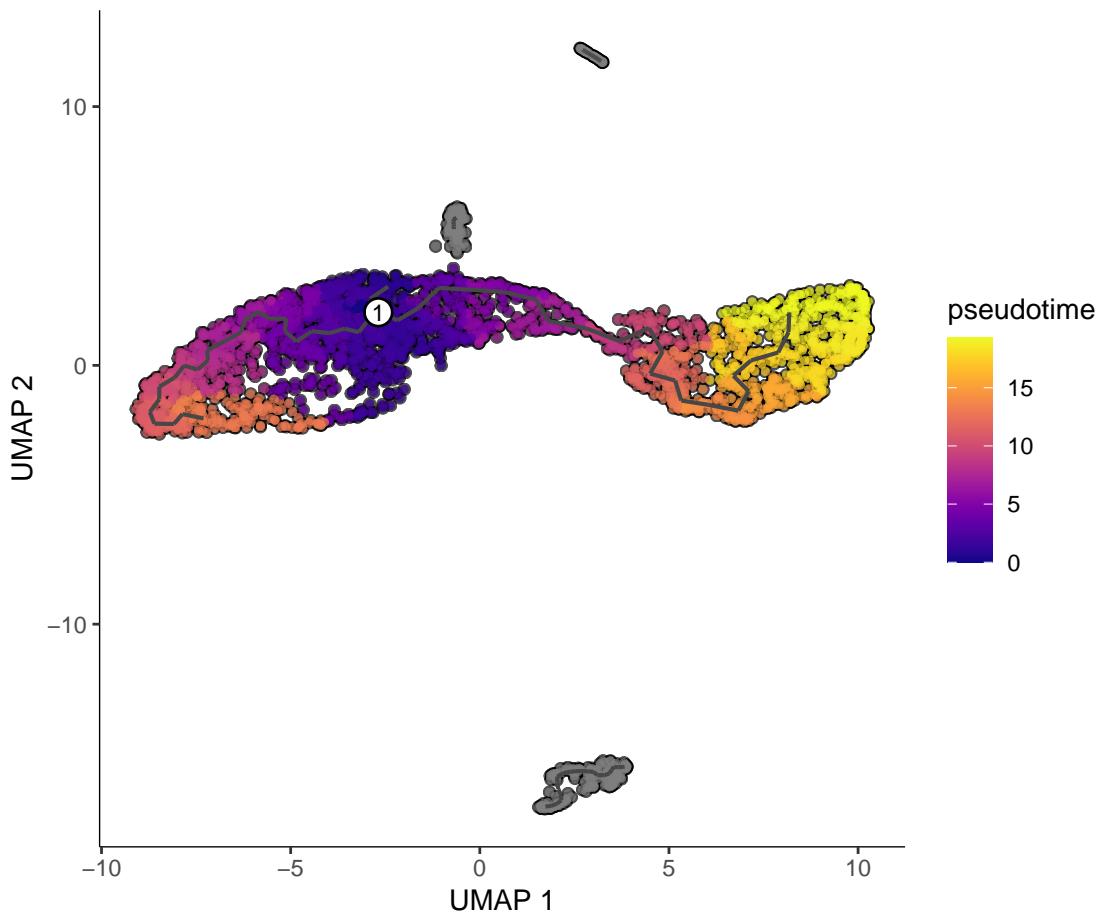


Figure 4: MI pseudotime

6.1.3.3 代谢通量预测

使用 scFEA 预测 Microglial cell 代谢通量。

Figure 5 (下方图) 为图 Convergency of the loss terms during training 概览。

(对应文件为 Figure+Table/Convergency-of-the-loss-terms-during-training.png)

