

乙酰化酶分析筛选

2024-04-17

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 | 2 |
| 4 分析结果 | 2 |
| 5 结论 | 2 |
| 6 附：分析流程 | 2 |
| 6.1 MI targets | 2 |
| 6.2 MI mice DEGs | 4 |
| 6.2.1 数据来源 | 4 |
| 6.2.2 差异分析 | 4 |
| 6.2.3 基因映射 | 6 |
| 6.2.4 FKBP5 的表达 | 7 |
| 6.3 MI intersection (MI_key_DEGs) | 8 |
| 6.4 乙酰化酶 | 8 |
| 6.4.1 使用的乙酰化酶及其相关信息 | 8 |
| 6.4.2 筛选差异表达的乙酰化酶 (CoA_DEGs) | 9 |
| 6.5 其它候选 | 10 |
| 6.5.1 以 PPI 网络筛选与 CoA_DEGs 相关的 MI_key_DEGs | 10 |
| 6.5.2 关联分析 | 12 |
| 6.5.3 存在 PPI 关联且关联分析显著的组合 | 14 |
| 6.5.4 富集分析 | 14 |
| 6.5.5 CoA-XX-pathways | 15 |
| Reference | 17 |

List of Figures

| | |
|---|----|
| 1 Overall targets number of datasets | 2 |
| 2 MI MI 7d vs MI sham DEGs | 5 |
| 3 Intersection of MI DEGs with MI targets | 8 |
| 4 Intersection of All CoA with MI DEGs | 10 |
| 5 Filtered and formatted PPI network | 12 |

| | | |
|---|----------------------------------|----|
| 6 | MI correlation heatmap | 13 |
| 7 | GO enrichment | 15 |
| 8 | CoA XX GOpathways | 16 |

List of Tables

| | | |
|---|---|----|
| 1 | GeneCards used data | 3 |
| 2 | MI data MI 7d vs MI sham DEGs | 6 |
| 3 | Mapped genes | 7 |
| 4 | Fkbp5 expression | 8 |
| 5 | All protein of CoA | 9 |
| 6 | PPI annotation | 11 |
| 7 | MI significant correlation | 13 |
| 8 | PPI interact and significant correlated in MI | 14 |
| 9 | All candidates and enriched GO BP pathways | 16 |

1 摘要

1.1 需求

利用生物信息学分析结合已有文献资料，筛选并验证与 XX 相关的乙酰化酶 AA

具体要求为：利用开源数据库，筛选心肌梗死机体的心脏细胞中关键差异表达基因 XX 以及与乙酰化相关酶基因的关联性。

1. 因客户之前所做基因为 FKBP5，故 XX 初步定为 FKBP5。假设 FKBP5 在心肌梗死机体心肌细胞中高表达，抑制 FKBP5 后可缓解心肌梗死。
2. 乙酰化酶 AA 备选：去乙酰化酶 sirtuin 1 (SIRT1) 可以直接与 FKBP5 相互作用。
3. 若方案中的 AA 选择为 HDAC6 (客户之前发表过 LncRNA NORAD-HDAC6-H3K9 -VEGF)，那么方案中的 XX 选择不一定非要是 FKBP5，若有创新点的更好的基因也可。

1.2 结果

- 结合数据库 MI 靶点和 MI 小鼠数据集获取一批 MI 基因 Fig. 3
- 从 epiFactor 数据库获取乙酰酶 (CoA) (Tab. 5)，筛选了 MI 中为差异表达的 CoA (Fig. 4)。
- 根据 Fig. 3 和 Fig. 4 建立 PPI 网络 (有实验基础的蛋白物理直接互作)，见 Fig. 5
- 筛选 CoA 与 DEGs 显著关联的组合，Fig. 6, Tab. 7
- 筛选上述关系：存在 PPI 关联且关联分析显著的组合 Tab. 8
- 将上述 DEGs GO 富集分析，Fig. 7, BP 结果指向了 MI。
- 建立 CoA-XX-pathways 网络关系图，Fig. 8, 数据见 Tab. 9。
- 最后，推荐 CoA-XX 组合为：
 - CoA:BRCA1, DEG:FLNA
 - CoA:HDAC9, DEG:PIK3CG
 - 以上 DEG 相关 GO 通路：cardiac muscle contraction; coagulation; muscle system process; regulation of body fluid levels; striated muscle contraction; wound healing
- 其它候选见 Tab. 9

注：

- FKBP5 (Fkbp5) 在 MI 中属于显著高表达，见 Tab. 4。
- FKBP5 在 Fig. 3 被筛离。
- 尝试单独建立 PPI，未发现 SIRT1 与 FKBP5 的直接结合作用。

2 前言

3 材料和方法

3.1 材料

All used GEO expression data and their design:

- **GSE236374**: Nine 8-week-old male C57BL/6JR mice were included in the experiment. The experiment was divided into 3 groups. Each group contained 3 mice, 2 groups of which required surgery to make models, called...

3.2 方法

Mainly used method:

- The **biomart** was used for mapping genes between organism (e.g., mgi_symbol to hgnc_symbol)¹.
- R package **ClusterProfiler** used for gene enrichment analysis².
- Database **EpiFactors** used for screening epigenetic regulators³.
- GEO <https://www.ncbi.nlm.nih.gov/geo/> used for expression dataset aquisition.
- Databases of **DisGeNet**, **GeneCards**, **PharmGKB** used for collating disease related targets⁴⁻⁶.
- The Human Gene Database **GeneCards** used for disease related genes prediction⁵.
- R package **ClusterProfiler** used for GSEA enrichment².
- R package **Limma** and **edgeR** used for differential expression analysis^{7,8}.
- R package **STEINGdb** used for PPI network construction^{9,10}.
- R version 4.3.2 (2023-10-31); 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 MI targets

使用以下合集：

Figure 1 (下方图) 为图 Overall targets number of datasets 概览。

(对应文件为 **Figure+Table/Overall-targets-number-of-datasets.pdf**)

