

网络药理学和表观遗传学修饰筛选靶点

2024-03-25

LiChuang Huang



@ 立效研究院

Contents

1 摘要	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	1
5 结论	2
6 附：分析流程	2
6.1 网络药理学分析	2
6.1.1 成分	2
6.1.2 成分靶点	4
6.1.3 疾病靶点	5
6.2 表观遗传学蛋白	5
6.3 筛选表观遗传学靶点	6
6.3.1 复方成分与表观遗传学靶点	6
6.4 脓毒症肠损伤的 GEO 数据分析	8
6.4.1 数据来源	8
6.4.2 fastp 质控	9
6.4.3 RNA 定量	9
6.4.4 差异分析	10
6.4.5 将小鼠基因 symbol 映射到人类 (hgnc_symbol)	11
6.4.6 富集分析 (GSEA)	12
6.5 复方成分表观遗传学靶点的通路调控	14
6.5.1 富集表观修饰蛋白的通路	14
6.5.2 表观修饰靶点	15
6.5.3 上下游	16
6.5.4 相关成分	18
6.6 分子对接	20
6.6.1 第一批对接	20
6.6.2 以口服利用度筛选其他成分	22
6.6.3 第二批对接	22
6.6.4 对接可视化 (Top 3)	24
Reference	27

List of Figures

1	Intersection of all compounds	4
2	Overall targets number of datasets	5
3	SLBJ network pharmacology of epigenetic target	7
4	SII model vs control DEGs	11
5	KEGG enrichment	13
6	KEGG enrichment with enriched genes	14
7	GSEA plot of the pathways	15
8	Hsa04920 visualization	16
9	Hsa04010 visualization	17
10	Hsa04657 visualization	18
11	SLBJ network pharmacology Target epigenetic related pathway	19
12	FIRST Overall combining Affinity	21
13	HOB 20 prediction	22
14	Overall combining Affinity	24
15	Docking 72326 into GADD45B	25
16	Docking 12313579 into GADD45B	26
17	Docking 5316891 into GADD45B	27

List of Tables

1	Herbs information	2
2	Components of Herbs	3
3	Tables of Herbs compounds and targets	4
4	All protein of epigenetic regulators	6
5	SLBJ Herbs compounds and targets of epigenetic target	7
6	SII metadata	9
7	Quantification	10
8	Mapped DEGs	12
9	KEGG enrichment data	13
10	SLBJ network pharmacology Target epigenetic related pathway data	19

1 摘要

网络药理学分析参苓白术散治疗脓毒症肠损伤的药物活性成分 XXX，对接下游的靶点 YYY（表观遗传学修饰蛋白），YYY 可以对靶点 ZZZ 进行表观遗传学修饰。YYY-ZZZ 的机制和调控脓毒症肠损伤相关。

最终筛得成分：Betulin (CID: 72326) (见 Fig. 14)，对接靶点 GADD45B (表观遗传学蛋白)，根据 KEGG (MARK) 通路图 (Fig. 9)，其可作用于 MAP3K4 和 TAOK3，与脓毒症肠损伤有关。

2 前言

3 材料和方法

3.1 材料

All used GEO expression data and their design:

- **GSE202261**: A total of 5 C57BL/6 mice were randomly divided into 2 groups: CLP group (n=3), Sham group (n=2). The disinfected abdomen was incised 1 cm at midline to expose the cecum. Silk thread was subsequentl...

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., `mgc_symbol` to `hgnc_symbol`)¹.
- R package **ClusterProfiler** used for gene enrichment analysis².
- Database **EpiFactors** used for screening epigenetic regulators³.
- **Fastp** used for Fastq data preprocessing⁴.
- GEO <https://www.ncbi.nlm.nih.gov/geo/> used for expression dataset acquisition.
- Databases of **DisGeNet**, **GeneCards**, **PharmGKB** used for collating disease related targets⁵⁻⁷.
- R package **ClusterProfiler** used for GSEA enrichment².
- Website **HERB** <http://herb.ac.cn/> used for data source⁸.
- Python tool of **HOB** was used for prediction of human oral bioavailability⁹.
- **Kallisto** used for RNA-seq mapping and quantification¹⁰.
- R package **Limma** and **edgeR** used for differential expression analysis^{11,12}.
- The CLI tools of **AutoDock vina** and **ADFR** software used for auto molecular docking¹³⁻¹⁷.
- R package **ChemmineR** used for similar chemical compounds clustering¹⁸.
- 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

- 从 HERB 数据库获取复方的成分信息和靶点信息，整理如 Tab. 3。

- 从多个数据库获取脓毒症肠损伤 (SH) 相关靶点，见 Fig. 2
- 表观遗传学蛋白的获取通过 EpiFactors 数据库。
- 在复方中存在的表观遗传学相关靶点见 Fig. 3
- 为了进一步筛选与 SH 相关的通路和靶点，以 GEO 的 SH 数据集做了差异分析 (Fig. 4)。
- 该数据集源于小鼠，这里将其映射到人类基因，然后富集分析 Fig. 5
- 在富集的通路中筛选包含表观修饰相关靶点，见 Fig. 6
- 通路具体可见 6.5.2, 6.5.3, 结果上述网络药理学，可发现相关化学成分为 Fig. 11
- 对 Fig. 11 所示的成分与靶点关系进行分子对接，对接结果见 Fig. 12。
- 由于 Fig. 12 所示对接能量均过高，这里随后尝试挖掘复方中其它可能作用于其相关靶点的化学成分。
- 复方中的成分复杂，为了减少过度的分子对接计算量，以计算 HOB 和化学结果相似聚类的方式，筛选了少数一批化学成分用以分子对接。
- 对接结果见 Fig. 14。其中，Betulin (CID: 72326) 对接 GADD45B 蛋白所需能量最少，对接可视化见 Fig. 15
- 联系 Fig. 9，可知 GADD45B 下游调控的靶点。其中，TAOK3 是 SH 中上调的基因，GADD45B-TAOK3 的作用，可能是最佳结果，对应成分为 Betulin (CID: 72326)。

5 结论

最终筛得成分：Betulin (CID: 72326) (见 Fig. 14)，对接靶点 GADD45B (表观遗传学蛋白)，根据 KEGG (MARK) 通路图 (Fig. 9)，其可作用于 MAP3K4 和 TAOK3，与脓毒症肠损伤有关。

6 附：分析流程

6.1 网络药理学分析

复方成分和靶点数据来源于 HERB 数据库。

6.1.1 成分

Table 1 (下方表格) 为表格 Herbs information 概览。

(对应文件为 **Figure+Table/Herbs-information.xlsx**)

注：表格共有 10 行 18 列，以下预览的表格可能省略部分数据；表格含有 10 个唯一 ‘Herb_’。

Table 1: Herbs information

Herb_	Herb_p...	Herb_c...	Herb_e...	Herb_l...	Proper...	Meridians	UsePart	Function	Indica...
HERB00...	BAI BI...	白扁豆	White ...	Semen ...	Minor ...	Spleen...	NA	To inv...	Treatm...
HERB00...	BAI ZHU	白术	rhizom...	Rhizom...	Warm; ...	Spleen...	root	To inv...	Spleen...
HERB00...	FU LING	茯苓	Indian...	Poria	Mild; ...	Spleen...	sclero...	To cau...	Neuras...
HERB00...	GAN CAO	甘草	Root o...	Radix ...	Mild; ...	Lung; ...	root a...	To rei...	1. Its...

Herb_	Herb_p...	Herb_c...	Herb_e...	Herb_l...	Proper...	Meridians	UsePart	Function	Indica...
HERB00...	JIE GENG	桔梗	Platyc...	Radix ...	Mild; ...	Lung	root	To rel...	Cough ...
HERB00...	LIAN ZI	莲子	Hindu ...	Nelumb...	Mild; ...	Spleen...	seed	To sup...	Chroni...
HERB00...	REN SHEN	人参	Ginseng	Radix ...	Minor ...	Lung; ...	root	To rei...	Qi vac...
HERB00...	SHA REN	砂仁	Villou...	Fructu...	Warm; ...	Spleen...	ripe f...	To eli...	Abdomi...
HERB00...	SHAN YAO	山药	Common...	Rhizom...	Mild; ...	Lung; ...	rhizome	To rep...	Reduce...
HERB00...	YI YI REN	薏苡仁	seed o...	semen ...	Minor ...	Lung; ...	seed	1. To ...	Edema,...