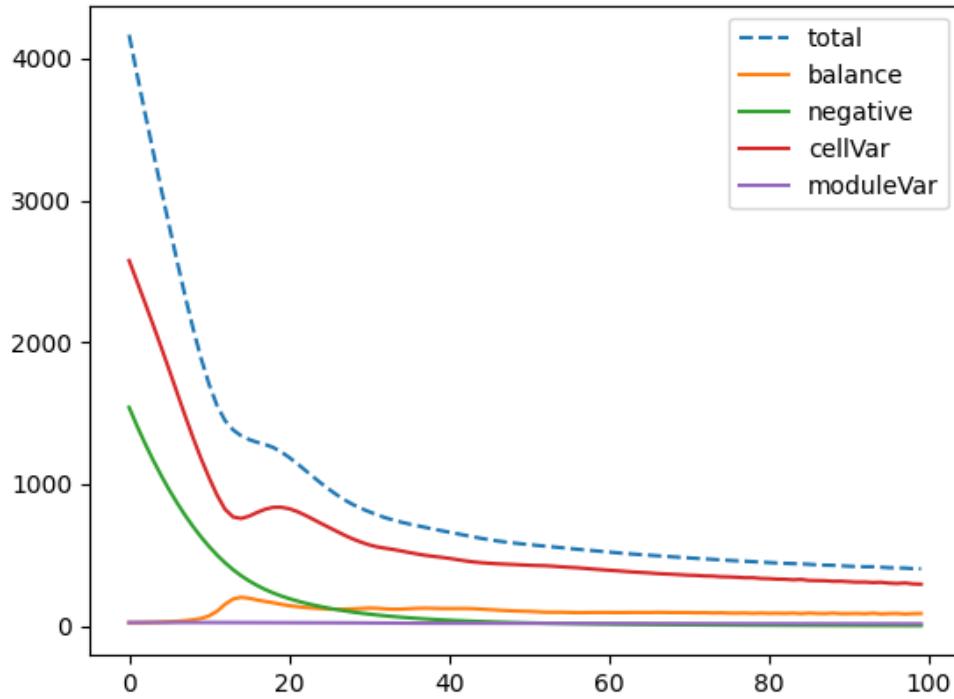


Figure 5: Convergency of the loss terms during training

6.1.3.4 代谢通量差异分析

Table 2 (下方表格) 为表格 SCF data Model vs Control 概览。

(对应文件为 `Figure+Table/SCF-data-Model-vs-Control.csv`)

注: 表格共有 13 行 8 列, 以下预览的表格可能省略部分数据; 含有 13 个唯一 ‘rownames’。

1. logFC: estimate of the log2-fold-change corresponding to the effect or contrast (for ‘topTableF’ there may be several columns of log-fold-changes)
2. AveExpr: average log2-expression for the probe over all arrays and channels, same as ‘Amean’ in the ‘MarrayLM’ object
3. t: moderated t-statistic (omitted for ‘topTableF’)
4. P.Value: raw p-value
5. B: log-odds that the gene is differentially expressed (omitted for ‘topTreat’)

Table 2: SCF data Model vs Control

rownames	logFC	AveExpr	t	P.Value	adj.P.Val	B	name
M_121	0.6117...	-3.117...	14.506...	1.1338...	1.9048...	97.444...	(Glc)3...

rownames	logFC	AveExpr	t	P.Value	adj.P.Val	B	name
M_122	0.5732...	4.8626...	13.592...	4.5332...	3.8079...	84.670...	(GlcNA...
M_132	0.5265...	-7.310...	12.485...	9.0708...	5.0796...	70.314...	(Gal)2...
M_89	0.4437...	-1.982...	10.521...	6.9549...	2.9210...	47.832...	B-Alan...
M_129	0.4401...	1.6913...	10.436...	1.7032...	5.7227...	46.948...	Protei...
M_119	0.4297...	1.0094...	10.190...	2.2013...	6.1638...	44.425...	Dolich...
M_80	0.4095...	-5.647...	9.7110...	2.7185...	6.5245...	39.680...	Cystei...
M_125	0.3755...	-3.824...	8.9058...	5.3173...	1.1166...	32.223...	Dolich...
M_95	0.3329...	-2.699...	7.8944...	2.9235...	5.4572...	23.771...	phenyl...
M_120	0.3083...	-8.293...	7.3119...	2.6374...	4.4308...	19.365...	(Glc)3...
M_109	0.3038...	-1.520...	7.2056...	5.7847...	8.8348...	18.598...	Glucos...
M_94	0.3028...	1.1379...	7.1816...	6.8961...	9.6546...	18.426...	Tyrosi...
M_146	0.3015...	1.8703...	7.1498...	8.6992...	1.1242...	18.199...	Xanthi...

[1] "Figure+Table/test1.csv"

Figure 6 (下方图) 为图 SCF Model vs Control 概览。

(对应文件为 Figure+Table/SCF-Model-vs-Control.pdf)

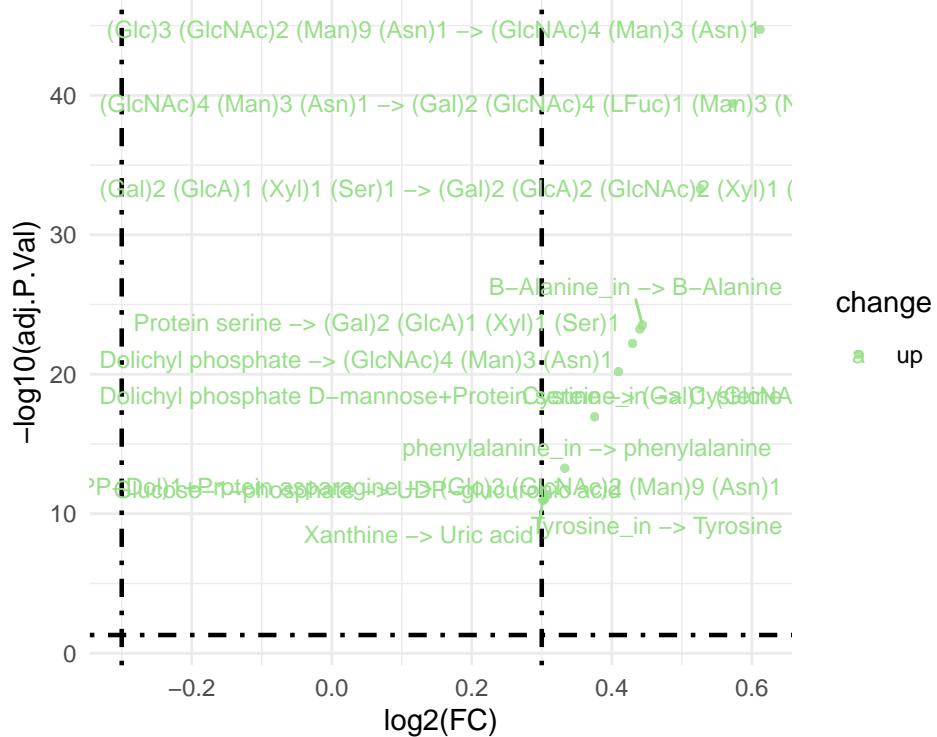


Figure 6: SCF Model vs Control

adj.P.Val cut-off :

