XX 基因通过促进糖酵解促进巨噬细胞 M1 极化

2024-05-06

LiChuang Huang

@ 立效研究院

Contents

List of Figures

List of Tables

1 摘要

1.1 生信需求

疾病：类风湿性关节炎 RA 物种：临床患者或者动物模型都可以细胞：巨噬细胞

目标：筛出 XX 基因，XX 基因满足，1、是糖酵解相关基因 2、与巨噬细胞极化相关（M1/M2）

设想：XX 基因在 RA 中上调，RA 中 M1 巨噬细胞上调，其中 XX 基因主要分布在 M1 巨噬细胞而非 M2 巨噬细胞上，XX 基因可能通过促进糖酵解促进巨噬细胞 M1 极化

M1 标志：iNOS，CD11c，CD86 等 M2 标志：CD206，IL-10，TGF-beta 等

1.2 结果

- 首先通过分析 GEO 单细胞数据，鉴定出巨噬细胞不同表型 Fig. 5。
- 该数据集为小鼠来源，鉴定 M0、M1、M2 的小鼠基因 Marker 参考[1]，实际使用的 Marker 见 Tab. 1
- 进行差异分析 (Tab. 2)：
  - XX 在 RA 中 M1 巨噬细胞上调: `GPI-day25-RA_Macrophage_M1` vs `Control_Macrophage_M1`
  - 其中 XX 基因主要分布在 M1 巨噬细胞而非 M2 巨噬细胞上: `GPI-day25-RA_Macrophage_M1` vs `GPI-day25-RA_Macrophage_M2`
- 以上两组差异基因交集见 Fig. 6
- 小鼠基因映射到人类 Tab. 3
- 其中糖酵解相关的基因见 Fig. 7
- 筛选到唯一的基因: PPARG (小鼠 Pparg)。其表达特征见 Fig. 8

1.3 进一步分析需求

利用开源数据库进行生物信息学分析，筛选并验证类风湿性关节炎临床患者和动物模型中与巨噬细胞极化和糖酵解相关的关键基因 XX 的表达情况

- XX (VWF) 表达水平与炎症因子、巨噬细胞浸润、巨噬细胞极化相关因子、糖酵解相关因子的相关性
- 与患者状态（例如血清类风湿因子（RF）、抗链球菌溶血素抗体（ASO）、红细胞沉降率（ESR）和 C 反应蛋白（CRP））的相关性

1.4 进一步分析结果

- 关联分析结果见 Fig. 9, Tab. 6。
- 未找到可用的 RA 表型数据集。

2 前言

3 材料和方法

3.1 材料

All used GEO expression data and their design:

- **GSE184609**: scRNA-Seq analysis of FACS-sorted live synovial cells isolated from naïve mice (two replicates) or from mice at day 6, 14, or 25 of GPI-induced arthritis (one replicate per time point).
- **GSE17755**: Peripheral blood was obtained from patients with RA (n=112), SLE (n=22), polyJIA (n=6), sJIA (n=51), HC (n=8), and HI (n=45). Blood samples from 8 HC and 45 HI are used as control.

3.2 方法

Mainly used method:

- The `biomart` was used for mapping genes between organism (e.g., mgi_symbol to hgnc_symbol)[2].
- The Human Gene Database `GeneCards` used for disease related genes prediction[3].
- GEO https://www.ncbi.nlm.nih.gov/geo/ used for expression dataset aquisition.
- R package `Limma` and `edgeR` used for differential expression analysis[4,5].
- The data in published article of Jablonski et al used for distinguishing macrophage phenotypes (M0/M1/M2)[1].
- The R package `Seurat` used for scRNA-seq processing[6,7].
- `SCSA` (python) used for cell type annotation[8].
- 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 scRNA-seq

6.1.1 数据来源

这是一批小鼠的单细胞测序数据。

Data Source ID :

GSE184609

data_processing :

10X Genomics Cell Ranger v3.1

data_processing.1 :

Gene-cell UMI matrix was generated for downstream analyses. Low-quality cells were removed based on their unique feature counts and mitochondrial gene content. Data was normalized and log transformed using the default setting of Seurat (version 3.1.4).

data_processing.2 :

Genome_build: mm10

data_processing.3 :

Supplementary_files_format_and_content: For each sample, there is one mtx file with filtered gene expressing UMI counts for each sample, one tsv file containing gene names, and one tsv file with cell barcodes.