Table 2 (下方表格) 为表格 Components of Herbs 概览。

(对应文件为 **Figure+Table/Components-of-Herbs.xlsx**)

注：表格共有 1870 行 4 列，以下预览的表格可能省略部分数据；表格含有 10 个唯一 'herb_id'。

Table 2: Components of Herbs

herb_id	Ingredient.id	Ingredient.name	Ingredient.alias
HERB006568	HBIN000890	1,2linoleic acid-...	NA
HERB006568	HBIN005609	2-ethyl-3-hydroxy...	NA
HERB006568	HBIN006106	2-Monoolein	2-monoolein; [2-h...
HERB006568	HBIN006366	[(2R)-2,3-dihydro...	(Z)-octadec-9-eno...
HERB006568	HBIN012837	(6Z,10E,14E,18E)-...	NA
HERB006568	HBIN015611	-monolinolein	NA
HERB006568	HBIN015675	-sitosterol	alpha-sitosterol
HERB006568	HBIN016562	Arabinose	arabinose
HERB006568	HBIN016720	arginine	AC1ODX8E; [(4S)-5...
HERB006568	HBIN018278	beta-sitosterol	24.alpha.-Ethylch...
HERB006568	HBIN019257	cadmium	NA
HERB006568	HBIN019351	calcium	NA
HERB006568	HBIN019475	campesterol	campesterol ; FT-...
HERB006568	HBIN019688	caprylic acid	Octanoic acid (mi...
HERB006568	HBIN021150	CLR	5-Cholesten-3b-ol...
...

Figure 1 (下方图) 为图 intersection of all compounds 概览。

(对应文件为 **Figure+Table/intersection-of-all-compounds.pdf**)

Ingred.....1	Herb_p...	Ingred.....3	Ingred.....4	Target.id	Target...	Databa...	Paper.id	...
HBIN00...	GAN CAO	11-deo...	11-Deo...	HBTAR0...	AR	NA	NA	...
HBIN00...	GAN CAO	11-deo...	11-Deo...	HBTAR0...	NR3C1	NA	NA	...
HBIN00...	GAN CAO	11-deo...	11-Deo...	HBTAR0...	AR	NA	NA	...
HBIN00...	FU LING	1,2,3,...	Pentag...	NA	NA	NA	NA	...
HBIN00...	BAI ZHU	12-(-...	12-(-...	NA	NA	NA	NA	...
HBIN00...	BAI ZHU	12-(-...	[(4E,6...	NA	NA	NA	NA	...
HBIN00...	YI YI REN	1,2lin...	NA	NA	NA	NA	NA	...
HBIN00...	GAN CAO	12-met...	12-met...	NA	NA	NA	NA	...
HBIN00...	REN SHEN	12-O-N...	12-o-n...	NA	NA	NA	NA	...
HBIN00...	BAI ZHU	12-Oxo...	(5aR,9...	HBTAR0...	NR3C1	NA	NA	...
...

6.1.3 疾病靶点

取以下基因集的合集：

Figure 2 (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**)

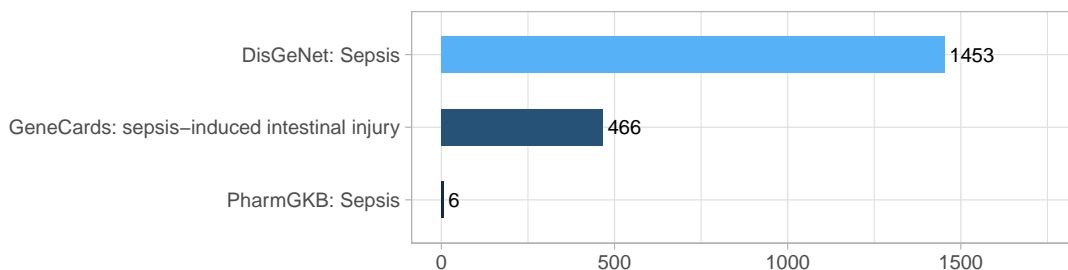


Figure 2: Overall targets number of datasets

6.2 表观遗传学蛋白

自数据库 <https://epifactors.autosome.org/> 获取相关蛋白。

Table 4 (下方表格) 为表格 All protein of epigenetic regulators 概览。

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

注：表格共有 801 行 25 列，以下预览的表格可能省略部分数据；表格含有 801 个唯一 ‘Id’。

Table 4: All protein of epigenetic regulators

Id	HGNC_s...	Status	HGNC_ID	HGNC_name	GeneID	UniPro.....7	UniPro.....8	Domain	MGI_sy...
1	A1CF	#	24086	APOBEC...	29974	Q9NQ94	A1CF_H...	DND1_D...	A1cf
2	ACINU	New	17066	Apopto...	22985	Q9UKV3	ACINU_...	PF1629...	Acin1
3	ACTB	#	132	actin,...	60	P60709	ACTB_H...	Actin ...	Actb
4	ACTL6A	#	24124	actin-...	86	O96019	ACL6A_...	Actin ...	Actl6a
5	ACTL6B	#	160	actin-...	51412	O94805	ACL6B_...	Actin ...	Actl6b
6	ACTR3B	#	17256	ARP3 a...	57180	Q9P1U1	ARP3B_...	Actin ...	Actr3b
7	ACTR5	#	14671	ARP5 a...	79913	Q9H9F9	ARP5_H...	Actin ...	Actr5
8	ACTR6	#	24025	ARP6 a...	64431	Q9GZN1	ARP6_H...	Actin ...	Actr6
9	ACTR8	#	14672	ARP8 a...	93973	Q9H981	ARP8_H...	Actin ...	Actr8
10	ADNP	#	15766	activi...	23394	Q9H2P0	ADNP_H...	Homeob...	Adnp
11	AEBP2	#	24051	AE bin...	121536	Q6ZN18	AEBP2_...	Pfam-B...	Aebp2
12	AICDA	#	13203	activa...	57379	Q9GZX7	AICDA_...	APOBEC...	Aicda
13	AIRE	#	360	autoim...	326	O43918	AIRE_H...	PHD PF...	Aire
14	ALKBH1	#	17911	alkB, ...	8846	Q13686	ALKB1_...	2OG-Fe...	Alkbh1
15	ALKBH1	New	17911	Nuclei...	8846	Q13686	ALKB1_...	PF13532	Alkbh1
...

6.3 筛选表观遗传学靶点

6.3.1 复方成分与表观遗传学靶点

Figure 3 (下方图) 为图 SLBJ network pharmacology of epigenetic target 概览。

(对应文件为 [Figure+Table/SLBJ-network-pharmacology-of-epigenetic-target.pdf](#))

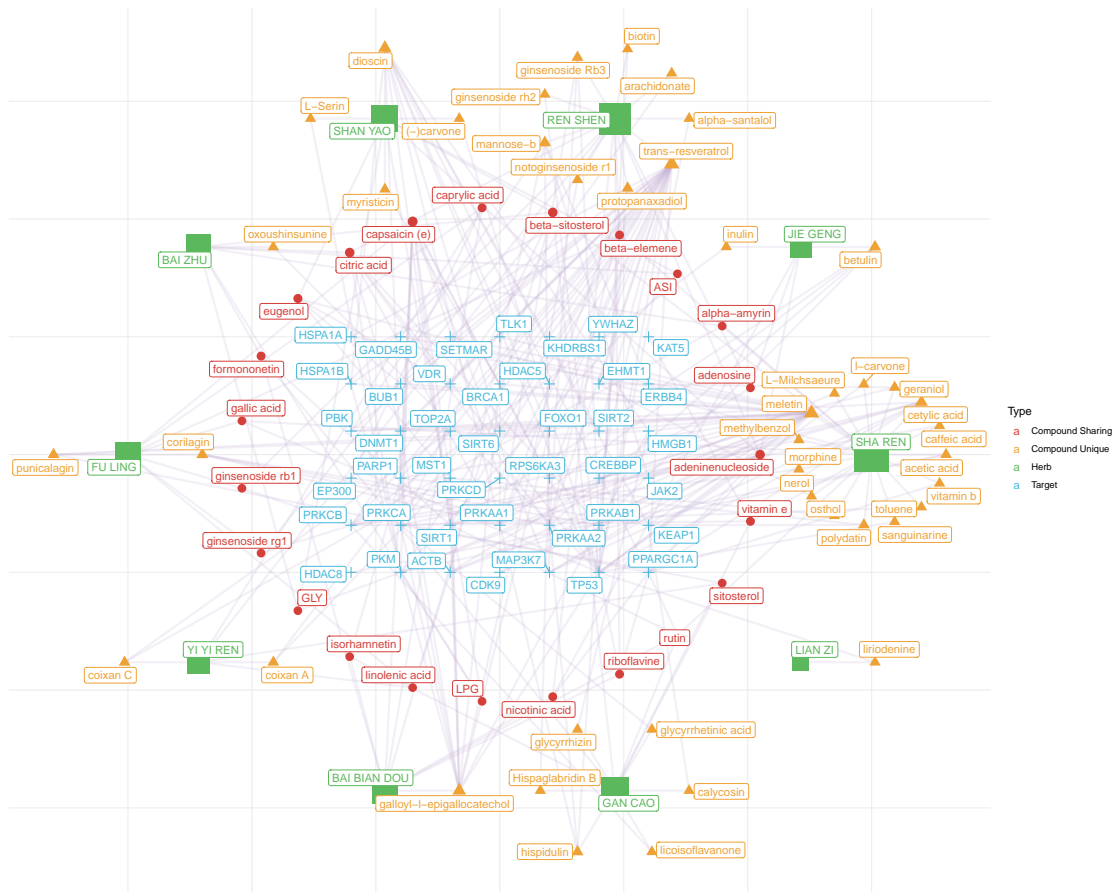


Figure 3: SLBJ network pharmacology of epigenetic target

Table 5 (下方表格) 为表格 SLBJ Herbs compounds and targets of epigenetic target 概览。

(对应文件为 **Figure+Table/SLBJ-Herbs-compounds-and-targets-of-epigenetic-target.csv**)