0.05

Log2(FC) cut-off :

0.3

(上述信息框内容已保存至 Figure+Table/SCF-Model-vs-Control-content)

6.1.3.5 差异代谢相关的基因

Figure 7 (下方图) 为图 TOPFLUX GO enrichment 概览。

(对应文件为 Figure+Table/TOPFLUX-GO-enrichment.pdf)

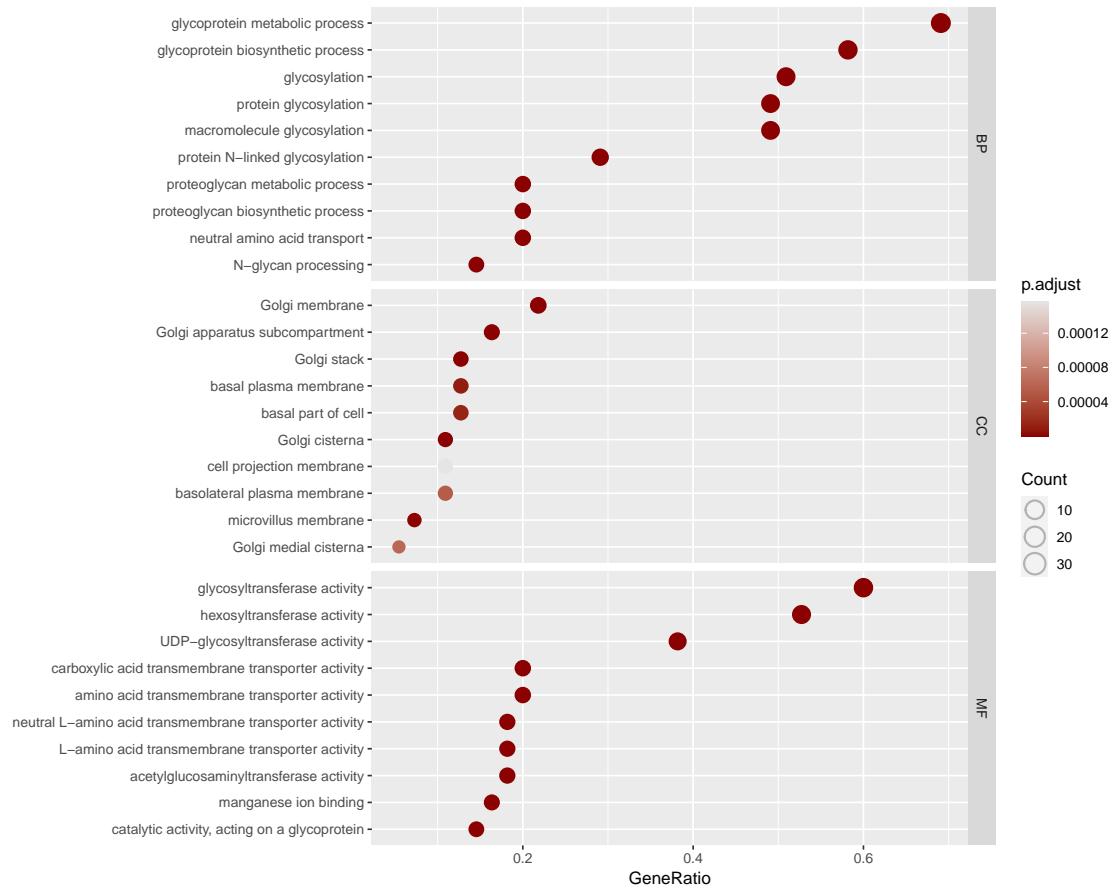


Figure 7: TOPFLUX GO enrichment

Figure 8 (下方图) 为图 MI Pseudotime heatmap of genes 概览。

(对应文件为 Figure+Table/MI-Pseudotime-heatmap-of-genes.pdf)