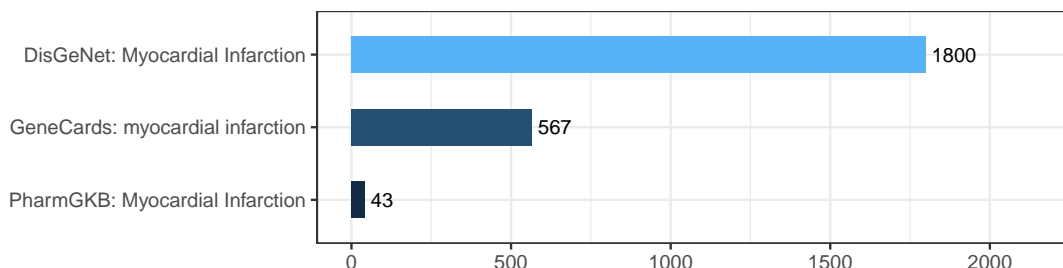


Figure 1: Overall targets number of datasets

The GeneCards data was obtained by querying :

myocardial infarction

Restrict (with quotes) :

TRUE

Filtering by Score: :

Score > 5

Table 1 (下方表格) 为表格 GeneCards used data 概览。

(对应文件为 **Figure+Table/GeneCards-used-data.xlsx**)

注：表格共有 567 行 7 列，以下预览的表格可能省略部分数据；含有 567 个唯一 ‘Symbol’。

Table 1: GeneCards used data

| Symbol | Description | Category | UniProt_ID | GIFtS | GC_id | Score |
|------------|---------------|----------------|------------|-------|-------------|-------|
| ACE | Angiotensi... | Protein Co... | P12821 | 60 | GC17P063477 | 75.08 |
| MIAT | Myocardial... | RNA Gene (...) | | 32 | GC22P026646 | 71.09 |
| F7 | Coagulatio... | Protein Co... | P08709 | 56 | GC13P113105 | 54.33 |
| ITGB3 | Integrin S... | Protein Co... | P05106 | 61 | GC17P112532 | 48.15 |
| LTA | Lymphotoxi... | Protein Co... | P01374 | 52 | GC06P134818 | 44.63 |
| OLR1 | Oxidized L... | Protein Co... | P78380 | 51 | GC12M029495 | 44.32 |
| PLAT | Plasminoge... | Protein Co... | P00750 | 58 | GC08M042174 | 39.78 |
| MCI2 | Myocardial... | Genetic Locus | | 4 | GC13U900611 | 39.39 |
| F13A1 | Coagulatio... | Protein Co... | P00488 | 56 | GC06M006144 | 39.35 |
| CDKN2B-AS1 | CDKN2B Ant... | RNA Gene (...) | | 31 | GC09P021994 | 39.31 |
| LGALS2 | Galectin 2 | Protein Co... | P05162 | 47 | GC22M037570 | 38.25 |
| MEF2A | Myocyte En... | Protein Co... | Q02078 | 54 | GC15P099565 | 38.14 |
| MIR499A | MicroRNA 499a | RNA Gene (...) | | 29 | GC20P034990 | 37.65 |
| ESR1 | Estrogen R... | Protein Co... | P03372 | 62 | GC06P151656 | 37.58 |
| MIR208B | MicroRNA 208b | RNA Gene (...) | | 27 | GC14M023417 | 35.34 |
| ... | ... | ... | ... | ... | ... | ... |

6.2 MI mice DEGs

6.2.1 数据来源

Data Source ID :

GSE236374

data__processing :

Raw reads were trimmed adaptor sequences and removed low-quality reads using TrimGalore with default settings

data__processing.1 :

Trimmed reads were aligned to the mm10 reference genome by STAR with default settings

data__processing.2 :

Read count extraction were performed using FeatureCounts

data__processing.3 :

Assembly: mm10

(Others) :

...

(上述信息框内容已保存至 `Figure+Table/GSE236374-content`)

6.2.2 差异分析

- MI-7d (7 day) vs Control

Figure 2 (下方图) 为图 MI MI 7d vs MI sham DEGs 概览。

(对应文件为 `Figure+Table/MI-MI-7d-vs-MI-sham-DEGs.pdf`)

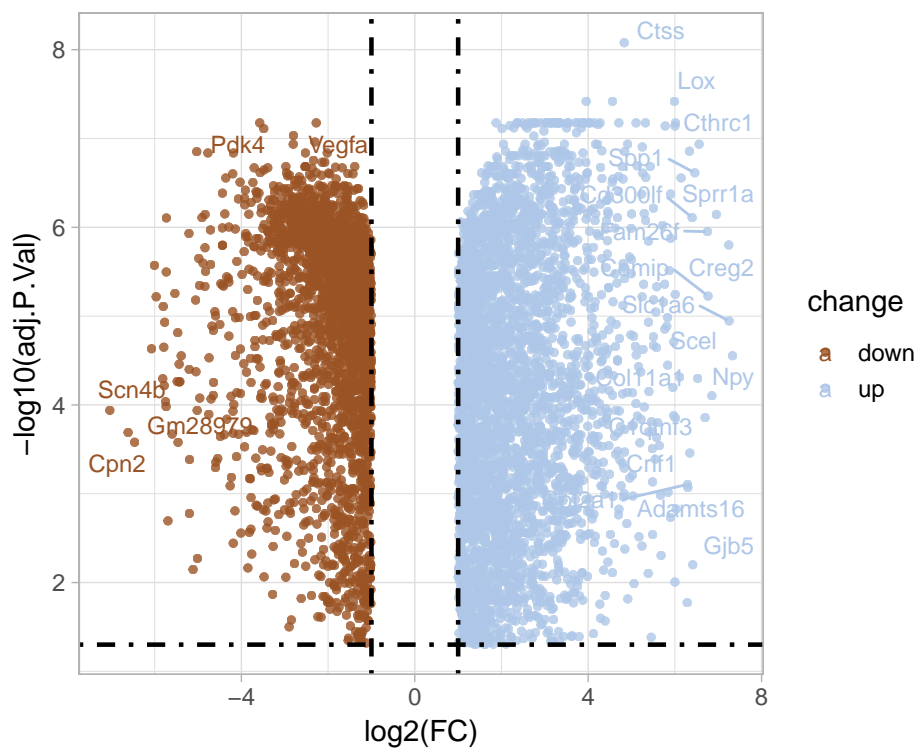


Figure 2: MI MI 7d vs MI sham DEGs

adj.P.Val cut-off :

