

中药-有效成分-乳腺癌相关靶点的网药分析

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	1
4	分析结果	2
5	结论	2
6	附：分析流程	2
6.1	网络药理学	2
6.1.1	成分	2
6.1.2	成分靶点	4
6.1.3	疾病靶点	4
6.1.4	疾病-成分-靶点网络图	5
6.1.5	富集分析	7
6.1.6	疾病-成分-靶点-通路网络图	9
	Reference	11

List of Figures

1	Intersection of all compounds	3
2	Network pharmacology with disease	6
3	Targets intersect with targets of diseases	6
4	KEGG enrichment	8
5	GO enrichment	9
6	Network pharmacology with disease and pathway	10

List of Tables

1	Herbs information	2
2	Compounds filtered by OB and DL	2
3	Tables of Herbs compounds and targets	4
4	Disease related targets from GeneCards	5
5	Network pharmacology with disease and pathway data	10

1 摘要

1.1 需求

网络药理学分析

- 药对：白花蛇舌草，半枝莲，浙贝母
- 疾病：乳腺癌
- 目标：提供中药-有效成分-乳腺癌相关靶点的网药分析

1.2 结果

- 数据来源于 TCMSP，以 OB、DL 筛选过化合物 Tab. 2。
- 疾病靶点来源于 GeneCards, Tab. 4
- 疾病成分靶点网络图：Fig. 3
- 包含通路：Fig. 6, Tab. 5

2 前言

3 材料和方法

3.1 材料

3.2 方法

Mainly used method:

- R package `ClusterProfiler` used for gene enrichment analysis¹.
- The Human Gene Database `GeneCards` used for disease related genes prediction².
- Website TCMSP <https://tcmsp-e.com/> used for data source³.
- The API of `UniProtKB` (https://www.uniprot.org/help/api_queries) used for mapping of names or IDs of proteins.
- 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 网络药理学

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.csv)

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

Table 1: Herbs information

Herb_pinyin_name	Herb_cn_name
Baihuasheshecao	白花蛇舌草
Banzhilian	半枝莲
Zhebeimu	浙贝母

Table 2 (下方表格) 为表格 Compounds filtered by OB and DL 概览。

(对应文件为 Figure+Table/Compounds-filtered-by-OB-and-DL.xlsx)

注：表格共有 43 行 15 列，以下预览的表格可能省略部分数据；含有 39 个唯一 ‘Mol ID’；含有 3 个唯一 ‘Herb_pinyin_name’。

OB (%) and DL cut-off :

OB \geq 30%; DL \geq 0.18

Table 2: Compounds filtered by OB and DL

Mol ID	Molecu...	MW	AlogP	Hdon	Hacc	OB (%)	Caco-2	BBB	DL
MOL001646	2,3-di...	282.310	3.262	0	4	34.858...	0.75128	0.17357	0.26255
MOL001659	Porife...	412.770	7.640	1	1	43.829...	1.43659	1.03472	0.75596
MOL001663	(4aS,6...	456.780	6.422	2	3	32.028...	0.60932	0.39268	0.75713
MOL001670	2-meth...	252.280	3.278	0	3	37.827...	0.72896	-0.12795	0.20517

Mol ID	Molecu...	MW	AlogP	Hdon	Hacc	OB (%)	Caco-2	BBB	DL
MOL000449	Stigma...	412.770	7.640	1	1	43.829...	1.44458	1.00045	0.75665
MOL000358	beta-s...	414.790	8.084	1	1	36.913...	1.32463	0.98588	0.75123
MOL000098	quercetin	302.250	1.504	5	7	46.433...	0.04842	-0.76890	0.27525
MOL001040	(2R)-5...	272.270	2.298	3	5	42.363...	0.37818	-0.47578	0.21141
MOL012245	5,7,4'...	302.300	2.281	3	6	36.626...	0.43274	-0.31890	0.26833
MOL012246	5,7,4'...	302.300	2.281	3	6	74.235...	0.37328	-0.43273	0.26479
MOL012248	5-hydr...	328.340	2.820	1	6	65.818...	0.84750	0.07437	0.32874
MOL012250	7-hydr...	298.310	2.836	1	5	43.716...	0.95759	0.22129	0.25376
MOL012251	Chrysi...	268.280	2.853	1	4	37.268...	0.90922	0.15556	0.20317
MOL012252	9,19-c...	426.800	7.554	1	1	38.685...	1.44891	1.16360	0.78074
MOL002776	Baicalin	446.390	0.639	6	11	40.123...	-0.84777	-1.74426	0.75264
...

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)