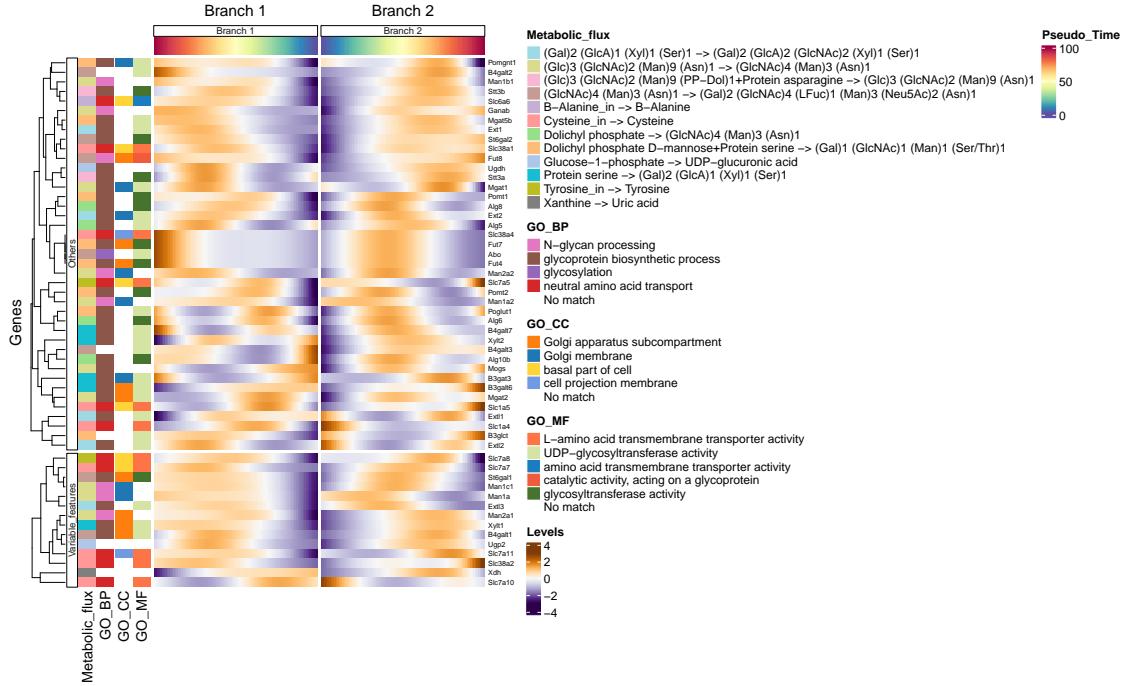


Figure 8: MI Pseudotime heatmap of genes

6.1.3.6 PPI 网络

根据 Fig. 8 中的基因, 将小鼠的基因 mgf symbol 映射为人类的基因 hgnc symbol, 构建 PPI 网络。

Figure 9 (下方图) 为图 TOPFLUX MCC score 概览。

(对应文件为 Figure+Table/TOPFLUX-MCC-score.pdf)

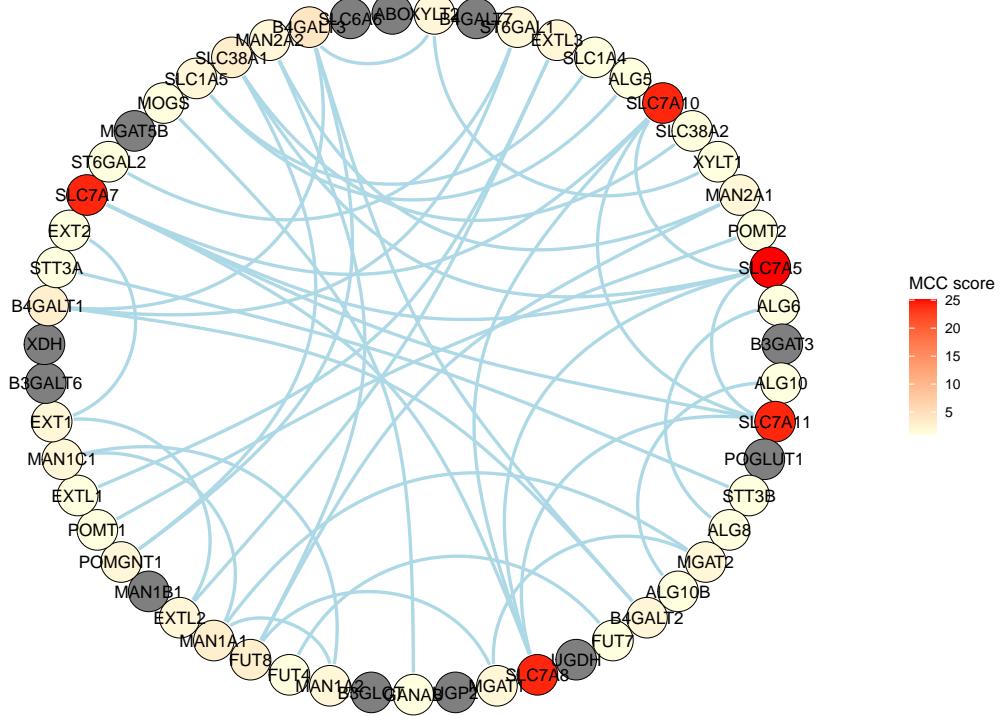


Figure 9: TOPFLUX MCC score

关注高变基因 (variable features, 差异水平更高) 与其它基因对应的蛋白的互作。

Figure 10 (下方图) 为图 TOPFLUX Top MCC score 概览。

(对应文件为 Figure+Table/TOPFLUX-Top-MCC-score.pdf)