0.05

Log2(FC) cut-off :

1

(上述信息框内容已保存至 Figure+Table/MI-MI-7d-vs-MI-sham-DEGs-content)

Table 2 (下方表格) 为表格 MI data MI 7d vs MI sham DEGs 概览。

(对应文件为 Figure+Table/MI-data-MI-7d-vs-MI-sham-DEGs.csv)

注：表格共有 5724 行 8 列，以下预览的表格可能省略部分数据；含有 5724 个唯一 ‘Genesymbol’。

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: MI data MI 7d vs MI sham DEGs

| rownames | Genesy... | logFC | AveExpr | t | P.Value | adj.P.Val | B |
|----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 7514 | Ctss | 4.8320... | 7.4632... | 50.601... | 5.8181... | 8.3472... | 19.814... |
| 14679 | Adamts2 | 3.9541... | 7.5930... | 37.675... | 9.6997... | 3.8153... | 17.557... |
| 23411 | Col14a1 | 4.5612... | 7.5634... | 37.311... | 1.0637... | 3.8153... | 17.437... |
| 11851 | Lox | 5.9882... | 7.0429... | 37.907... | 9.1490... | 3.8153... | 17.419... |
| 21619 | Fstl1 | 3.9252... | 9.4422... | 33.841... | 2.6934... | 6.6702... | 16.550... |
| 1261 | Ctsh | 2.6147... | 6.6709... | 31.959... | 4.6403... | 6.6702... | 16.144... |
| 13487 | Pla2g7 | 4.0298... | 4.6625... | 32.933... | 3.4885... | 6.6702... | 16.129... |
| 22176 | Laptm5 | 3.3558... | 6.9162... | 31.874... | 4.7596... | 6.6702... | 16.105... |
| 1490 | Sparc | 3.2522... | 11.160... | 32.579... | 3.8660... | 6.6702... | 16.079... |
| 6315 | Hexb | 3.1220... | 6.3869... | 31.264... | 5.7173... | 6.6702... | 15.929... |
| 5004 | Ctsz | 3.0952... | 6.9421... | 30.777... | 6.6372... | 6.6702... | 15.790... |
| 21174 | Fbln5 | 3.7685... | 7.2452... | 30.367... | 7.5384... | 6.6702... | 15.656... |
| 1805 | Litaf | 2.3676... | 5.9412... | 30.219... | 7.8956... | 6.6702... | 15.624... |
| 12260 | Nckap1l | 3.3359... | 5.8304... | 29.954... | 8.5853... | 6.6702... | 15.519... |
| 3893 | Gusb | 2.3568... | 6.0931... | 29.740... | 9.1894... | 6.6702... | 15.480... |
| ... | ... | ... | ... | ... | ... | ... | ... |

6.2.3 基因映射

将小鼠基因映射到人类

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 5274 行 2 列，以下预览的表格可能省略部分数据；含有 5123 个唯一 ‘mgi_symbol’；含有 5146 个唯一 ‘hgnc_symbol’。

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

Table 3: Mapped genes

| mgc_symbol | hgnc_symbol |
|------------|-------------|
| Tmsb4x | TMSB4Y |
| Hopx | HOPX |
| Cyth4 | CYTH4 |
| Col6a2 | COL6A2 |
| Pacsin2 | PACsin2 |
| Fbln1 | FBLN1 |
| Sh3bp2 | SH3BP2 |
| Abcg1 | ABCG1 |
| Mipep | MIPEP |
| Itgb2 | ITGB2 |
| Pmepa1 | PMEPA1 |
| Maged2 | MAGED2 |
| Postn | POSTN |
| Slc39a6 | SLC39A6 |
| Sirpa | SIRPG |
| ... | ... |

6.2.4 FKBP5 的表达

FKBP5 (Fkbp5) 在 MI 中属于显著高表达。

Table 4 (下方表格) 为表格 Fkbp5 expression 概览。

(对应文件为 `Figure+Table/Fkbp5-expression.csv`)

注：表格共有 1 行 10 列，以下预览的表格可能省略部分数据；含有 1 个唯一 ‘hgnc_symbol’。

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

Table 4: Fkbp5 expression

| hgnc_s... | mgis_sy... | rownames | Genesy... | logFC | AveExpr | t | P.Value | adj.P.Val | B |
|-----------|------------|----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| FKBP5 | Fkbp5 | 9124 | Fkbp5 | 1.5635... | 5.3072... | 5.7027... | 0.0002... | 0.0005... | 0.0172... |

6.3 MI intersection (MI_key_DEGs)

Figure 3 (下方图) 为图 Intersection of MI DEGs with MI targets 概览。

(对应文件为 Figure+Table/Intersection-of-MI-DEGs-with-MI-targets.pdf)