(上述信息框内容已保存至 `Figure+Table/GSE184609-content`)

6.1.2 细胞聚类与初步注释

使用 SCSA 对细胞类型注释。

Figure 1 (下方图) 为图 SCSA Cell type annotation 概览。

(对应文件为 `Figure+Table/SCSA-Cell-type-annotation.pdf`)

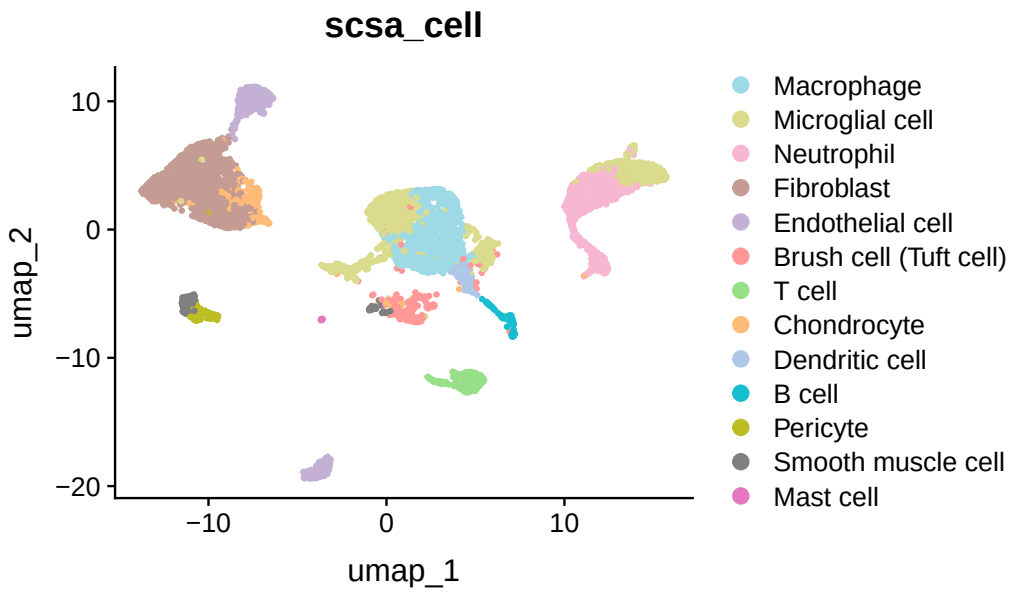

Figure 1: SCSA Cell type annotation

6.1.3 巨噬细胞重聚类

Figure 2 (下方图) 为图 Microphage UMAP Clustering 概览。

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

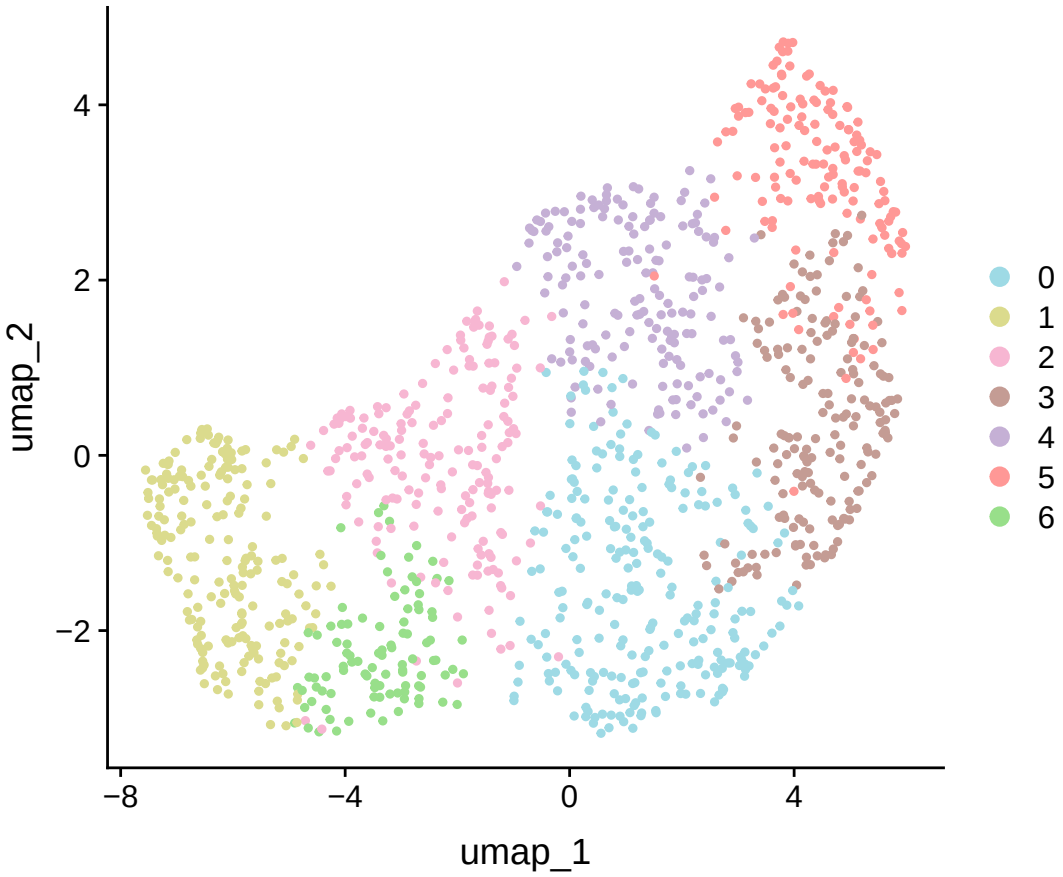

Figure 2: Microphage UMAP Clustering

6.1.4 巨噬细胞表型 M0、M1、M2 鉴定 Markers

Table 1 (下方表格) 为表格 The markers for Macrophage phenotypes annotation 概览。

(对应文件为 `Figure+Table/The-markers-for-Macrophage-phenotypes-annotation.csv`)