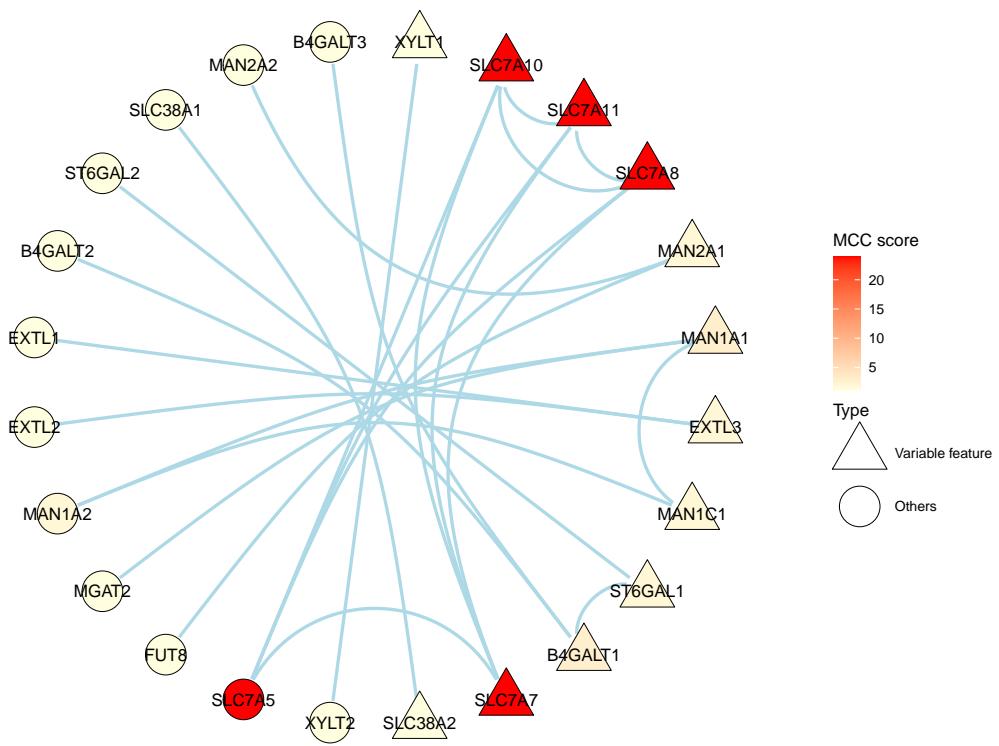


Figure 10: TOPFLUX Top MCC score

6.1.3.7 上游的转录因子

寻找 Fig. 8 中的基因的上游转录因子。

Table 3 (下方表格) 为表格 Transcription Factor binding sites 概览。

(对应文件为 Figure+Table/Transcription-Factor-binding-sites.csv)

注: 表格共有 15113 行 10 列, 以下预览的表格可能省略部分数据; 含有 56 个唯一 ‘target’。

1. Start: 起始点

Table 3: Transcription Factor binding sites

target	TF_symbol	Motif	Source	Strand	Start	Stop	PValue	MatchS...	Overla...
XDH	FOXB1	FOXB1_...	SELEX	+	31634777	31634794	2.0E-06	TTGATA...	18
XDH	FOXB1	FOXB1_...	SELEX	-	31634777	31634794	4.0E-06	ATAGTC...	18
XDH	POU4F2	POU4F2...	SELEX	+	31635992	31636007	4.0E-06	TTTAAT...	16
XDH	HOXD12	HOXD12...	SELEX	+	31636000	31636008	5.0E-06	ATAATAAAA	9

target	TF_symbol	Motif	Source	Strand	Start	Stop	PValue	MatchS...	Overla...
XDH	HOXD12	HOXD12...	SELEX	+	31636072	31636080	2.0E-06	CTAATAAAA	9
XDH	FOXJ2	FOXJ2_...	SELEX	+	31635996	31636008	2.0E-06	ATAAAAT...	13
XDH	FOXJ2	FOXJ2_...	SELEX	+	31637732	31637744	9.0E-06	GCAAAC...	13
XDH	HOXC10	Hoxc10...	SELEX	+	31636000	31636009	6.0E-06	ATAATA...	10
XDH	HOXC10	Hoxc10...	SELEX	+	31636072	31636081	3.0E-06	CTAATA...	10
XDH	HOXC10	Hoxc10...	SELEX	-	31639392	31639401	7.0E-06	ACAATA...	10
XDH	SOX21	SOX21_...	SELEX	-	31636149	31636163	0.0E+00	AGCAAT...	15
XDH	SOX4	SOX4_H...	SELEX	-	31636149	31636164	1.0E-06	CAGCAA...	16
XDH	HOXD9	Hoxd9_...	SELEX	-	31639392	31639401	3.0E-06	ACAATA...	10
XDH	POU2F1	POU2F1...	SELEX	+	31635995	31636008	5.0E-06	AATAAA...	14
XDH	POU2F1	POU2F1...	SELEX	+	31637713	31637726	6.0E-06	TTTACA...	14
...