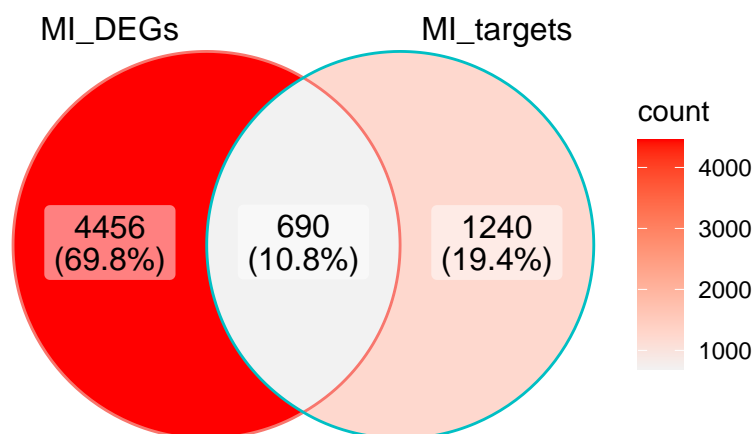


Figure 3: Intersection of MI DEGs with MI targets

Intersection :

ABCG1, ITGB2, POSTN, EGLN3, PPARGC1A, LTBP2, CYBB, C3AR1, THBS1, SERPINE1, CLU, SFRP2, TGFB3, IGFBP4, TNC, LCP1, GAS6, CTSZ, HPGDS, BGN, VLDLR, GUCY1A1, CYP4F3, LIPA, NCAM1, GLA, HLA-DMB, FERMT3, LGALS3, TLR2, MMP2, GPNMB, CYBA, ALCAM, KDR, TNNT3, ARNTL, IGFBP7, ANPEP, PPM1L, TNFRSF1B, SERPINF1, ...

(上述信息框内容已保存至 Figure+Table/Intersection-of-MI-DEGs-with-MI-targets-content)

6.4 乙酰化酶

6.4.1 使用的乙酰化酶及其相关信息

Table 5 (下方表格) 为表格 All protein of CoA 概览。

(对应文件为 Figure+Table/All-protein-of-CoA.xlsx)

注：表格共有 145 行 25 列，以下预览的表格可能省略部分数据；含有 142 个唯一‘HGNC_symbol’。

Table 5: All protein of CoA

| HGNC_s... | Modifi... | Id | Status | HGNC_ID | HGNC_name | GeneID | UniPro.....8 | UniPro.....9 | Domain |
|-----------|-----------|-----|--------|---------|-----------|--------|--------------|--------------|-----------|
| ARID4A | Histon... | 36 | # | 9885 | AT ric... | 5926 | P29374 | ARI4A_... | ARID P... |
| ARID4B | Histon... | 37 | # | 15550 | AT ric... | 51742 | Q4LE39 | ARI4B_... | ARID P... |
| ATF2 | Histon... | 49 | # | 784 | activa... | 1386 | P15336 | ATF2_H... | bZIP_1... |
| ATXN7 | Histon... | 55 | # | 10560 | ataxin 7 | 6314 | O15265 | ATX7_H... | Pfam-B... |
| BANP | Histon... | 62 | # | 13450 | BTG3 a... | 54971 | Q8N9N5 | BANP_H... | BEN PF... |
| BAZ2A | Histon... | 67 | # | 962 | bromod... | 11176 | Q9UIF9 | BAZ2A_... | Bromod... |
| BCORL1 | Histon... | 70 | # | 25657 | BCL6 c... | 63035 | Q5H9F3 | BCORL_... | Ank_2 ... |
| BRCA1 | Histon... | 73 | # | 1100 | breast... | 672 | P38398 | BRCA1_... | BRCT P... |
| BRCA2 | Histon... | 74 | # | 1101 | breast... | 675 | P51587 | BRCA2_... | BRCA-2... |
| BRMS1L | Histon... | 86 | # | 20512 | breast... | 84312 | Q5PSV4 | BRM1L_... | Sds3 P... |
| BRPF3 | Histon... | 88 | # | 14256 | bromod... | 27154 | Q9ULD4 | BRPF3_... | Bromod... |
| CDY1 | Histon... | 115 | # | 1809 | chromo... | 9085 | Q9Y6F8 | CDY1_H... | Chromo... |
| CDY1B | Histon... | 116 | # | 23920 | chromo... | 253175 | Q9Y6F8 | CDY1_H... | Chromo... |
| CDY2A | Histon... | 117 | # | 1810 | chromo... | 9426 | Q9Y6F7 | CDY2_H... | Chromo... |
| CDY2B | Histon... | 118 | # | 23921 | chromo... | 203611 | Q9Y6F7 | CDY2_H... | Chromo... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |

6.4.2 筛选差异表达的乙酰化酶 (CoA_DEGs)

使用 MI 数据 (6.2) 的 DEGs，筛选差异表达的乙酰化酶。

以 `mg_i_symbol` 取交集。

Figure 4 (下方图) 为图 Intersection of All CoA with MI DEGs 概览。

(对应文件为 `Figure+Table/Intersection-of-All-CoA-with-MI-DEGs.pdf`)

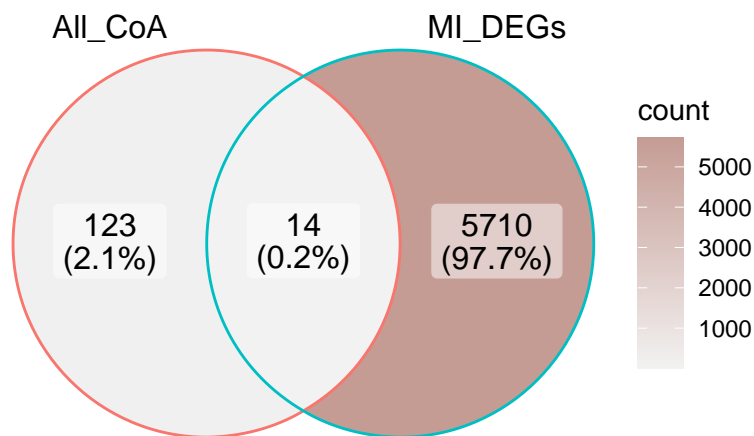


Figure 4: Intersection of All CoA with MI DEGs

Intersection :