> 注：表格共有 26 行 2 列，以下预览的表格可能省略部分数据；含有 3 个唯一 'cell'。

Table 1: The markers for Macrophage phenotypes annotation

cell	markers
Macrophage_M0	Sh2d3c
Macrophage_M0	Slc13a3
Macrophage_M0	Rcan1
Macrophage_M0	Trp53inp1
Macrophage_M0	Slc40a1
Macrophage_M0	Il16
Macrophage_M1	Cfb

cell	markers
Macrophage_M1	Slfn4
Macrophage_M1	H2-Q6
Macrophage_M1	Fpr1
Macrophage_M1	Slfn1
Macrophage_M1	Ccrl2
Macrophage_M1	Fpr2
Macrophage_M1	Cxcl10
Macrophage_M1	Oasl1
...	...

Figure 3 (下方图) 为图 Heatmap show the reference genes 概览。

(对应文件为 `Figure+Table/Heatmap-show-the-reference-genes.pdf`)

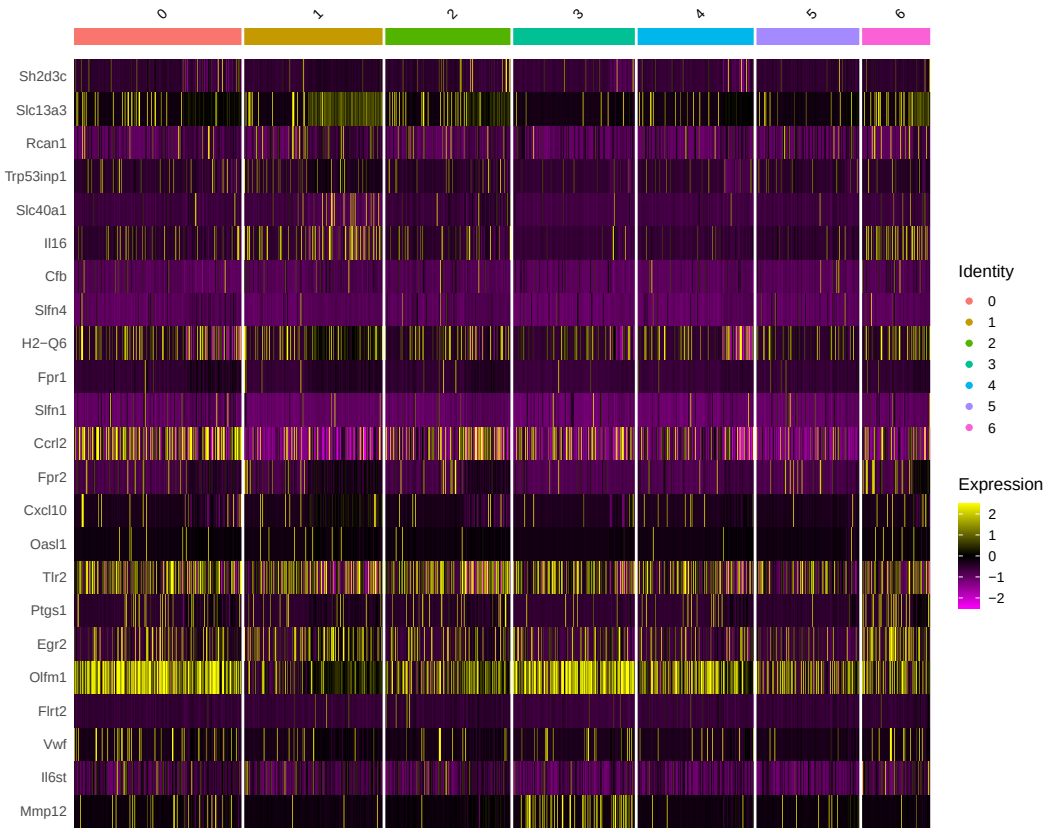

Figure 3: Heatmap show the reference genes

Figure 4 (下方图) 为图 Macrophage phenotypes type annotation 概览。

(对应文件为 `Figure+Table/Macrophage-phenotypes-type-annotation.pdf`)