6.1.4 星形细胞胶质瘢痕相关基因

6.1.4.1 差异分析

Table 4 (下方表格) 为表格 DEGs of the contrasts Astrocyte 概览。

(对应文件为 `Figure+Table/DEGs-of-the-contrasts-Astrocyte.csv`)

注: 表格共有 1058 行 7 列, 以下预览的表格可能省略部分数据; 含有 1 个唯一 ‘contrast’。

Table 4: DEGs of the contrasts Astrocyte

contrast	p_val	avg_log2FC	pct.1	pct.2	p_val_adj	gene
Astrocyte_...	4.37264108...	8.78114541...	0.024	0.107	1.31179232...	Dnah11
Astrocyte_...	3.64754326...	2.65537976...	0.093	0.243	1.09426297...	H2-Aa
Astrocyte_...	2.70284056...	2.40818618...	0.056	0.148	8.10852170...	Cd53
Astrocyte_...	6.14268687...	8.30800266...	0.08	0.147	1.84280606...	Thbs1
Astrocyte_...	9.59253467...	2.25479722...	0.062	0.198	2.87776040...	Mpeg1
Astrocyte_...	1.90263565...	14.3762318...	0.073	0.274	5.70790697...	Top2a
Astrocyte_...	3.92239217...	4.31264893...	0.071	0.192	1.17671765...	Cebpb
Astrocyte_...	7.06233131...	7.35551440...	0.058	0.21	2.11869939...	Grap2
Astrocyte_...	4.42147374...	2.20926332...	0.069	0.205	1.32644212...	A630001G21Rik
Astrocyte_...	1.77494299...	0.93300639...	0.106	0.063	5.32482897...	Stab1
Astrocyte_...	6.62357870...	-1.2610776...	0.075	0.165	1.98707361...	Cd247
Astrocyte_...	1.03473035...	-1.9719507...	0.136	0.196	3.10419105...	Gpnmb
Astrocyte_...	3.73638612...	12.9129020...	0.14	0.225	1.12091583...	Bcl3
Astrocyte_...	3.00567566...	4.02484698...	0.14	0.248	9.01702698...	Ifitm3
Astrocyte_...	1.91228260...	5.32650753...	0.047	0.244	5.73684782...	Rnf17

contrast	p_val	avg_log2FC	pct.1	pct.2	p_val_adj	gene
...

将这些差异基因映射到人类的基因 hgnc_symbol

Table 5 (下方表格) 为表格 mapped genes Astrocyte DEGs 概览。

(对应文件为 **Figure+Table/mapped-genes-Astrocyte-DEGs.csv**)

注: 表格共有 921 行 2 列, 以下预览的表格可能省略部分数据; 含有 892 个唯一 ‘mgi_symbol’; 含有 914 个唯一 ‘hgnc_symbol’。

1. hgnc_symbol: 基因名 (Human)
2. mgi_symbol: 基因名 (Mice)

Table 5: Mapped genes Astrocyte DEGs

mgi_symbol	hgnc_symbol
Man1a	MAN1A1
Slc7a7	SLC7A7
B4galt1	B4GALT1
Xdh	XDH
St6gal1	ST6GAL1
mt-Atp8	MT-ATP8
Egln3	EGLN3
Mis18bp1	MIS18BP1
Cdc42ep1	CDC42EP1
Sdc4	SDC4
Cd93	CD93
Shroom2	SHROOM2
Jag1	JAG1
Fbxo7	FBXO7
Parp4	PARP4
...	...

6.1.4.2 胶质瘢痕

从 genecards 获取胶质瘢痕相关基因。

Table 6 (下方表格) 为表格 GLIALSCAR disease related targets from GeneCards 概览。

(对应文件为 **Figure+Table/GLIALSCAR-disease-related-targets-from-GeneCards.csv**)

注：表格共有 90 行 7 列，以下预览的表格可能省略部分数据；含有 90 个唯一 ‘Symbol’。

The GeneCards data was obtained by querying :

Glial scar

Restrict (with quotes) :

TRUE

Filtering by Score:

Score > 0

Table 6: GLIALSCAR disease related targets from GeneCards

Symbol	Description	Category	UniProt_ID	GIFTs	GC_id	Score
BDNF-AS	BDNF Antis...	RNA Gene (...)		29	GC11P027466	4.85
TRA-TGC7-1	TRNA-Ala (...)	RNA Gene (...)		14	GC06M093612	2.82
CSPG4	Chondroiti...	Protein Co...	Q6UVK1	54	GC15M075674	2.4
GFAP	Glial Fibr...	Protein Co...	P14136	57	GC17M077883	1.91
MAG	Myelin Ass...	Protein Co...	P20916	55	GC19P035292	1.78
TNR	Tenascin R	Protein Co...	Q92752	51	GC01M175291	1.78
MYOC	Myocilin	Protein Co...	Q99972	50	GC01M171604	1.78
MMP9	Matrix Met...	Protein Co...	P14780	62	GC20P046008	1.7
S100B	S100 Calci...	Protein Co...	P04271	53	GC21M053599	1.7
PDGFRB	Platelet D...	Protein Co...	P09619	62	GC05M150113	1.59
CST3	Cystatin C	Protein Co...	P01034	53	GC20M023930	1.59
NES	Nestin	Protein Co...	P48681	50	GC01M156668	1.59
FGFR4	Fibroblast...	Protein Co...	P22455	60	GC05P177086	1.32
FGF2	Fibroblast...	Protein Co...	P09038	54	GC04P122826	1.32
GALNS	Galactosam...	Protein Co...	P34059	56	GC16M088813	0.7
...