Brcal, Eid1, Eid2b, Hdac11, Hdac9, Hif1an, Jdp2, Morf4l2, Ncoa1, Nsl1, Sirt7, Smarca1, Taf7, Zbtb16

(上述信息框内容已保存至 Figure+Table/Intersection-of-All-CoA-with-MI-DEGs-content)

6.5 其它候选

6.5.1 以 PPI 网络筛选与 CoA_DEGs 相关的 MI_key_DEGs

根据 Fig. 3 和 Fig. 4 建立 PPI 网络 (有实验基础的蛋白物理直接互动)。

STRINGdb network type: :

physical

Filter experiments score: :

At least score 100

Filter textmining score: :

At least score 0

Table 6 (下方表格) 为表格 PPI annotation 概览。

(对应文件为 Figure+Table/PPI-annotation.csv)

注：表格共有 1364 行 10 列，以下预览的表格可能省略部分数据；含有 381 个唯一 ‘from’。

1. experiments: 相关实验。

Table 6: PPI annotation

| from | to | homology | experi.....4 | experi.....5 | database | databa... | textmi.....8 | textmi.....9 | ... |
|----------|----------|----------|--------------|--------------|----------|-----------|--------------|--------------|-----|
| TNFRSF1A | RIPK3 | 0 | 292 | 0 | 0 | 0 | 473 | 0 | ... |
| DCN | PLAT | 0 | 205 | 0 | 0 | 0 | 0 | 0 | ... |
| DCN | TGFB1 | 0 | 457 | 0 | 500 | 0 | 979 | 60 | ... |
| MMP2 | TGFB1 | 0 | 548 | 0 | 0 | 0 | 118 | 0 | ... |
| PLAT | SERPINE1 | 0 | 955 | 0 | 700 | 0 | 982 | 0 | ... |
| MYH9 | ACTA2 | 0 | 205 | 97 | 900 | 0 | 0 | 91 | ... |
| MMP2 | COL1A1 | 0 | 292 | 0 | 0 | 0 | 0 | 0 | ... |
| TGFB1 | VDR | 0 | 292 | 0 | 0 | 0 | 0 | 0 | ... |
| COL1A1 | VDR | 0 | 292 | 0 | 0 | 0 | 0 | 0 | ... |
| MMP2 | LOX | 0 | 238 | 0 | 0 | 0 | 0 | 0 | ... |
| COL1A1 | LOX | 0 | 230 | 0 | 0 | 0 | 0 | 0 | ... |
| COL1A1 | SPARC | 0 | 457 | 0 | 0 | 0 | 89 | 90 | ... |
| VDR | IL12B | 0 | 292 | 0 | 0 | 0 | 0 | 0 | ... |
| ACTA2 | CTSD | 0 | 229 | 0 | 0 | 0 | 0 | 0 | ... |
| VDR | EGR1 | 0 | 292 | 0 | 0 | 0 | 0 | 0 | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |

获取 CoA -> DEGs 的网络：

Figure 5 (下方图) 为图 Filtered and formatted PPI network 概览。

(对应文件为 Figure+Table/Filtered-and-formated-PPI-network.pdf)