注：表格共有 583 行 3 列，以下预览的表格可能省略部分数据；表格含有 10 个唯一 ‘Herb_pinyin_name’。

Table 5: SLBJ Herbs compounds and targets of epigenetic target

Herb_pinyin_name	Ingredient.name	Target.name
SHA REN	acetic acid	HDAC8
SHA REN	acetic acid	PKM
SHA REN	acetic acid	HDAC8
SHA REN	acetic acid	PKM
JIE GENG	adeninenucleoside	ACTB
JIE GENG	adeninenucleoside	CDK9
JIE GENG	adeninenucleoside	MAP3K7
JIE GENG	adeninenucleoside	TP53
JIE GENG	adeninenucleoside	PPARGC1A

Herb_pinyin_name	Ingredient.name	Target.name
JIE GENG	adeninenucleoside	ACTB
JIE GENG	adeninenucleoside	CDK9
JIE GENG	adeninenucleoside	MAP3K7
JIE GENG	adeninenucleoside	TP53
JIE GENG	adeninenucleoside	PPARGC1A
JIE GENG	adeninenucleoside	ACTB
...

6.4 脓毒症肠损伤的 GEO 数据分析

6.4.1 数据来源

注：由于该数据的原作者没有上传定量后的原始 Count，不利于差异分析；因此，这里下载了 SRA 原始数据，使用 Kallisto 重新定量。

Data Source ID :

GSE202261

data__processing :

Raw data (raw reads) of fastq format were firstly processed using fastp(version 0.20.0) software with default parameters in pair end mode.

data__processing.1 :

The clean data were obtained for downstream analyses by removing reads containing adapter, reads containing ploy-N and low-quality reads from raw data.

data__processing.2 :

The clean reads were mapped to reference of mouse mRNAs and lncRNAs using bowtie2(version 2.3.1) with parameters -q -L 16 -phred64 -p 6

data__processing.3 :

TPM value and read counts of each transcript were obtained by eXpress(version 1.5.1) with parameters -no-update-check -rf-stranded.

(Others) :

...

(上述信息框内容已保存至 `Figure+Table/SII-GSE202261-content`)

Table 6 (下方表格) 为表格 SII metadata 概览。

(对应文件为 `Figure+Table/SII-metadata.csv`)

注：表格共有 4 行 8 列，以下预览的表格可能省略部分数据；表格含有 4 个唯一 ‘sample’。

1. sample: 样品名称
2. group: 分组名称

Table 6: SII metadata

sample	group	lib.size	norm.f...	file	directory	sample1	gsm
SRR211...	control	206657...	1	kallis...	kallis...	SRR211...	GSM646...
SRR211...	control	194806...	1	kallis...	kallis...	SRR211...	GSM646...
SRR211...	model	205782...	1	kallis...	kallis...	SRR211...	GSM646...
SRR211...	model	230056...	1	kallis...	kallis...	SRR211...	GSM646...

6.4.2 fastp 质控

‘Fastp QC’ 数据已全部提供。

(对应文件为 `./fastp_report/`)

注：文件夹 `./fastp_report/` 共包含 4 个文件。

1. SRR21101636.html
2. SRR21101637.html
3. SRR21101638.html
4. SRR21101639.html

6.4.3 RNA 定量

使用小鼠 cDNA 作为参考基因组 (`Mus_musculus.GRCm39.cdna.all.fa.gz`), `Kallisto` 定量。

Table 7 (下方表格) 为表格 Quantification 概览。

(对应文件为 `Figure+Table/Quantification.csv`)

注：表格共有 116873 行 5 列，以下预览的表格可能省略部分数据；表格含有 116873 个唯一 ‘target_id’。

Table 7: Quantification

target_id	SRR211016361	SRR211016371	SRR211016381	SRR211016391
ENSMUST00000196221	0	0	0	0
ENSMUST00000179664	0	0	0	0
ENSMUST00000177564	0	0	0	0
ENSMUST00000178537	0	0	0	0
ENSMUST00000178862	0	0	0	0
ENSMUST00000179520	0	0	0	0
ENSMUST00000179883	0	0	0	0
ENSMUST00000195858	0	0	0	0
ENSMUST00000179932	0	0	0	0
ENSMUST00000180001	0	0	0	0
ENSMUST00000178815	0	0	0	0
ENSMUST00000177965	0	0	0	0
ENSMUST00000178909	0	0	0	0
ENSMUST00000177646	0	0	0	0
ENSMUST00000178230	0	0	0	0
...

6.4.4 差异分析

使用 limma 差异分析

Figure 4 (下方图) 为图 SII model vs control DEGs 概览。

(对应文件为 **Figure+Table/SII-model-vs-control-DEGs.pdf**)

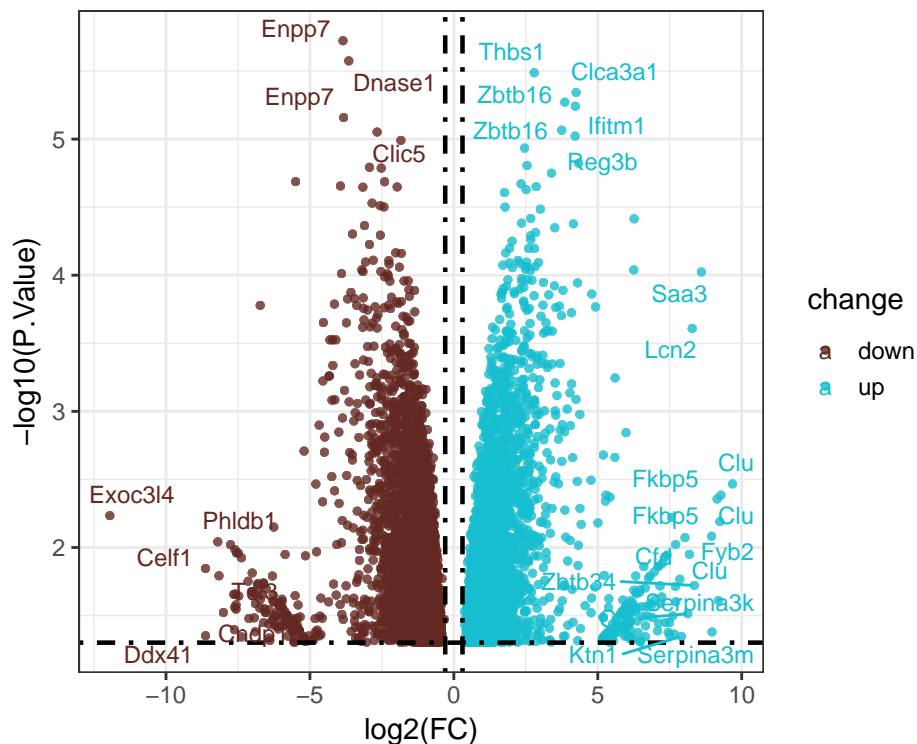


Figure 4: SII model vs control DEGs

6.4.5 将小鼠基因 symbol 映射到人类 (hgnc_symbol)

使用 biomaRt 将基因映射

Table 8 (下方表格) 为表格 Mapped DEGs 概览。

(对应文件为 Figure+Table/Mapped-DEGs.xlsx)

注：表格共有 4409 行 13 列，以下预览的表格可能省略部分数据；表格含有 4338 个唯一 'hgnc_symbol'。

1. hgnc_symbol: 基因名 (Human)
2. mgi_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 'topTableF')

Table 8: Mapped DEGs

hgnc_s...	mgc_sy...	logFC	P.Value	rownames	ensembl.....6	ensembl.....7	entrez...	descri...	...
ENPP7	Enpp7	-3.851...	1.8936...	43380	ENSMUS...	ENSMUS...	238011	ectonu...	...
THBS1	Thbs1	2.7918...	3.2577...	94883	ENSMUS...	ENSMUS...	21825	thromb...	...
DNASE1	Dnase1	-3.650...	2.6636...	26241	ENSMUS...	ENSMUS...	13419	deoxyr...	...
ANPEP	Anpep	-1.838...	1.0218...	50217	ENSMUS...	ENSMUS...	16790	alanyl...	...
CLIC5	Clic5	-2.664...	8.8784...	45278	ENSMUS...	ENSMUS...	224796	chlori...	...
ZBTB16	Zbtb16	3.8536...	5.3567...	5449	ENSMUS...	ENSMUS...	235320	zinc f...	...
CEP85	Cep85	2.4595...	1.1643...	45584	ENSMUS...	ENSMUS...	70012	centro...	...
REG1B	Reg3b	4.2139...	9.5036...	25543	ENSMUS...	ENSMUS...	18489	regene...	...
IGFBP5	Igfbp5	2.5428...	1.5655...	72899	ENSMUS...	ENSMUS...	16011	insuli...	...
PMAIP1	Pmaip1	3.3882...	1.7795...	14956	ENSMUS...	ENSMUS...	58801	phorbo...	...
CYP4F2	Cyp4f14	-1.967...	2.2471...	66083	ENSMUS...	ENSMUS...	64385	cytoch...	...
PRR15	Prr15	-2.404...	2.0582...	38200	ENSMUS...	ENSMUS...	78004	prolin...	...
FKBP5	Fkbp5	2.8596...	2.2370...	54319	ENSMUS...	ENSMUS...	14229	FK506
ERRFI1	Errfi1	2.3357...	2.1310...	90849	ENSMUS...	ENSMUS...	74155	ERBB r...	...
LGMN	Lgmn	2.5094...	2.3522...	59306	ENSMUS...	ENSMUS...	19141	leguma...	...
...