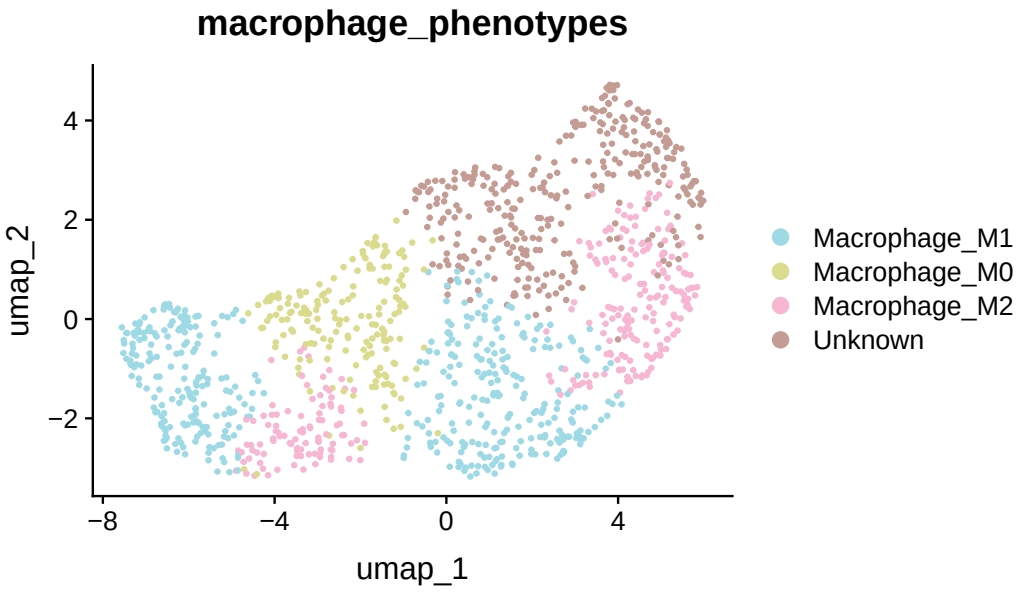

Figure 4: Macrophage phenotypes type annotation

6.1.5 RA 与 Control 的巨噬细胞表型

随后，根据数据集的来源 (RA 或 Control，将巨噬细胞分类)

Figure 5 (下方图) 为图 The Phenotypes 概览。

(对应文件为 `Figure+Table/The-Phenotypes.pdf`)

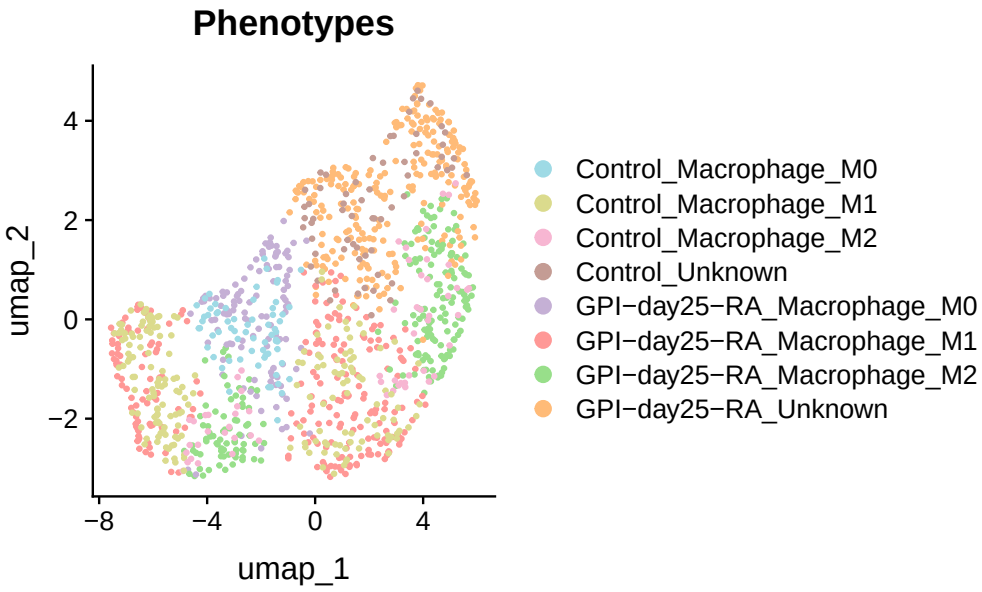

Figure 5: The Phenotypes

6.1.6 差异分析

- XX 在 RA 中 M1 巨噬细胞上调: `GPI-day25-RA_Macrophage_M1` vs `Control_Macrophage_M1`
- 其中 XX 基因主要分布在 M1 巨噬细胞而非 M2 巨噬细胞上: `GPI-day25-RA_Macrophage_M1` vs `GPI-day25-RA_Macrophage_M2`