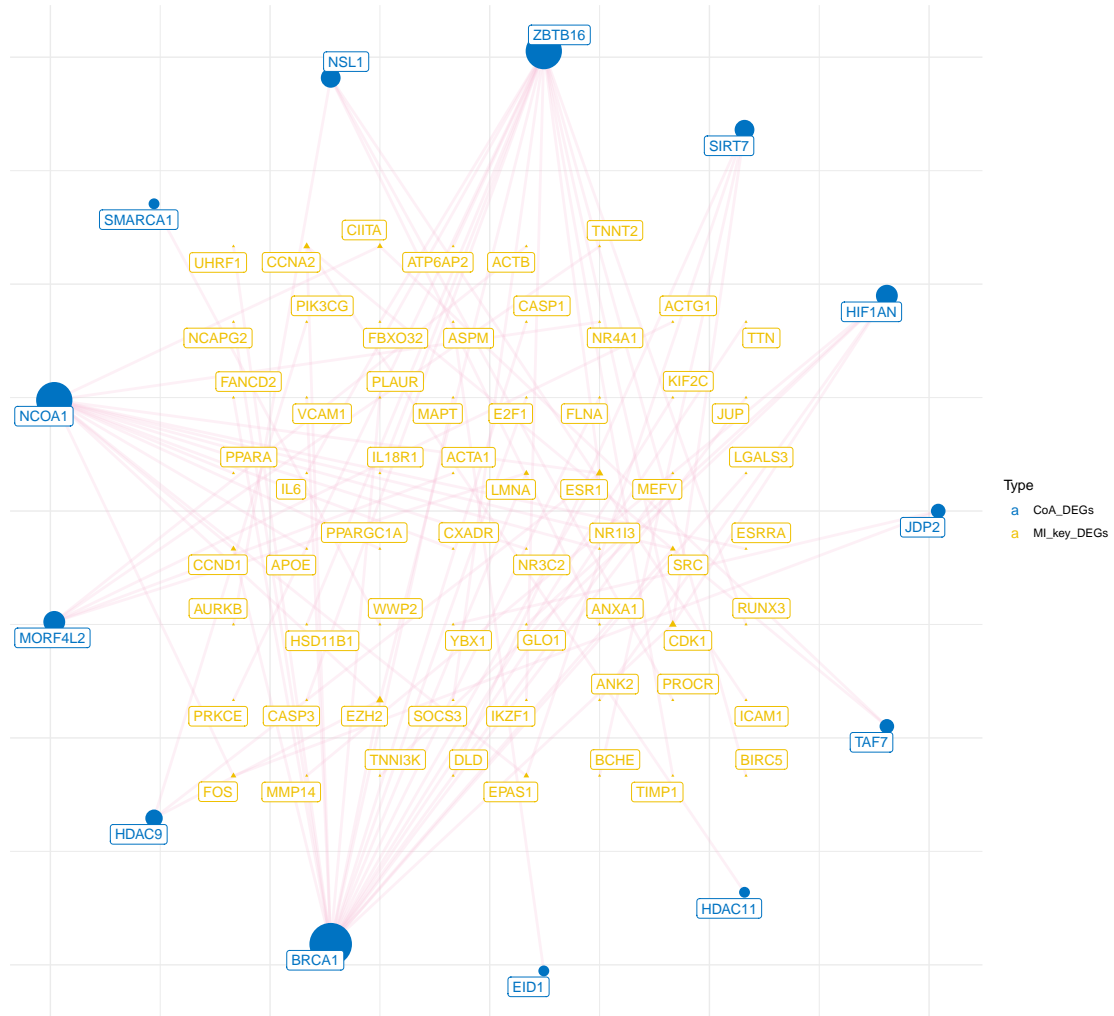


Figure 5: Filtered and formatted PPI network

6.5.2 关联分析

根据 Fig. 5, 以小鼠数据集 (6.2) 进行关联分析。

Figure 6 (下方图) 为图 MI correlation heatmap 概览。

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

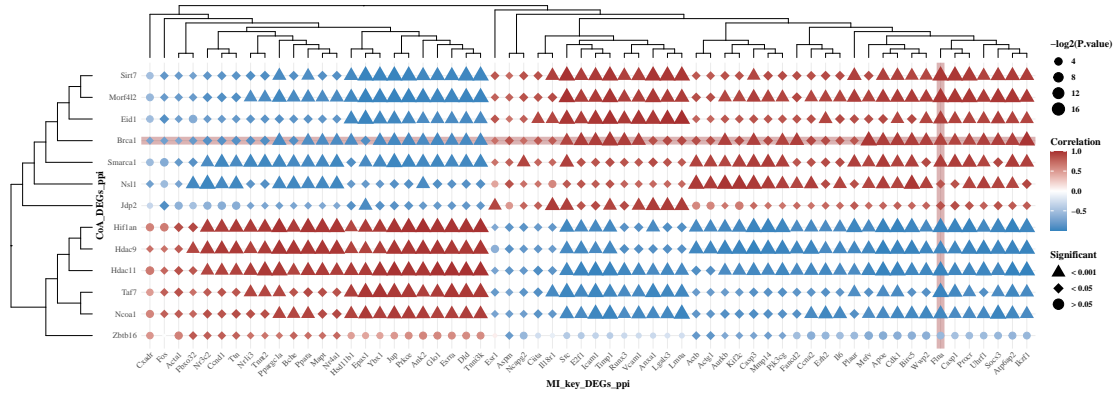


Figure 6: MI correlation heatmap

Table 7 (下方表格) 为表格 MI significant correlation 概览。

(对应文件为 **Figure+Table/MI-significant-correlation.csv**)

注：表格共有 738 行 7 列，以下预览的表格可能省略部分数据；含有 13 个唯一 ‘CoA_DEGs_ppi’。

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 7: MI significant correlation

| CoA_DEGs_ppi | MI_key_DEG... | cor | pvalue | -log2(P.va... | significant | sign |
|--------------|---------------|-------|--------|---------------|-------------|------|
| Morf4l2 | Ppargc1a | -0.95 | 1e-04 | 13.2877123... | < 0.001 | ** |
| Hdac9 | Ppargc1a | 0.98 | 0 | 16.6096404... | < 0.001 | ** |
| Sirt7 | Ppargc1a | -0.92 | 5e-04 | 10.9657842... | < 0.001 | ** |
| Nsl1 | Ppargc1a | -0.94 | 2e-04 | 12.2877123... | < 0.001 | ** |
| Taf7 | Ppargc1a | 0.9 | 0.001 | 9.96578428... | < 0.001 | ** |
| Ncoa1 | Ppargc1a | 0.91 | 7e-04 | 10.4803574... | < 0.001 | ** |
| Jdp2 | Ppargc1a | -0.74 | 0.0217 | 5.52616114... | < 0.05 | * |
| Hif1an | Ppargc1a | 0.99 | 0 | 16.6096404... | < 0.001 | ** |
| Brcal | Ppargc1a | -0.93 | 3e-04 | 11.7027498... | < 0.001 | ** |
| Smarca1 | Ppargc1a | -0.95 | 1e-04 | 13.2877123... | < 0.001 | ** |
| Hdac11 | Ppargc1a | 0.96 | 0 | 16.6096404... | < 0.001 | ** |
| Eid1 | Ppargc1a | -0.87 | 0.0024 | 8.70274987... | < 0.05 | * |
| Zbtb16 | Ppargc1a | 0.67 | 0.0483 | 4.37183300... | < 0.05 | * |
| Morf4l2 | Il18r1 | 0.88 | 0.0018 | 9.11778737... | < 0.05 | * |
| Hdac9 | Il18r1 | -0.78 | 0.013 | 6.26534456... | < 0.05 | * |

| CoA_DEGs_ppi | MI_key_DEG... | cor | pvalue | -log2(P.va... | significant | sign |
|--------------|---------------|-----|--------|---------------|-------------|------|
| ... | ... | ... | ... | ... | ... | ... |

6.5.3 存在 PPI 关联且关联分析显著的组合

结合 6.5.1 和 6.5.2 筛选 CoA 与 XX

Table 8 (下方表格) 为表格 PPI interact and significant correlated in MI 概览。

(对应文件为 **Figure+Table/PPI-interact-and-significant-correlated-in-MI.csv**)