Figure 11 (下方图) 为图 Intersection of DB GlialScar with Astrocyte DEGs 概览。

(对应文件为 Figure+Table/Intersection-of-DB-GlialScar-with-Astrocyte-DEGs.pdf)

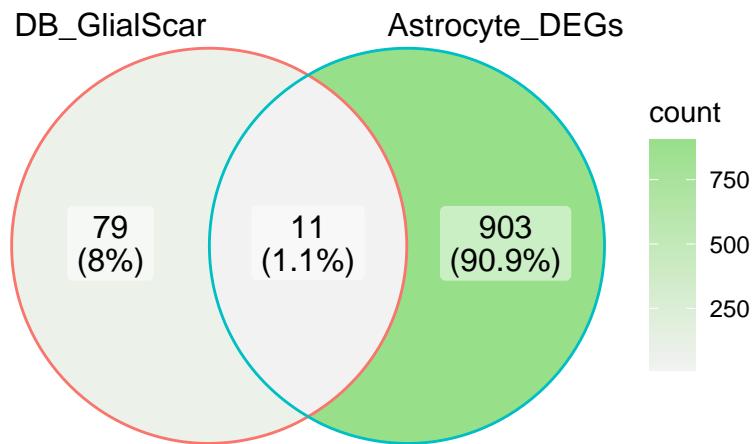


Figure 11: Intersection of DB GlialScar with Astrocyte DEGs

Intersection :

GFAP, PDGFRB, CST3, VIM, BDNF, DCN, TGM2, VCAN, TNC, MRC1, CCL15-CCL14

(上述信息框内容已保存至 [Figure+Table/Intersection-of-DB-GliaScar-with-Astrocyte-DEGs-content](#))

6.1.5 调控小胶质细胞代谢且与星形细胞胶质瘢痕相关基因

6.1.5.1 交集基因

Figure 12 (下方图) 为图 UpSet plot of Genes sources 概览。

(对应文件为 [Figure+Table/UpSet-plot-of-Genes-sources.pdf](#))

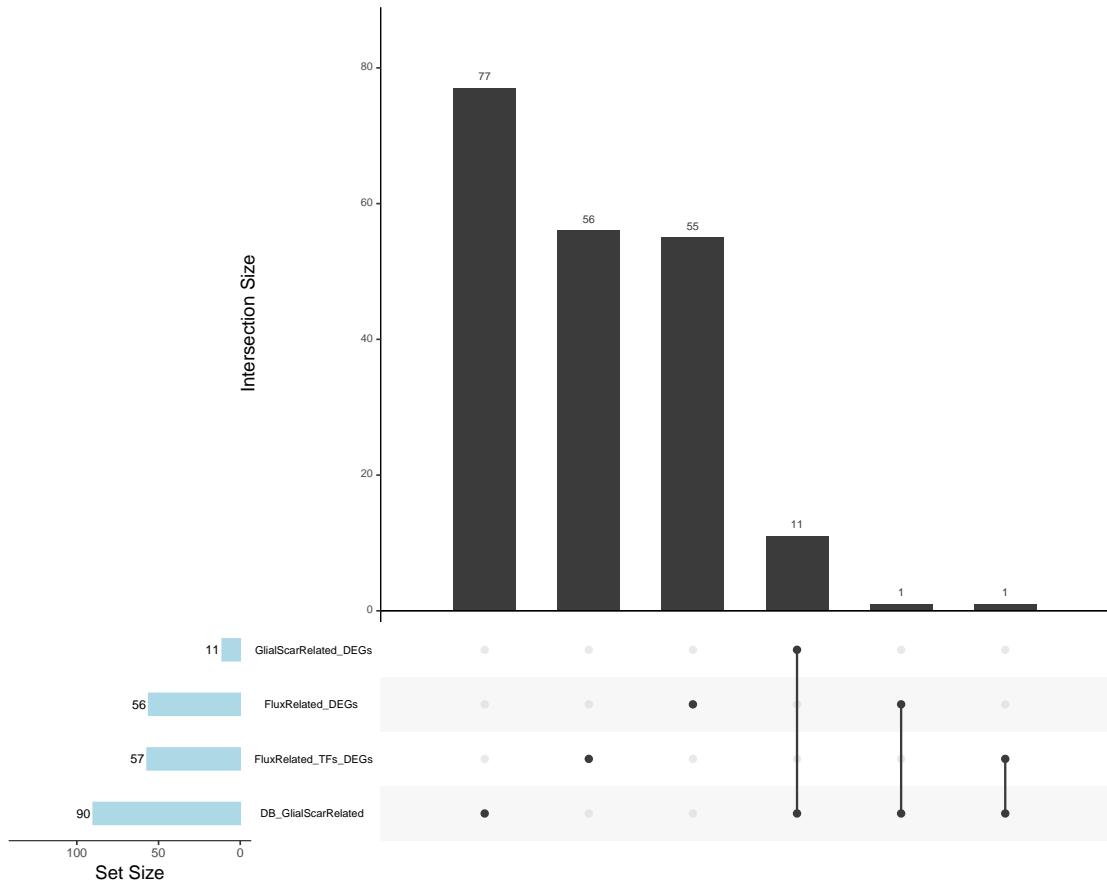


Figure 12: UpSet plot of Genes sources

All_intersection :