6.4.6 富集分析 (GSEA)

Figure 5 (下方图) 为图 KEGG enrichment 概览。

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

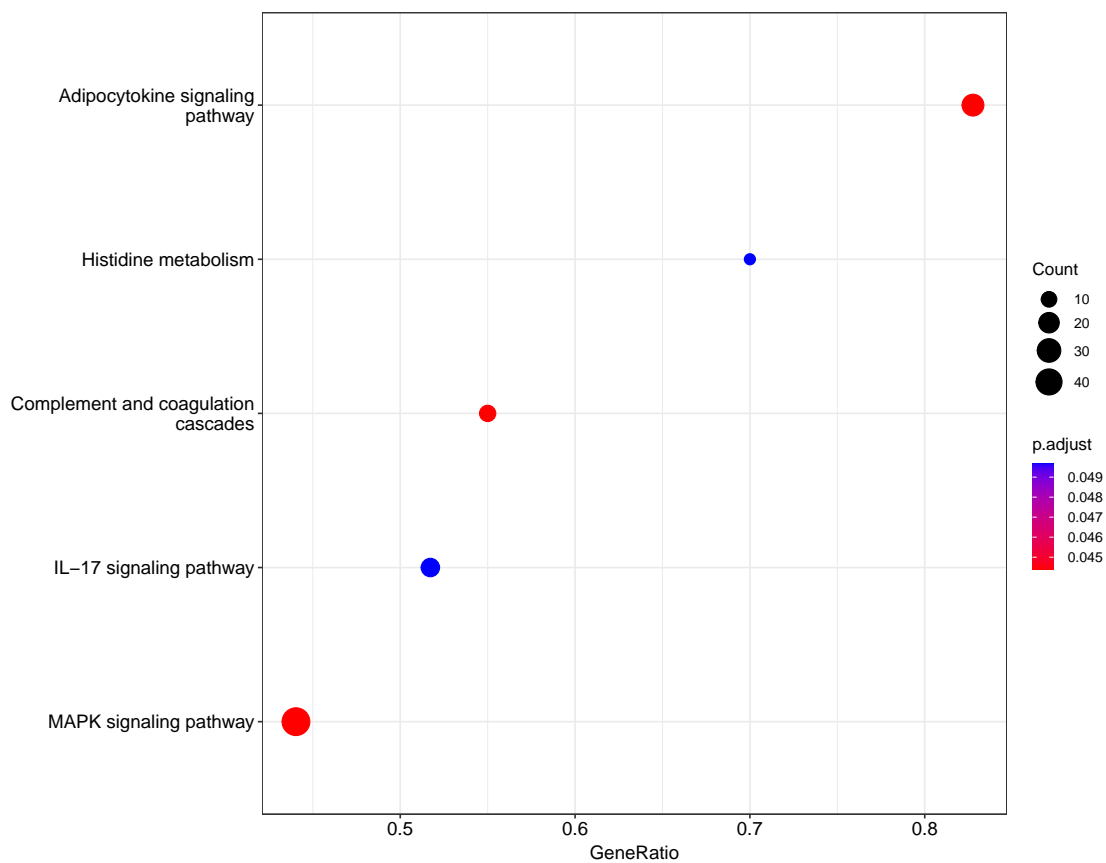


Figure 5: KEGG enrichment

Table 9 (下方表格) 为表格 KEGG enrichment data 概览。

(对应文件为 **Figure+Table/KEGG-enrichment-data.xlsx**)

注：表格共有 5 行 13 列，以下预览的表格可能省略部分数据；表格含有 5 个唯一 ‘ID’。

1. pvalue: 显著性 P。

Table 9: KEGG enrichment data

ID	Descri...	setSize	enrich...	NES	pvalue	p.adjust	qvalue	rank	leadin...
hsa04610	Comple...	20	0.6129...	2.1798...	0.0002...	0.0444...	0.0415...	789	tags=5...
hsa04920	Adipoc...	29	0.5336...	2.1079...	0.0004...	0.0444...	0.0415...	1541	tags=8...
hsa04010	MAPK s...	109	0.3242...	1.7682...	0.0004...	0.0444...	0.0415...	1026	tags=4...
hsa00340	Histid...	10	-0.726...	-2.053...	0.0008...	0.0496...	0.0464...	536	tags=7...
hsa04657	IL-17 ...	29	0.5129...	2.0264...	0.0008...	0.0496...	0.0464...	704	tags=5...

6.5 复方成分表观遗传学靶点的通路调控

6.5.1 富集表观修饰蛋白的通路

以 Fig. 3 的靶点筛选, 发现存在三条通路:

Figure 6 (下方图) 为图 KEGG enrichment with enriched genes 概览。

(对应文件为 Figure+Table/KEGG-enrichment-with-enriched-genes.pdf)

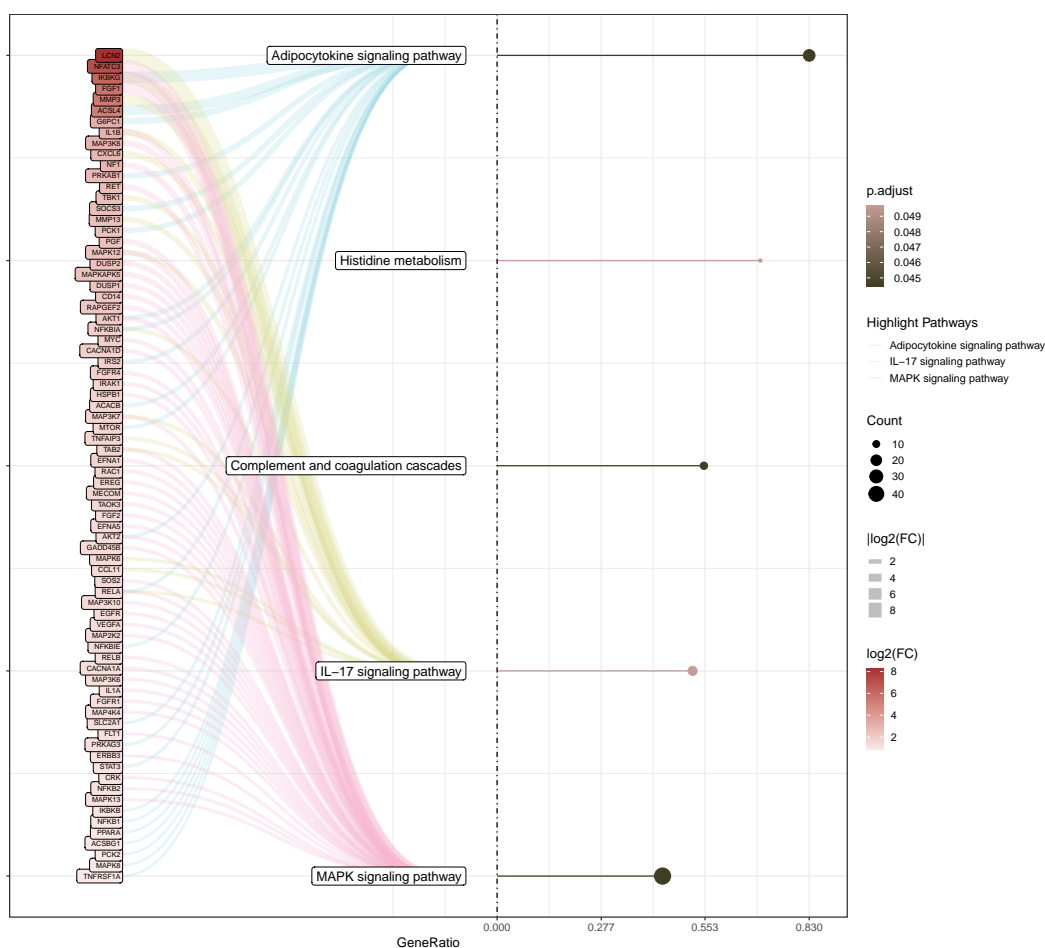


Figure 6: KEGG enrichment with enriched genes

Figure 7 (下方图) 为图 GSEA plot of the pathways 概览。

(对应文件为 Figure+Table/GSEA-plot-of-the-pathways.pdf)