Table 2 (下方表格) 为表格 DEGs of the contrasts 概览。

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

> 注：表格共有 355 行 7 列，以下预览的表格可能省略部分数据；含有 2 个唯一 'contrast；含有 335 个唯一 'gene'。

Table 2: DEGs of the contrasts

contrast	p_val	avg_log2FC	pct.1	pct.2	p_val_adj	gene
GPI-day25-...	1.75831850...	2.17770295...	0.044	0.385	5.27495551...	Adora3
GPI-day25-...	2.23900708...	9.52901476...	0.069	0.427	6.71702126...	F7
GPI-day25-...	6.04375313...	9.97788827...	0.093	0.536	1.81312594...	Hal
GPI-day25-...	1.83819770...	13.5930067...	0.052	0.641	5.51459312...	Cxcl13
GPI-day25-...	1.00690808...	4.79632080...	0.153	0.583	3.02072424...	Ifi44
GPI-day25-...	2.16444867...	8.41136649...	0.153	0.87	6.49334601...	Slc13a3
GPI-day25-...	8.87142099...	7.46793928...	0.081	0.391	2.66142629...	Cd4
GPI-day25-...	1.52778766...	0.95745400...	0.141	0.307	4.58336298...	Tnfsf14
GPI-day25-...	4.80404528...	3.93765968...	0.169	0.484	1.44121358...	Cd79b
GPI-day25-...	8.42060311...	3.12179649...	0.06	0.651	2.52618093...	Cd209e
GPI-day25-...	8.42067724...	1.96704094...	0.145	0.786	2.52620317...	Adgre4
GPI-day25-...	8.24915617...	8.83724798...	0.161	0.766	2.47474685...	Pparg
GPI-day25-...	1.31157015...	8.50505864...	0.153	0.292	3.93471045...	F10
GPI-day25-...	2.28779328...	2.51813054...	0.024	0.411	6.86337986...	Apoc4
GPI-day25-...	2.75634208...	4.11075823...	0.073	0.755	8.26902626...	Il10
...	...	...	...	...	...	...

Figure 6 (下方图) 为图 Intersection of RA M1 up with M1 not M2 概览。

(对应文件为 `Figure+Table/Intersection-of-RA-M1-up-with-M1-not-M2.pdf`)

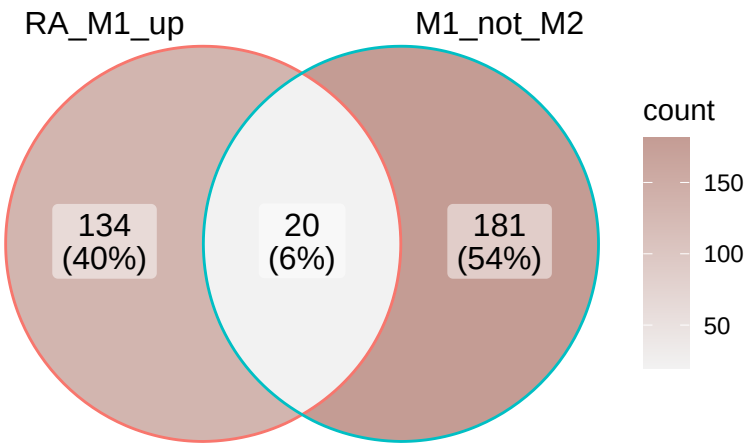

Figure 6: Intersection of RA M1 up with M1 not M2

Intersection :

Ifi44, Adgre4, Pparg, Dppa3, Cadm1, P2ry14, Gm1673, Vwf, Ednrb, Fam43a, Bambi, Slc28a2, Plk2, Rcn3, Rrm1, Ifi204, Bmp2, Gfra2, Spon1, Gstm1

(上述信息框内容已保存至 `Figure+Table/Intersection-of-RA-M1-up-with-M1-not-M2-content`)

6.2 小鼠基因映射到人类基因

Table 3 (下方表格) 为表格 Mapped genes 概览。

(对应文件为 `Figure+Table/Mapped-genes.csv`)

注: 表格共有 19 行 2 列, 以下预览的表格可能省略部分数据; 含有 19 个唯一 'mgi_symbol; 含有 19 个唯一 'hgnc_symbol'。

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

Table 3: Mapped genes

mgi_symbol	hgnc_symbol
Bmp2	BMP2
Ednrb	EDNRB
Dppa3	DPPA3
Spon1	SPON1
Gfra2	GFRA2
Bambi	BAMBI

mgi_symbol	hgnc_symbol
Cadm1	CADM1
Slc28a2	SLC28A2
Rrm1	RRM1
Ifi44	IFI44
Gm1673	C4orf48
Ifi204	MNDA
P2ry14	P2RY14
Rcn3	RCN3
Gstm1	GSTM1
…	…

6.3 糖酵解相关基因

Table 4 (下方表格) 为表格 Glycolysis related genes from GeneCards 概览。

(对应文件为 `Figure+Table/Glycolysis-related-genes-from-GeneCards.xlsx`)

注：表格共有 118 行 7 列，以下预览的表格可能省略部分数据；含有 118 个唯一 'Symbol'。

The GeneCards data was obtained by querying :

Glycolysis

Restrict (with quotes) :

FALSE

Filtering by Score: :