Other intersection :

”FluxRelated_DEGs” WITH ”DB_GlialScarRelated”: XYL1

”FluxRelated_TFs_DEGs” WITH ”DB_GlialScarRelated”: STAT3

(上述信息框内容已保存至 [Figure+Table/UpSet-plot-of-Genes-sources-content](#))

6.1.5.2 在小胶质细胞中的表达

可以发现，‘Xylt1’ 主要集中表达于模型组的 Microglial 中。

Figure 13 (下方图) 为图 Dimension plot of expression level of the genes 概览。

(对应文件为 [Figure+Table/Dimension-plot-of-expression-level-of-the-genes.pdf](#))

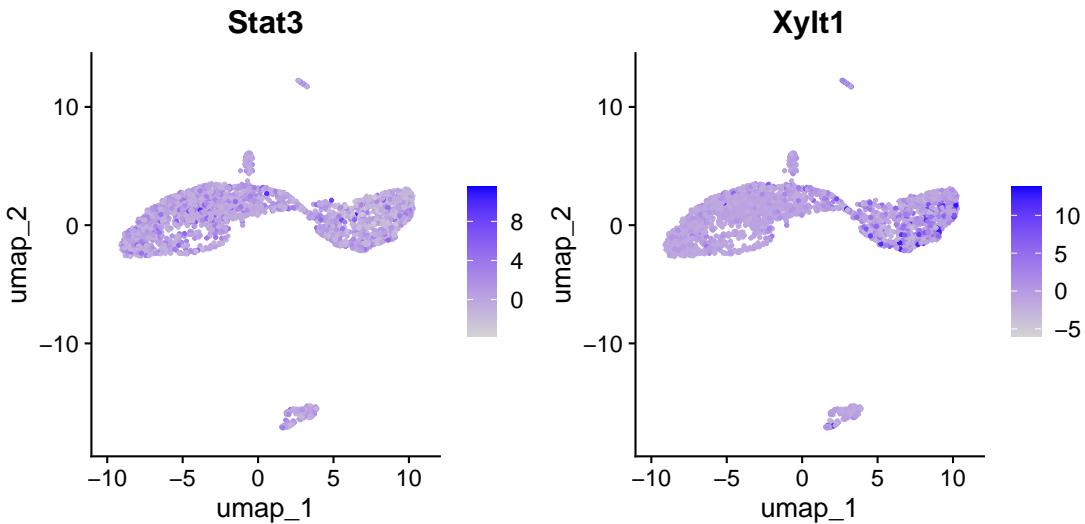


Figure 13: Dimension plot of expression level of the genes

Table 7 (下方表格) 为表格 Xylt1 related metabolic flux 概览。

(对应文件为 [Figure+Table/Xylt1-related-metabolic-flux.csv](#))

注：表格共有 1 行 2 列，以下预览的表格可能省略部分数据；含有 1 个唯一 ‘gene’。

Table 7: Xylt1 related metabolic flux

gene	Metabolic_flux
Xylt1	Protein serine -> (Gal)2 (G...

Reference

1. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the r/bioconductor package biomaRt. *Nature protocols* **4**, 1184–1191 (2009).
2. Alghamdi, N. *et al.* A graph neural network model to estimate cell-wise metabolic flux using single-cell rna-seq data. *Genome research* **31**, 1867–1884 (2021).
3. Wu, T. *et al.* ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* **2**, (2021).
4. Stelzer, G. *et al.* The genecards suite: From gene data mining to disease genome sequence analyses. *Current protocols in bioinformatics* **54**, 1.30.1–1.30.33 (2016).
5. Ritchie, M. E. *et al.* Limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47 (2015).
6. Chen, Y., McCarthy, D., Ritchie, M., Robinson, M. & Smyth, G. EdgeR: Differential analysis of sequence read count data user’s guide. 119.

7. Qiu, X. *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nature Methods* **14**, (2017).
8. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature Biotechnology* **32**, (2014).
9. Szklarczyk, D. *et al.* The string database in 2021: Customizable proteinprotein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research* **49**, D605–D612 (2021).
10. Chin, C.-H. *et al.* CytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Systems Biology* **8**, S11 (2014).
11. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, (2021).
12. Stuart, T. *et al.* Comprehensive integration of single-cell data. *Cell* **177**, (2019).
13. Plaisier, C. L. *et al.* Causal mechanistic regulatory network for glioblastoma deciphered using systems genetics network analysis. *Cell systems* **3**, 172–186 (2016).
14. Cao, Y., Wang, X. & Peng, G. SCSA: A cell type annotation tool for single-cell rna-seq data. *Frontiers in genetics* **11**, (2020).