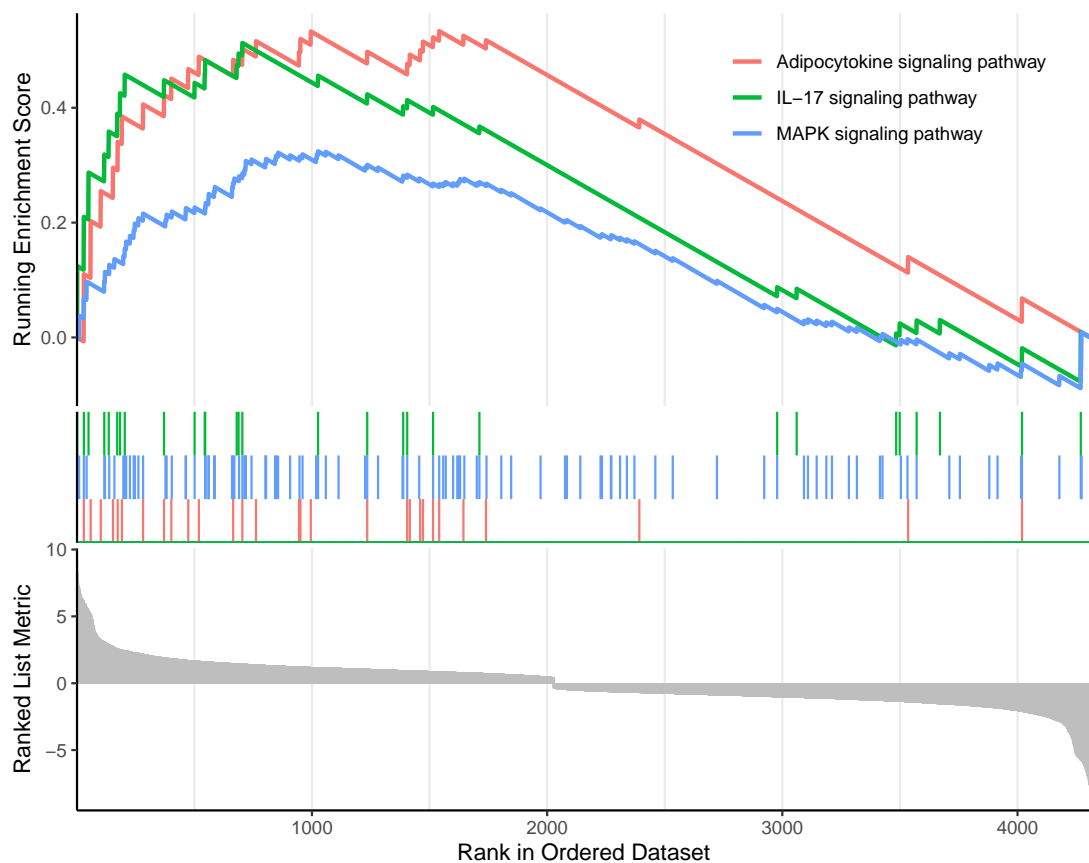


Figure 7: GSEA plot of the pathways

6.5.2 表观修饰靶点

这三条通路存在的表观遗传学修饰靶点为：

Adipocytokine signaling pathway :

PRKAB1

MAPK signaling pathway :

MAP3K7, GADD45B

IL-17 signaling pathway :

MAP3K7

(上述信息框内容已保存至 Figure+Table/unnamed-chunk-32-content)

6.5.3 上下游

Figure 8 (下方图) 为图 Hsa04920 visualization 概览。

(对应文件为 Figure+Table/hsa04920.pathview.png)

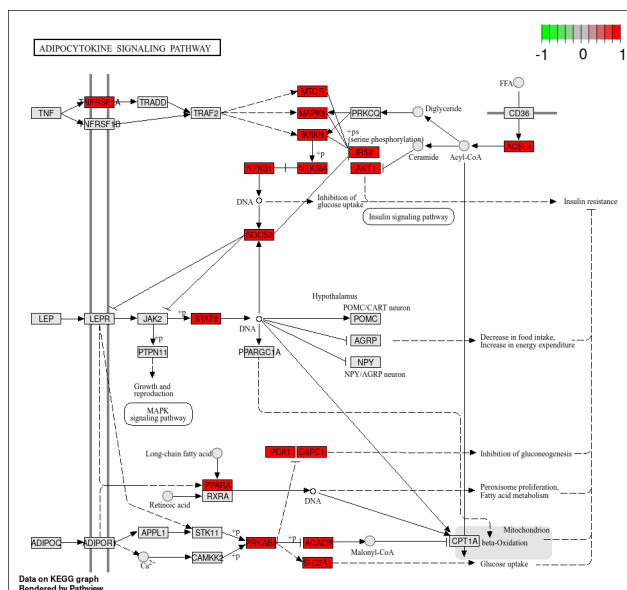


Figure 8: Hsa04920 visualization

Interactive figure :

<https://www.genome.jp/pathway/hsa04920>

Figure 9 (下方图) 为图 Hsa04010 visualization 概览。

(对应文件为 Figure+Table/hsa04010.pathview.png)

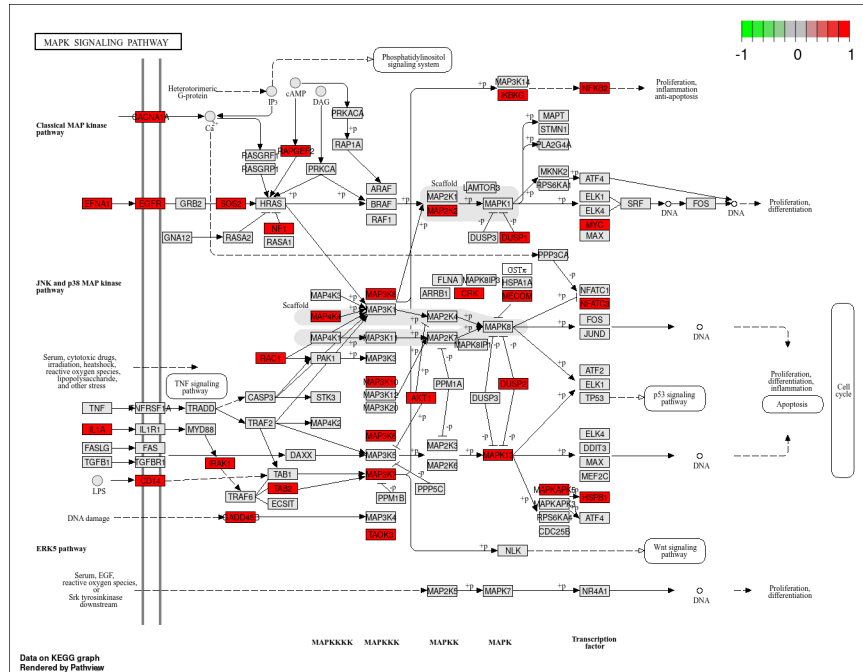


Figure 9: Hsa04010 visualization

Interactive figure :

<https://www.genome.jp/pathway/hsa04010>

Figure 10 (下方图) 为图 Hsa04657 visualization 概览。

(对应文件为 Figure+Table/hsa04657.pathview.png)

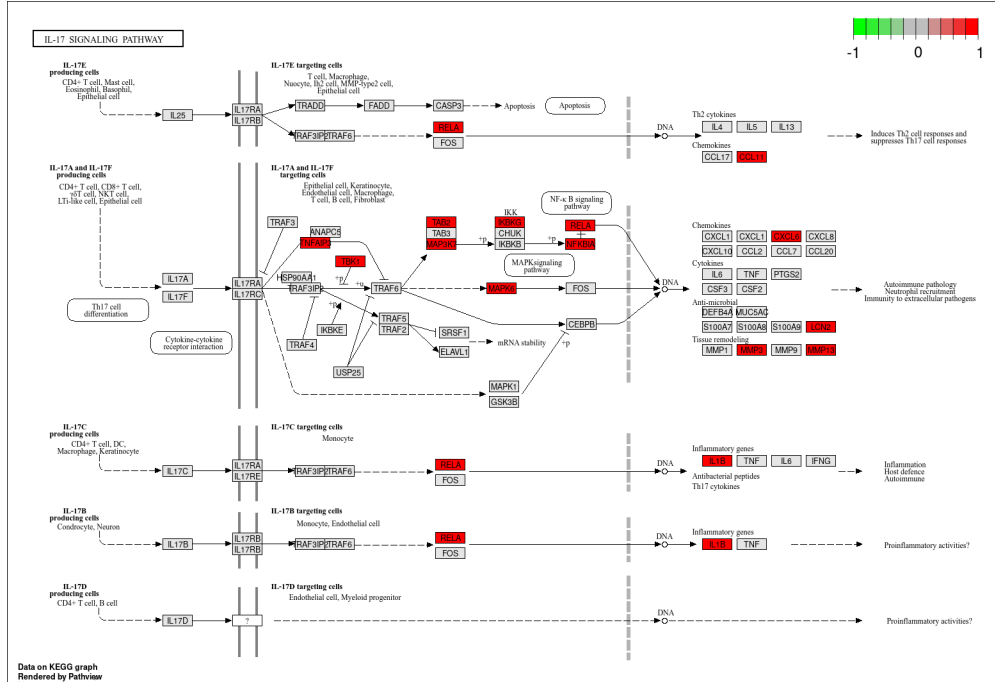


Figure 10: Hsa04657 visualization

Interactive figure :