Score > 3

Table 4: Glycolysis related genes from GeneCards

Symbol	Description	Category	UniProt_ID	GIFtS	GC_id	Score
TIGAR	TP53 Induc…	Protein Co…	Q9NQ88	45	GC12P038924	22.4
PKM	Pyruvate K…	Protein Co…	P14618	58	GC15M072199	20.77
HK2	Hexokinase 2	Protein Co…	P52789	55	GC02P074947	19.42
GAPDH	Glyceralde…	Protein Co…	P04406	59	GC12P038965	17.14
LDHA	Lactate De…	Protein Co…	P00338	59	GC11P018394	15.81
HIF1A	Hypoxia In…	Protein Co…	Q16665	57	GC14P061695	15.1
RRAD	RRAD, Ras …	Protein Co…	P55042	46	GC16M067483	15.1

Symbol	Description	Category	UniProt_ID	GIFtS	GC_id	Score
HK1	Hexokinase 1	Protein Co...	P19367	59	GC10P069269	14.64
PKLR	Pyruvate K...	Protein Co...	P30613	55	GC01M155289	13.37
ENO1	Enolase 1	Protein Co...	P06733	56	GC01M008861	13.36
ENO3	Enolase 3	Protein Co...	P13929	54	GC17P004948	13.33
PFKP	Phosphofru...	Protein Co...	Q01813	53	GC10P003066	13.19
TPI1	Triosephos...	Protein Co...	P60174	55	GC12P006867	13.18
GLTC1	Glycolysis...	RNA Gene (...		2	GC11U909607	12.97
PGK1	Phosphogly...	Protein Co...	P00558	57	GC0XP078166	12.94
...	...	...	...	...	...	...

Figure 7 (下方图) 为图 Intersection of RA M1M2 related with Glycolysis related 概览。

(对应文件为 `Figure+Table/Intersection-of-RA-M1M2-related-with-Glycolysis-related.pdf`)

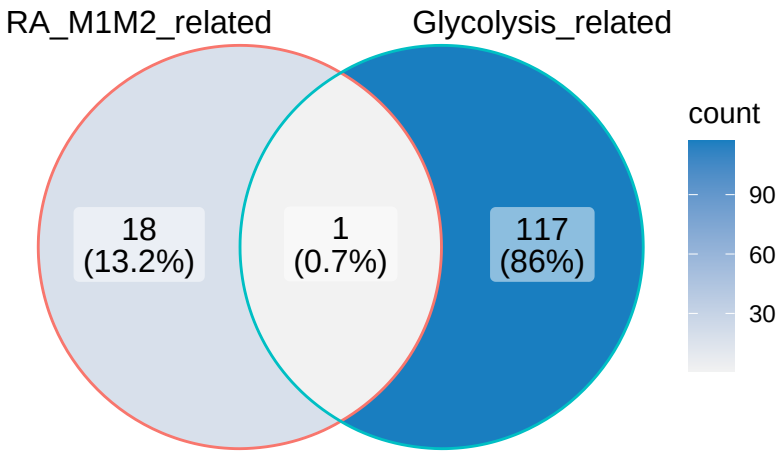

Figure 7: Intersection of RA M1M2 related with Glycolysis related

Intersection :

PPARG

(上述信息框内容已保存至 `Figure+Table/Intersection-of-RA-M1M2-related-with-Glycolysis-related-content`)

6.4 交集基因的表达 (小鼠单细胞数据)

Figure 8 (下方图) 为图 Violing plot of expression level of the Pparg 概览。

(对应文件为 `Figure+Table/Violing-plot-of-expression-level-of-the-Pparg.pdf`)