注：表格共有 64 行 9 列，以下预览的表格可能省略部分数据；含有 13 个唯一 ‘CoA_DEGs_ppi’。

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 8: PPI interact and significant correlated in MI

| CoA_DE... | MI_key... | cor | pvalue | -log2(... | signif... | sign | CoA_hg... | DEG_hg... |
|-----------|-----------|-------|--------|-----------|-----------|------|-----------|-----------|
| Brcal | Casp1 | 0.91 | 8e-04 | 10.287... | < 0.001 | ** | BRCA1 | CASP1 |
| Brcal | Ccna2 | 0.9 | 0.0011 | 9.8282... | < 0.05 | * | BRCA1 | CCNA2 |
| Brcal | Ccnd1 | -0.87 | 0.0024 | 8.7027... | < 0.05 | * | BRCA1 | CCND1 |
| Brcal | Cdk1 | 0.95 | 1e-04 | 13.287... | < 0.001 | ** | BRCA1 | CDK1 |
| Brcal | E2f1 | 0.95 | 1e-04 | 13.287... | < 0.001 | ** | BRCA1 | E2F1 |
| Brcal | Esr1 | 0.7 | 0.0356 | 4.8119... | < 0.05 | * | BRCA1 | ESR1 |
| Brcal | Ezh2 | 0.88 | 0.002 | 8.9657... | < 0.05 | * | BRCA1 | EZH2 |
| Brcal | Fancd2 | 0.94 | 2e-04 | 12.287... | < 0.001 | ** | BRCA1 | FANCD2 |
| Brcal | Flna | 0.92 | 4e-04 | 11.287... | < 0.001 | ** | BRCA1 | FLNA |
| Brcal | Jup | -0.94 | 2e-04 | 12.287... | < 0.001 | ** | BRCA1 | JUP |
| Brcal | Kif2c | 0.89 | 0.0011 | 9.8282... | < 0.05 | * | BRCA1 | KIF2C |
| Brcal | Lgals3 | 0.89 | 0.0013 | 9.5872... | < 0.05 | * | BRCA1 | LGALS3 |
| Brcal | Lmna | 0.89 | 0.0015 | 9.3808... | < 0.05 | * | BRCA1 | LMNA |
| Brcal | Mapt | -0.91 | 8e-04 | 10.287... | < 0.001 | ** | BRCA1 | MAPT |
| Brcal | Mefv | 0.97 | 0 | 16.609... | < 0.001 | ** | BRCA1 | MEFV |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |

6.5.4 富集分析

将 Tab. 8 中的 DEGs 进行富集分析

Figure 7 (下方图) 为图 GO enrichment 概览。

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

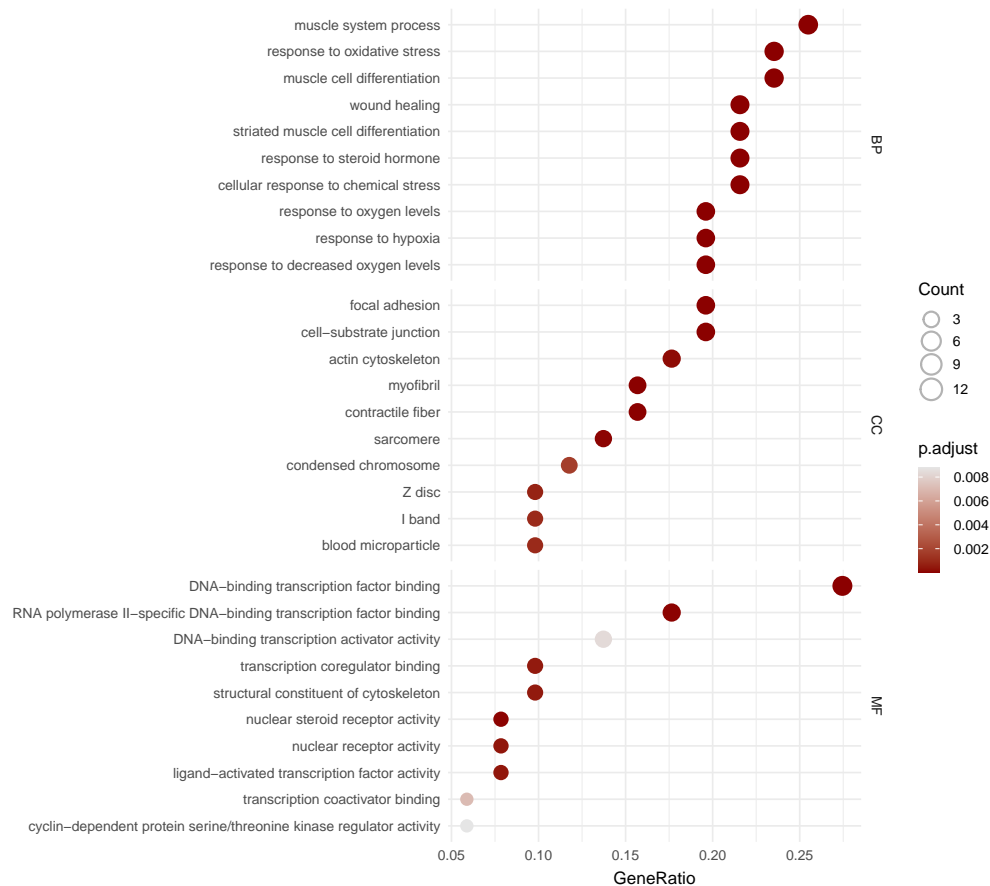


Figure 7: GO enrichment

6.5.5 CoA-XX-pathways

Figure 8 (下方图) 为图 CoA XX Gopathways 概览。

(对应文件为 **Figure+Table/CoA-XX-Gopathways.pdf**)