<https://www.genome.jp/pathway/hsa04657>

6.5.4 相关成分

Figure 11 (下方图) 为图 SLBJ network pharmacology Target epigenetic related pathway 概览。

(对应文件为 Figure+Table/SLBJ-network-pharmacology-Target-epigenetic-related-pathway.pdf)

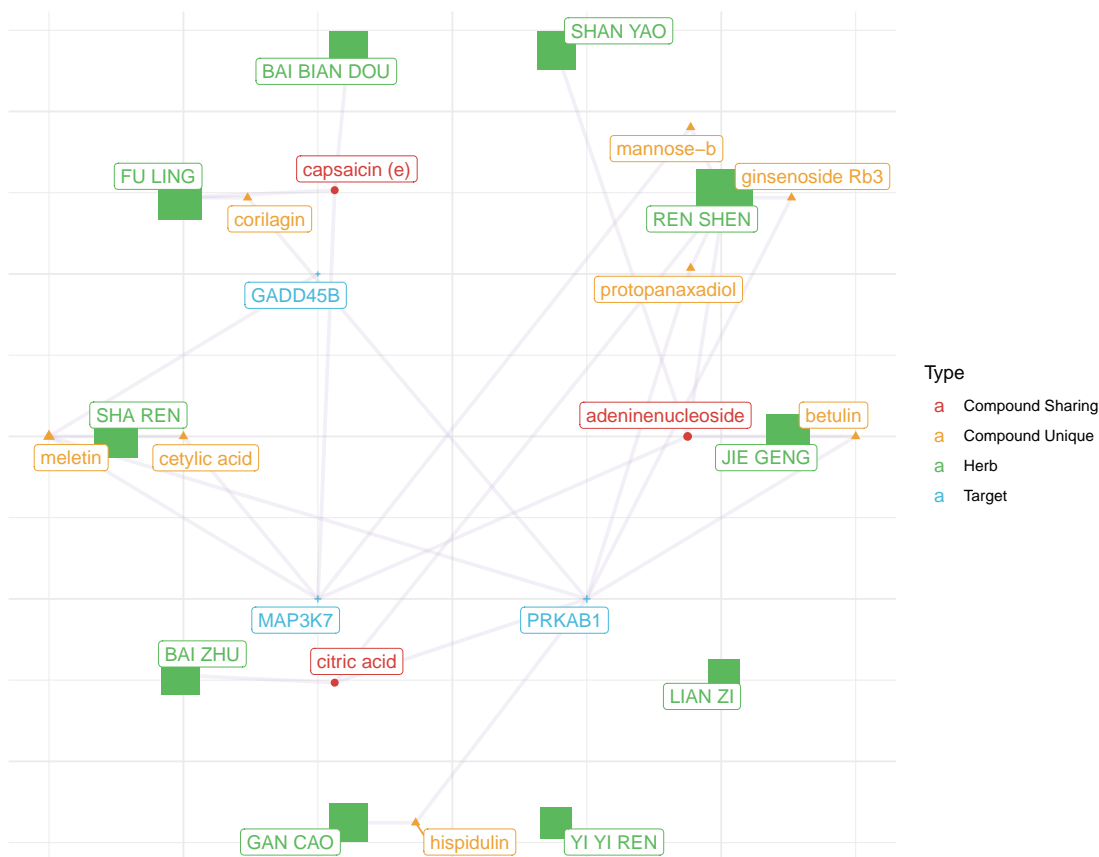


Figure 11: SLBJ network pharmacology Target epigenetic related pathway

Table 10 (下方表格) 为表格 SLBJ network pharmacology Target epigenetic related pathway data 概览。

(对应文件为 **Figure+Table/SLBJ-network-pharmacology-Target-epigenetic-related-pathway-data.csv**)

注：表格共有 42 行 3 列，以下预览的表格可能省略部分数据；表格含有 8 个唯一 ‘Herb_pinyin_name’。

Table 10: SLBJ network pharmacology Target epigenetic related pathway data

Herb_pinyin_name	Ingredient.name	Target.name
JIE GENG	adeninenucleoside	MAP3K7
JIE GENG	adeninenucleoside	MAP3K7
JIE GENG	adeninenucleoside	MAP3K7
JIE GENG	adeninenucleoside	MAP3K7
REN SHEN	adeninenucleoside	MAP3K7
REN SHEN	adeninenucleoside	MAP3K7
REN SHEN	adeninenucleoside	MAP3K7
REN SHEN	adeninenucleoside	MAP3K7
SHAN YAO	adeninenucleoside	MAP3K7

Herb_pinyin_name	Ingredient.name	Target.name
SHAN YAO	adeninenucleoside	MAP3K7
SHAN YAO	adeninenucleoside	MAP3K7
SHAN YAO	adeninenucleoside	MAP3K7
JIE GENG	betulin	PRKAB1
JIE GENG	betulin	PRKAB1
JIE GENG	betulin	PRKAB1
...

6.6 分子对接

6.6.1 第一批对接

根据 HERB 数据库记录的成分靶点信息，Fig. 11 中的成分能作用于其对应的靶点。以下尝试分子对接。

Figure 12 (下方图) 为图 FIRST Overall combining Affinity 概览。

(对应文件为 `Figure+Table/FIRST-Overall-combining-Affinity.pdf`)

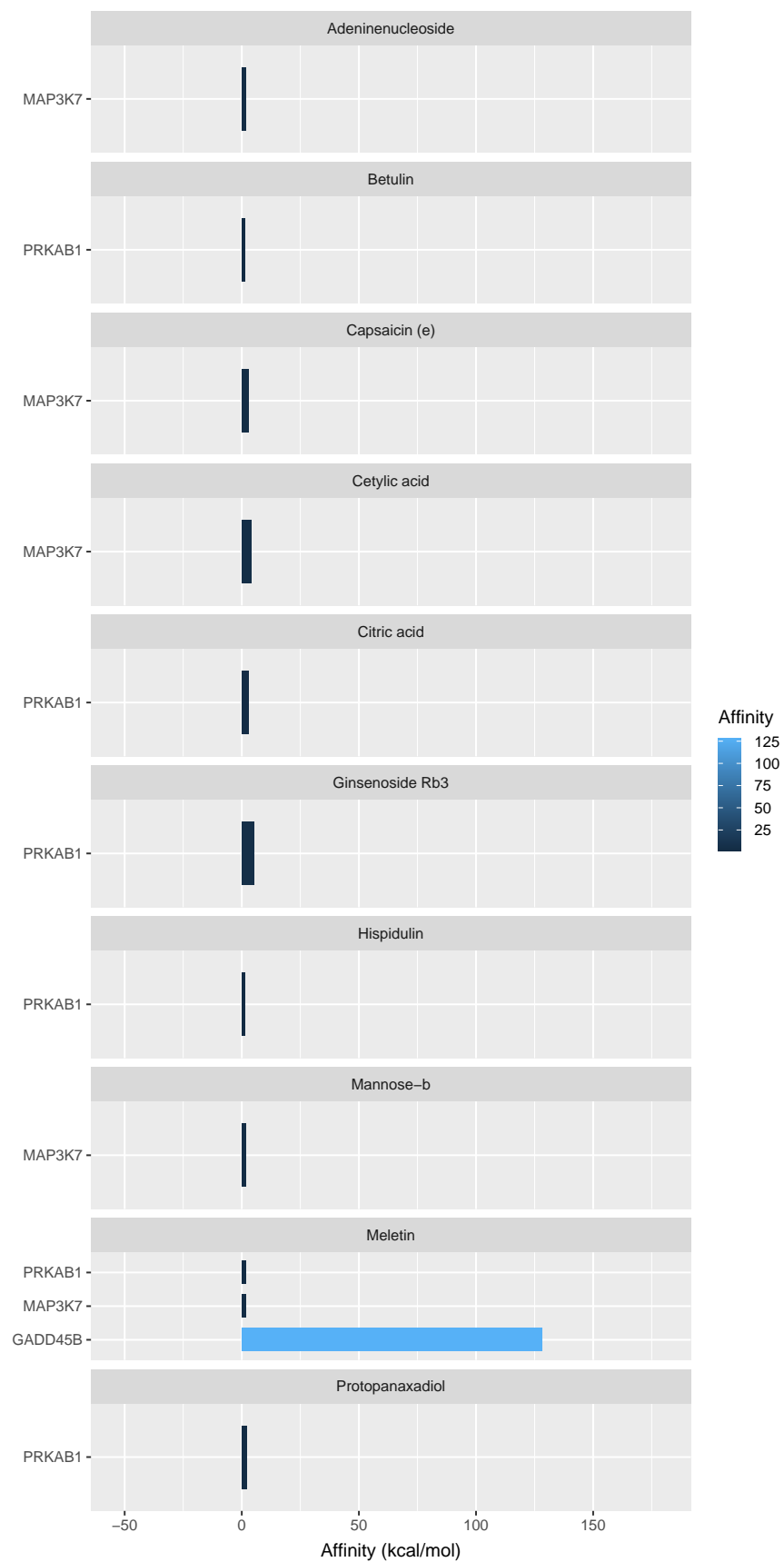


Figure 12: FIRST Overall combining Affinity

Fig. 12 显示，候选分子与对应靶点的对接能量均较大。

6.6.2 以口服利用度筛选其他成分

由于 Fig. 12 所示对接能量过大，以下尝试挖掘复方中其它能够作用于 Fig. 11 表观遗传修饰靶点的成分。

以下通过 HOB 筛选成分 (预测是否达到 20% HOB)。

Figure 13 (下方图) 为图 HOB 20 prediction 概览。

(对应文件为 Figure+Table/HOB-20-prediction.pdf)