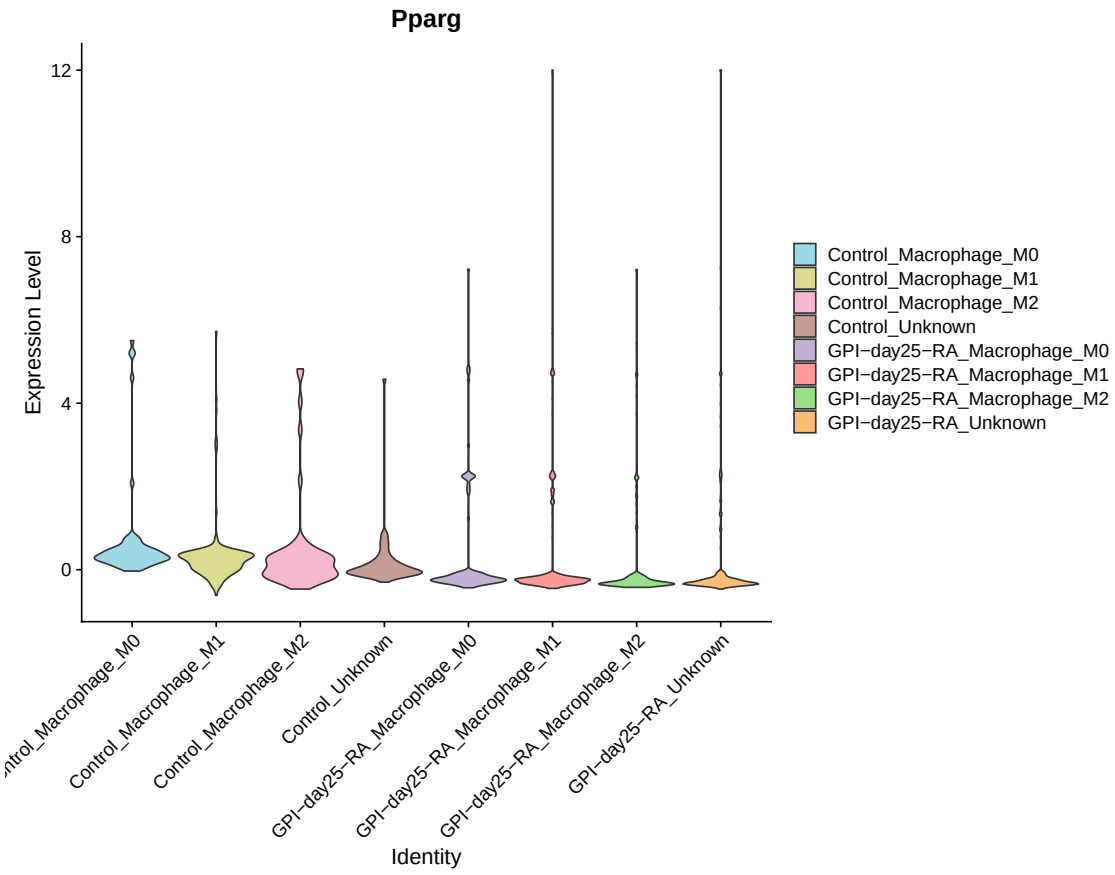

Figure 8: Violing plot of expression level of the Pparg

7 进一步分析

7.1 数据来源

Data Source ID :

GSE17755

data_processing :

Log2 ratios of Cy3 to Cy5 were calculated and normalized by the method of global ratio median normalization.

(上述信息框内容已保存至 `Figure+Table/HUMAN-GSE17755-content`)

7.2 炎症因子、巨噬细胞浸润、巨噬细胞极化相关因子、糖酵解相关因子

使用 genecards 获取相关基因 (各取前 50 基因):

- IF: Inflammatory factors 炎症因子
- MI: Macrophage infiltration 巨噬细胞浸润
- MP: Macrophage polarization 巨噬细胞极化
- G: Glycolysis 糖酵解

Table 5 (下方表格) 为表格 All Factors 概览。

(对应文件为 `Figure+Table/All-Factors.csv`)

> 注：表格共有 200 行 2 列，以下预览的表格可能省略部分数据；含有 4 个唯一 'type'。

Table 5: All Factors

type	name
Inflammatory factors	IL6
Inflammatory factors	TNF
Inflammatory factors	CRP
Inflammatory factors	BDNF-AS
Inflammatory factors	IL1B
Inflammatory factors	LINC02605
Inflammatory factors	TLR4
Inflammatory factors	MIR146B
Inflammatory factors	ADIPOQ
Inflammatory factors	LINC01672
Inflammatory factors	CXCL8
Inflammatory factors	IL1A
Inflammatory factors	NFKB1
Inflammatory factors	CERNA3
Inflammatory factors	IL18
...	...

对上述基因集去重复后，关联分析。

7.3 关联分析

Figure 9 (下方图) 为图 HUMAN correlation heatmap 概览。

(对应文件为 `Figure+Table/HUMAN-correlation-heatmap.pdf`)

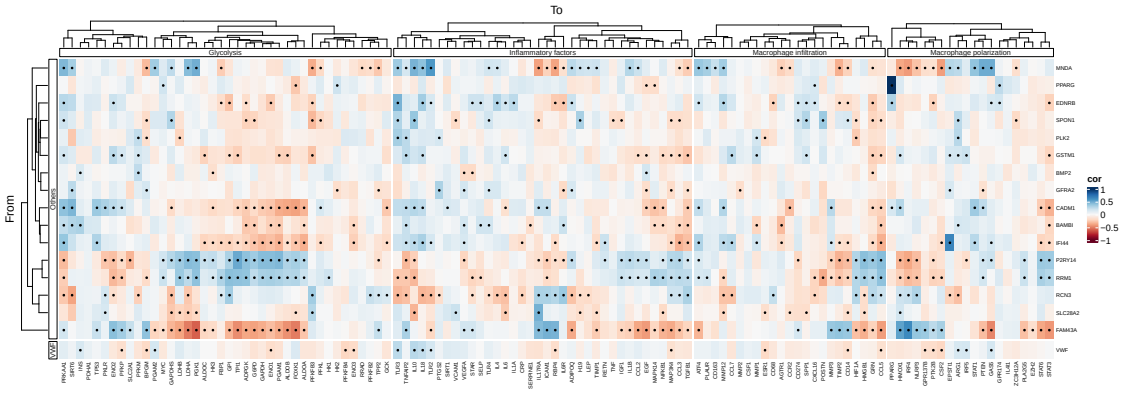

Figure 9: HUMAN correlation heatmap

Figure 10 (下方图) 为图 HUMAN correlation heatmap VWF significant 概览。

(对应文件为 `Figure+Table/HUMAN-correlation-heatmap-VWF-significant.pdf`)