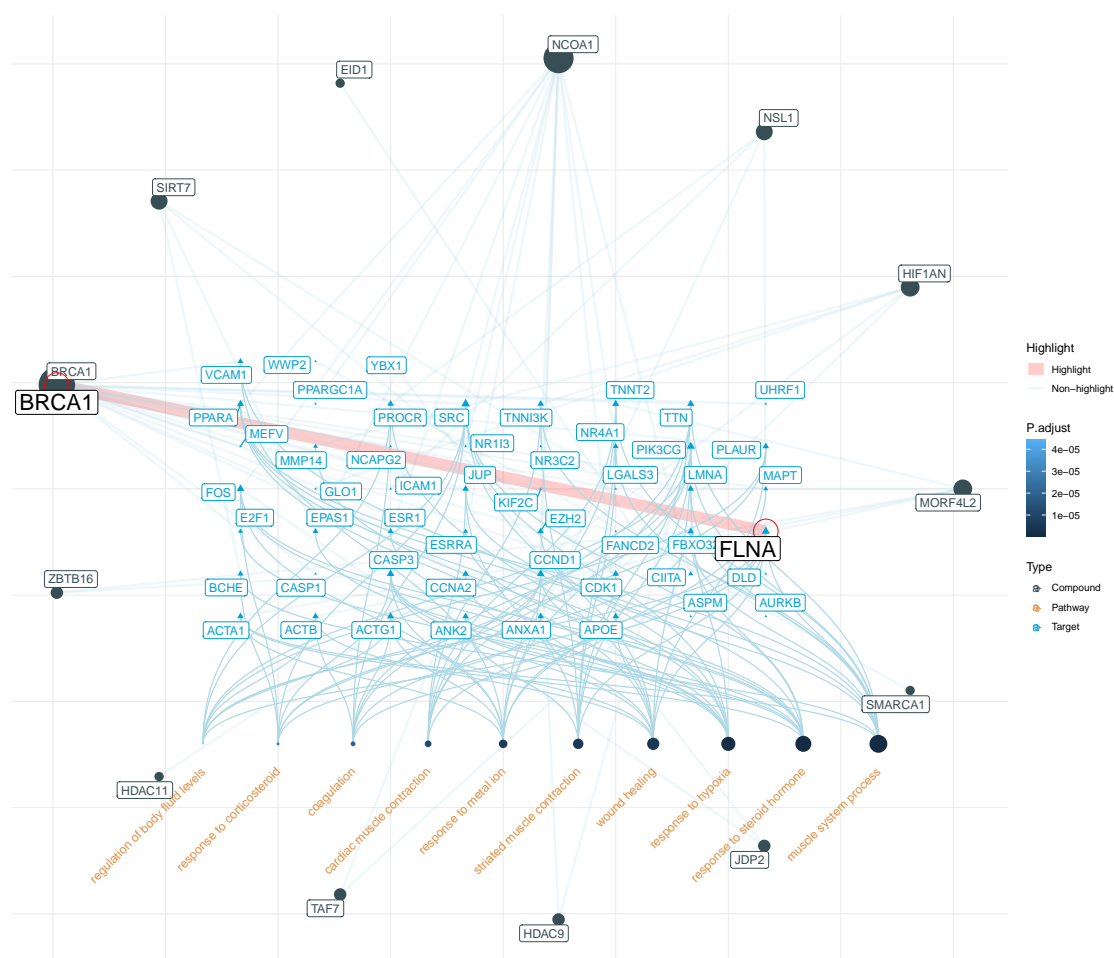


Figure 8: CoA XX GOpathways

Table 9 (下方表格) 为表格 All candidates and enriched GO BP pathways 概览。

(对应文件为 **Figure+Table/All-candidates-and-enriched-GO-BP-pathways.csv**)

注：表格共有 64 行 4 列，以下预览的表格可能省略部分数据；含有 13 个唯一 ‘CoA_hgnc_symbol’；含有 51 个唯一 ‘DEG_hgnc_symbol’。

Table 9: All candidates and enriched GO BP pathways

| CoA_hgnc_symbol | DEG_hgnc_symbol | Hit_pathway_number | Enriched_pathways |
|-----------------|-----------------|--------------------|----------------------|
| BRCA1 | FLNA | 6 | cardiac muscle co... |
| BRCA1 | SRC | 6 | coagulation; regu... |
| HDAC9 | PIK3CG | 6 | cardiac muscle co... |
| NCOA1 | SRC | 6 | coagulation; regu... |
| ZBTB16 | CASP3 | 5 | response to corti... |
| BRCA1 | CCND1 | 4 | regulation of bod... |
| BRCA1 | TTN | 4 | cardiac muscle co... |

| CoA_hgnc_symbol | DEG_hgnc_symbol | Hit_pathway_number | Enriched_pathways |
|-----------------|-----------------|--------------------|----------------------|
| JDP2 | FOS | 4 | response to corti... |
| MORF4L2 | ACTG1 | 4 | coagulation; regu... |
| MORF4L2 | TNNT2 | 4 | cardiac muscle co... |
| NCOA1 | CCND1 | 4 | regulation of bod... |
| NCOA1 | FOS | 4 | response to corti... |
| NCOA1 | PPARA | 4 | muscle system pro... |
| BRCA1 | JUP | 3 | cardiac muscle co... |
| BRCA1 | PLAUR | 3 | coagulation; regu... |
| ... | ... | ... | ... |

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. Wu, T. *et al.* ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* **2**, (2021).
3. Marakulina, D. *et al.* EpiFactors 2022: Expansion and enhancement of a curated database of human epigenetic factors and complexes. *Nucleic acids research* **51**, D564–D570 (2023).
4. Piñero, J. *et al.* The disgenet knowledge platform for disease genomics: 2019 update. *Nucleic Acids Research* (2019) doi:10.1093/nar/gkz1021.
5. 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).
6. Barbarino, J. M., Whirl-Carrillo, M., Altman, R. B. & Klein, T. E. PharmGKB: A worldwide resource for pharmacogenomic information. *Wiley interdisciplinary reviews. Systems biology and medicine* **10**, (2018).
7. Ritchie, M. E. *et al.* Limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47 (2015).
8. Chen, Y., McCarthy, D., Ritchie, M., Robinson, M. & Smyth, G. EdgeR: Differential analysis of sequence read count data user’s guide. 119.
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).