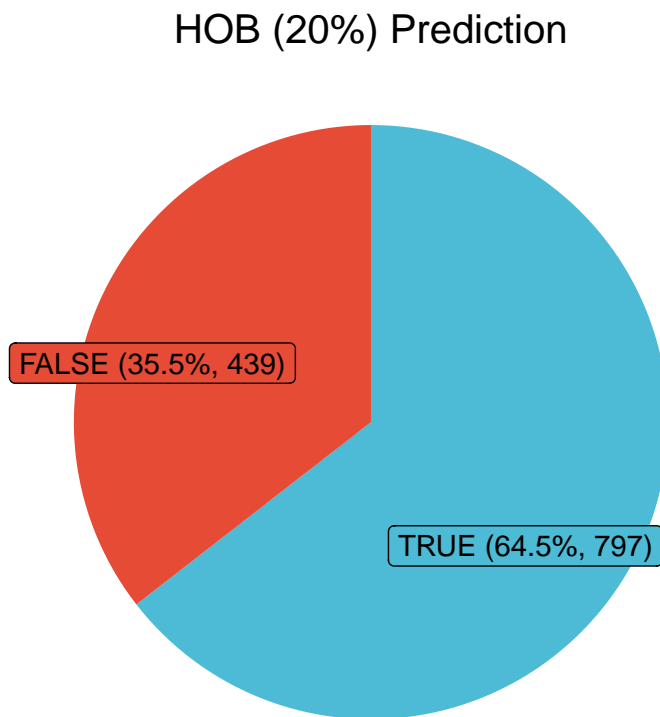


Figure 13: HOB 20 prediction

6.6.3 第二批对接

对于 Fig. 13 满足 HOB 条件的化合物，尝试分子对接；然而由于化合物数量过多，运算过于将过于耗时，这里，以 ChemmineR 对化合物结构聚类 (0.4 cut-off)，每个聚类团随机抽取三个化合物，最后用于分子对接。

实际对接的有：

543854 :

7NTI, 7MYJ, gadd45b

480873 :

7NTI, 7MYJ, gadd45b

6992244 :

7NTI, 7MYJ, gadd45b

6553876 :

7NTI, 7MYJ, gadd45b

20055133 :

7NTI, 7MYJ, gadd45b

(Others) :

...

(上述信息框内容已保存至 **Figure+Table/unnamed-chunk-43-content**)

对每个靶点都选择了对接能量最小的 Top 5，结果如下：

Figure 14 (下方图) 为图 Overall combining Affinity 概览。

(对应文件为 **Figure+Table/Overall-combining-Affinity.pdf**)

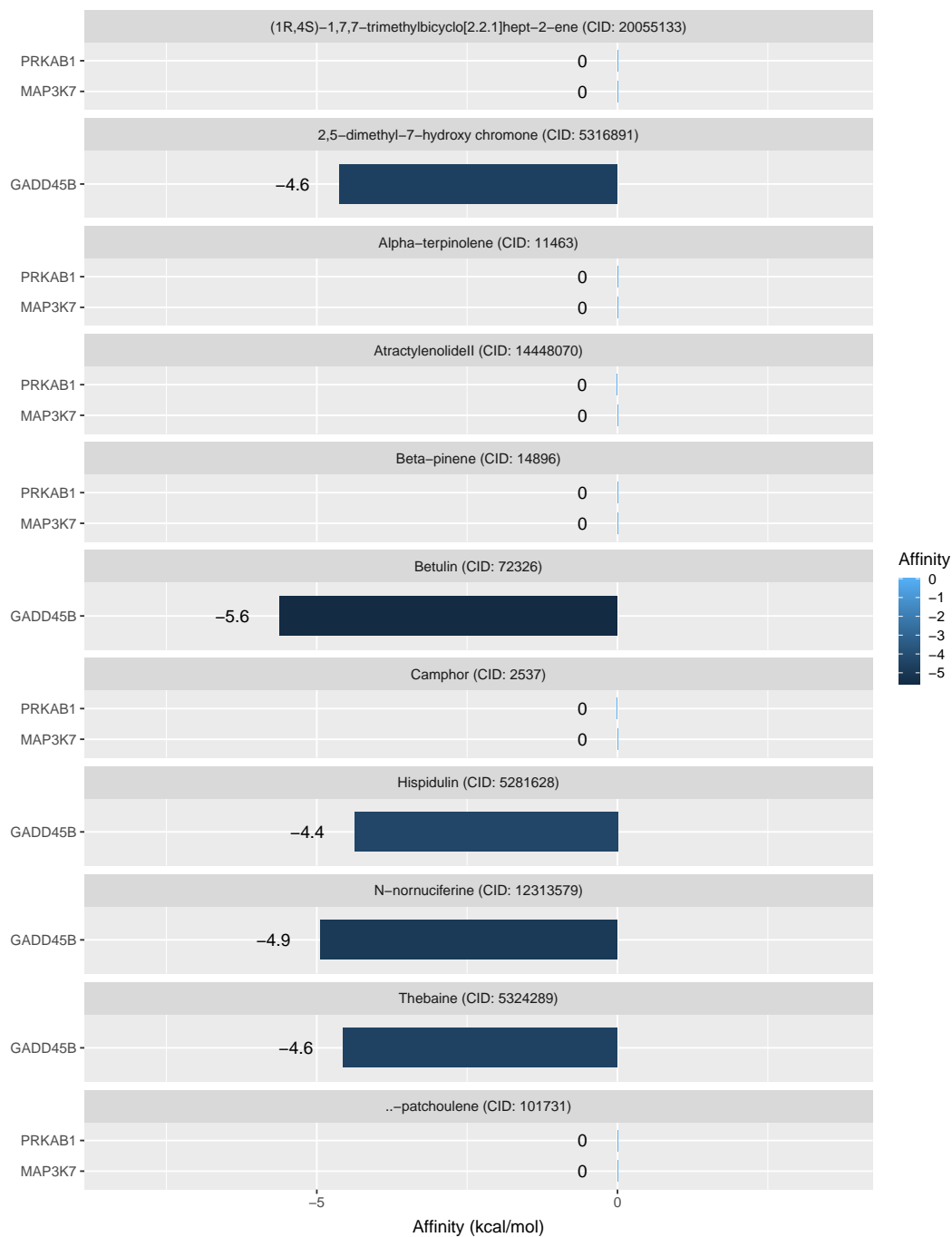


Figure 14: Overall combining Affinity

6.6.4 对接可视化 (Top 3)

Figure 15 (下方图) 为图 Docking 72326 into GADD45B 概览。

(对应文件为 Figure+Table/72326_into_GADD45B.png)