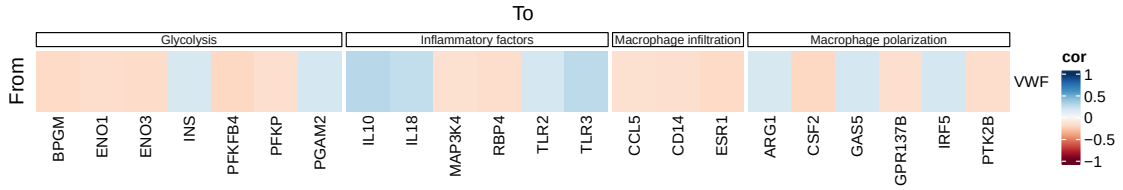

Figure 10: HUMAN correlation heatmap VWF significant

'HUMAN regression VWF significant' 数据已全部提供。

(对应文件为 `Figure+Table/HUMAN-regression-VWF-significant`)

> 注：文件夹 Figure+Table/HUMAN-regression-VWF-significant 共包含 4 个文件。
>
> 1. 1_Glycolysis.pdf
> 2. 2_Inflammatory factors.pdf
> 3. 3_Macrophage infiltration.pdf
> 4. 4_Macrophage polarization.pdf

Table 6 (下方表格) 为表格 HUMAN correlation 概览。

(对应文件为 `Figure+Table/HUMAN-correlation.csv`)

> 注：表格共有 2023 行 9 列，以下预览的表格可能省略部分数据；含有 17 个唯一 'From'。

1. cor: 皮尔逊关联系数，正关联或负关联。
2. pvalue: 显著性 P。
3. -log2(P.value): P 的对数转化。
4. significant: 显著性。
5. sign: 人为赋予的符号，参考 significant。

Table 6: HUMAN correlation

From	To	cor	pvalue	-log2(...	signif...	sign	Factors	Type
EDNRB	GAPDH	-0.14	0.0931	3.4250...	> 0.05	-	Glycol...	Others
PPARG	GAPDH	-0.06	0.4812	1.0552...	> 0.05	-	Glycol...	Others
CADM1	GAPDH	-0.26	9e-04	10.117...	< 0.001	**	Glycol...	Others
BMP2	GAPDH	-0.1	0.225	2.1520...	> 0.05	-	Glycol...	Others
SLC28A2	GAPDH	0.06	0.4919	1.0235...	> 0.05	-	Glycol...	Others
RRM1	GAPDH	0.33	0	16.609...	< 0.001	**	Glycol...	Others
BAMBI	GAPDH	-0.16	0.0796	3.6510...	> 0.05	-	Glycol...	Others
PLK2	GAPDH	-0.08	0.3648	1.4548...	> 0.05	-	Glycol...	Others
P2RY14	GAPDH	0.43	0	16.609...	< 0.001	**	Glycol...	Others
MNDA	GAPDH	-0.02	0.7616	0.3928...	> 0.05	-	Glycol...	Others
GSTM1	GAPDH	-0.16	0.073	3.7759...	> 0.05	-	Glycol...	Others
IFI44	GAPDH	-0.33	0	16.609...	< 0.001	**	Glycol...	Others
RCN3	GAPDH	0.12	0.148	2.7563...	> 0.05	-	Glycol...	Others
SPON1	GAPDH	-0.14	0.0882	3.5030...	> 0.05	-	Glycol...	Others
GFRA2	GAPDH	-0.07	0.3662	1.4492...	> 0.05	-	Glycol...	Others
...	...	...	...	...	...	...	...	...

Reference

1. Jablonski, K. A. *et al.* Novel markers to delineate murine m1 and m2 macrophages. *PloS one* **10**, (2015).

2. 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).

3. 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).

4. Ritchie, M. E. *et al.* Limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47 (2015).

5. Chen, Y., McCarthy, D., Ritchie, M., Robinson, M. & Smyth, G. EdgeR: Differential analysis of sequence read count data user's guide. 119.

6. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, (2021).

7. Stuart, T. *et al.* Comprehensive integration of single-cell data. *Cell* **177**, (2019).

8. Cao, Y., Wang, X. & Peng, G. SCSA: A cell type annotation tool for single-cell rna-seq data. *Frontiers in genetics* **11**, (2020).