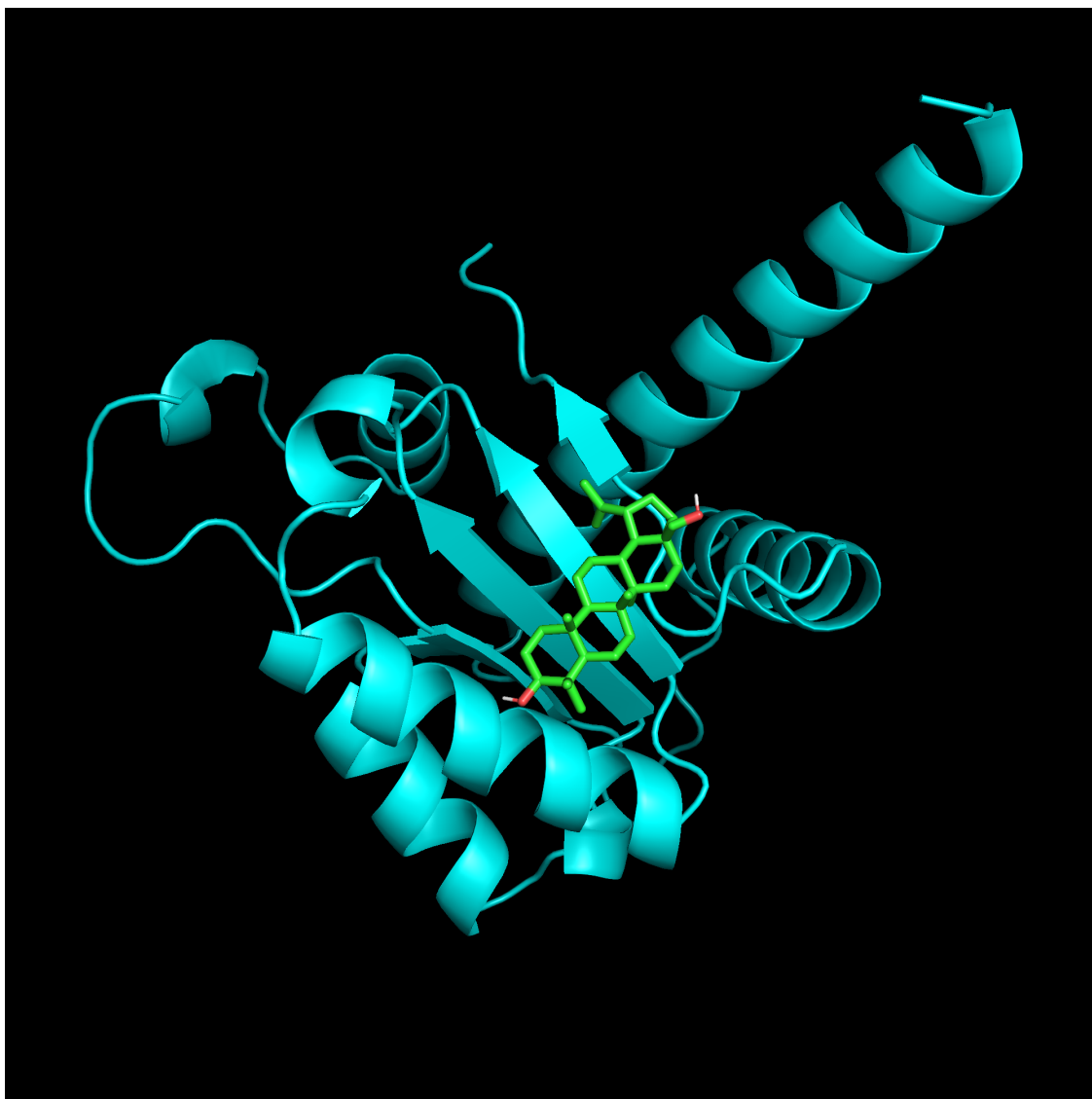


Figure 15: Docking 72326 into GADD45B

Figure 16 (下方图) 为图 Docking 12313579 into GADD45B 概览。

(对应文件为 Figure+Table/12313579_into_GADD45B.png)



Figure 16: Docking 12313579 into GADD45B

Figure 17 (下方图) 为图 Docking 5316891 into GADD45B 概览。

(对应文件为 Figure+Table/5316891_into_GADD45B.png)

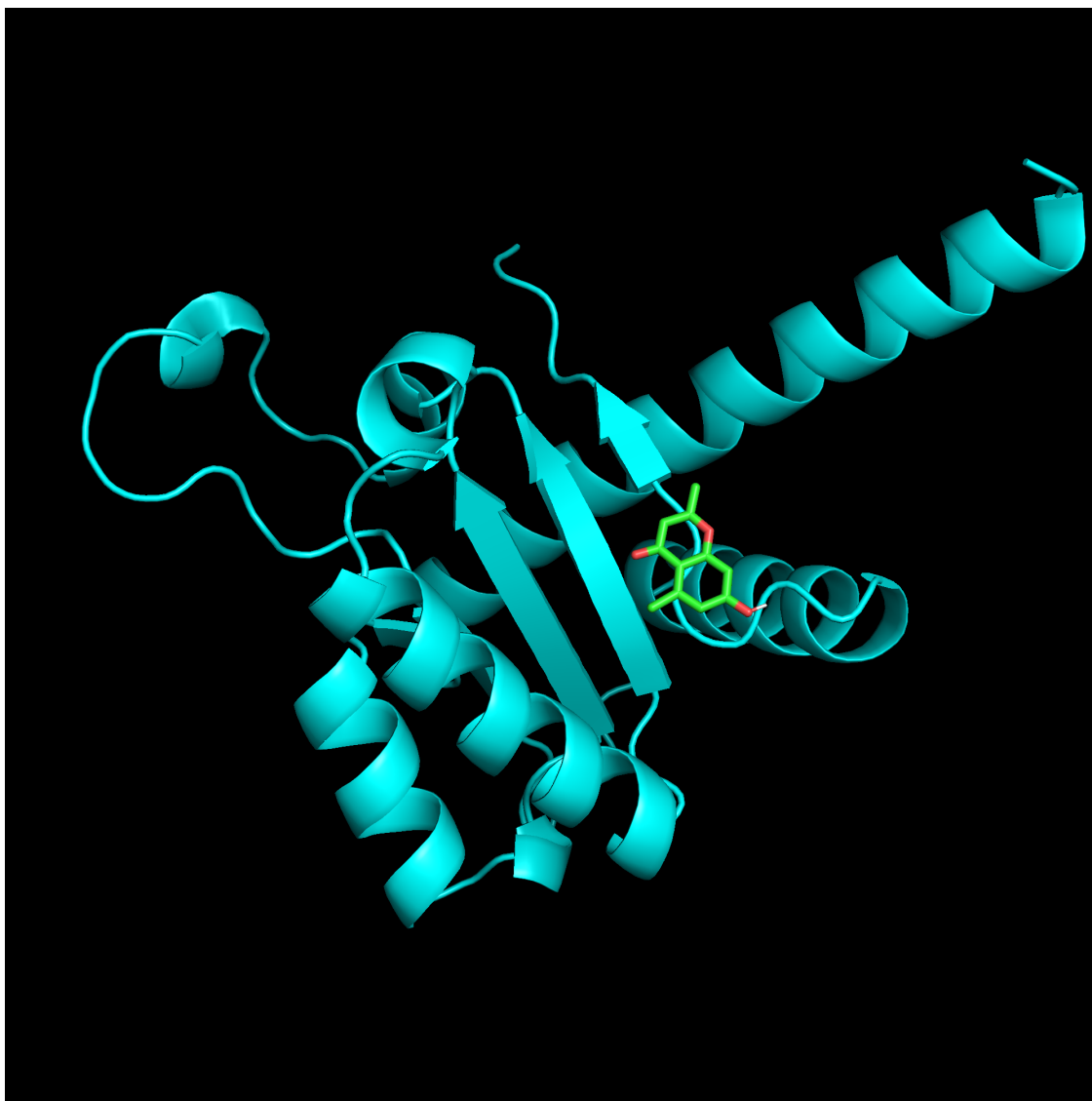


Figure 17: Docking 5316891 into GADD45B

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. Chen, S. Ultrafast one-pass fastq data preprocessing, quality control, and deduplication using fastp. *iMeta* **2**, (2023).

5. Piñero, J. *et al.* The disgenet knowledge platform for disease genomics: 2019 update. *Nucleic Acids Research* (2019) doi:10.1093/nar/gkz1021.
6. 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).
7. 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).
8. Fang, S. *et al.* HERB: A high-throughput experiment- and reference-guided database of traditional chinese medicine. *Nucleic Acids Research* **49**, D1197–D1206 (2021).
9. Wei, M. *et al.* HobPre: Accurate prediction of human oral bioavailability for small molecules. *Journal of Cheminformatics* **14**, (2022).
10. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic rna-seq quantification. *Nature Biotechnology* **34**, (2016).
11. Ritchie, M. E. *et al.* Limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47 (2015).
12. Chen, Y., McCarthy, D., Ritchie, M., Robinson, M. & Smyth, G. EdgeR: Differential analysis of sequence read count data user’s guide. 119.
13. Eberhardt, J., Santos-Martins, D., Tillack, A. F. & Forli, S. AutoDock vina 1.2.0: New docking methods, expanded force field, and python bindings. *Journal of Chemical Information and Modeling* **61**, 3891–3898 (2021).
14. Zhang, Y., Forli, S., Omelchenko, A. & Sanner, M. F. AutoGridFR: Improvements on autodock affinity maps and associated software tools. *Journal of computational chemistry* **40**, 2882–2886 (2019).
15. Zhang, Y. & Sanner, M. F. AutoDock crankpep: Combining folding and docking to predict protein-peptide complexes. *Bioinformatics (Oxford, England)* **35**, 5121–5127 (2019).
16. Ravindranath, P. A. & Sanner, M. F. AutoSite: An automated approach for pseudo-ligands prediction-from ligand-binding sites identification to predicting key ligand atoms. *Bioinformatics (Oxford, England)* **32**, 3142–3149 (2016).
17. Ravindranath, P. A., Forli, S., Goodsell, D. S., Olson, A. J. & Sanner, M. F. AutoDockFR: Advances in protein-ligand docking with explicitly specified binding site flexibility. *PLoS computational biology* **11**, (2015).
18. Cao, Y., Charisi, A., Cheng, L.-C., Jiang, T. & Girke, T. ChemmineR: A compound mining framework for r. *Bioinformatics (Oxford, England)* **24**, 1733–1734 (2008).