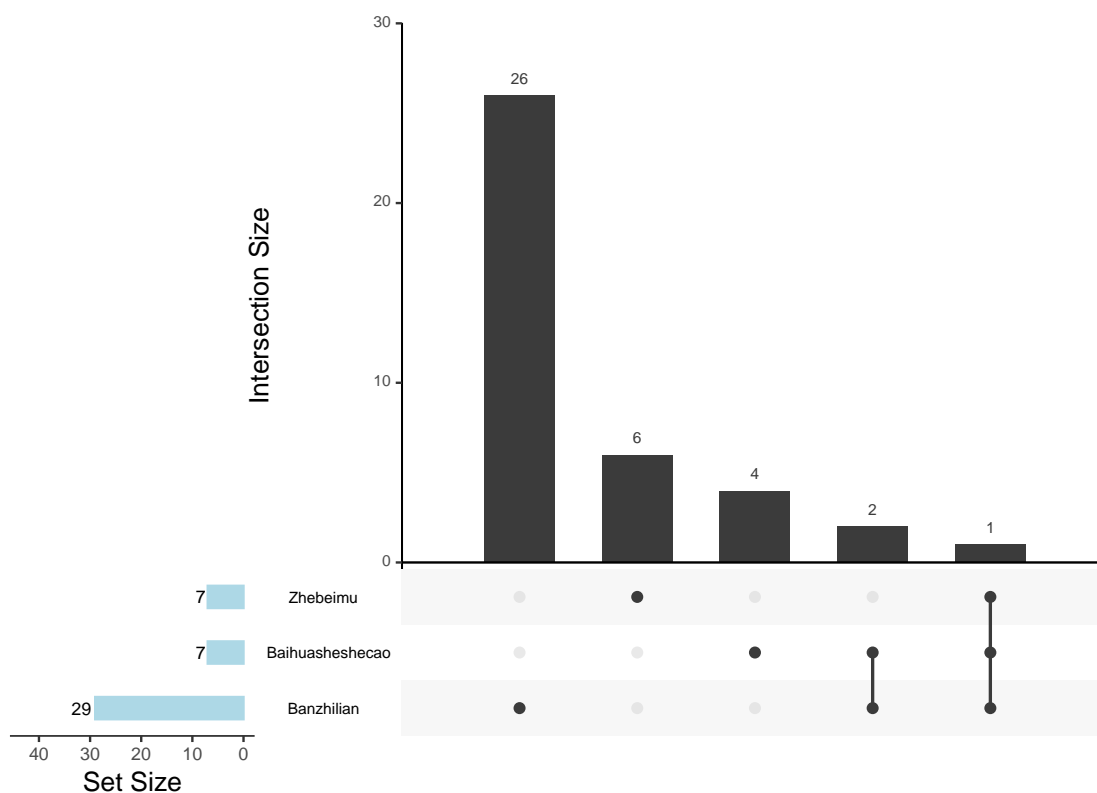


Figure 1: Intersection of all compounds

All_intersection :

MOL000358

(上述信息框内容已保存至 Figure+Table/intersection-of-all-compounds-content)

6.1.2 成分靶点

Table 3 (下方表格) 为表格 tables of Herbs compounds and targets 概览。

(对应文件为 Figure+Table/tables-of-Herbs-compounds-and-targets.xlsx)

注：表格共有 1846 行 4 列，以下预览的表格可能省略部分数据；含有 3 个唯一 ‘Herb_pinyin_name’。

Table 3: Tables of Herbs compounds and targets

Herb_pinyin_name	Molecule name	symbols	protein.names
Banzhilian	luteolin	NA	NA
Banzhilian	luteolin	MMP2	72 kDa type IV co...
Banzhilian	luteolin	CLG4A	72 kDa type IV co...
Banzhilian	luteolin	ADCY2	Adenylate cyclase...
Banzhilian	luteolin	KIAA1060	Adenylate cyclase...
Banzhilian	luteolin	APP	Amyloid-beta prec...
Banzhilian	luteolin	A4	Amyloid-beta prec...
Banzhilian	luteolin	AD1	Amyloid-beta prec...
Banzhilian	luteolin	AR	Androgen receptor...
Banzhilian	luteolin	DHTR	Androgen receptor...
Banzhilian	luteolin	NR3C4	Androgen receptor...
Banzhilian	luteolin	XIAP	E3 ubiquitin-prot...
Banzhilian	luteolin	API3	E3 ubiquitin-prot...
Banzhilian	luteolin	BIRC4	E3 ubiquitin-prot...
Banzhilian	luteolin	IAP3	E3 ubiquitin-prot...
...

6.1.3 疾病靶点

Table 4 (下方表格) 为表格 Disease related targets from GeneCards 概览。

(对应文件为 Figure+Table/Disease-related-targets-from-GeneCards.xlsx)

注：表格共有 1746 行 7 列，以下预览的表格可能省略部分数据；含有 1746 个唯一 ‘Symbol’。

The GeneCards data was obtained by querying :

breast cancer

Restrict (with quotes) :

FALSE

Filtering by Score: :

Score > 15

Table 4: Disease related targets from GeneCards

Symbol	Description	Category	UniProt_ID	GIFtS	GC_id	Score
BRCA2	BRCA2 DNA ...	Protein Co...	P51587	56	GC13P032315	584.27
BRCA1	BRCA1 DNA ...	Protein Co...	P38398	59	GC17M043044	565.02
PALB2	Partner An...	Protein Co...	Q86YC2	53	GC16M023603	366.84
ATM	ATM Serine...	Protein Co...	Q13315	61	GC11P108223	340.7
CHEK2	Checkpoint...	Protein Co...	O96017	63	GC22M028687	336.43
BRIP1	BRCA1 Inte...	Protein Co...	Q9BX63	57	GC17M061679	325.07
CDH1	Cadherin 1	Protein Co...	P12830	58	GC16P068737	306.68
BARD1	BRCA1 Asso...	Protein Co...	Q99728	55	GC02M214725	291.41
TP53	Tumor Prot...	Protein Co...	P04637	62	GC17M007661	287.34
MSH6	MutS Homol...	Protein Co...	P52701	58	GC02P047695	239.29
MSH2	MutS Homol...	Protein Co...	P43246	57	GC02P047403	231.87
MLH1	MutL Homol...	Protein Co...	P40692	58	GC03P036993	223.25
C11orf65	Chromosome...	Protein Co...	Q8NCR3	40	GC11M108308	218.43
LOC126862571	BRD4-Indep...	Functional...		10	GC17P114574	215.91
APC	APC Regula...	Protein Co...	P25054	58	GC05P112707	199.23
...

6.1.4 疾病-成分-靶点网络图

Figure 2 (下方图) 为图 Network pharmacology with disease 概览。

(对应文件为 Figure+Table/Network-pharmacology-with-disease.pdf)

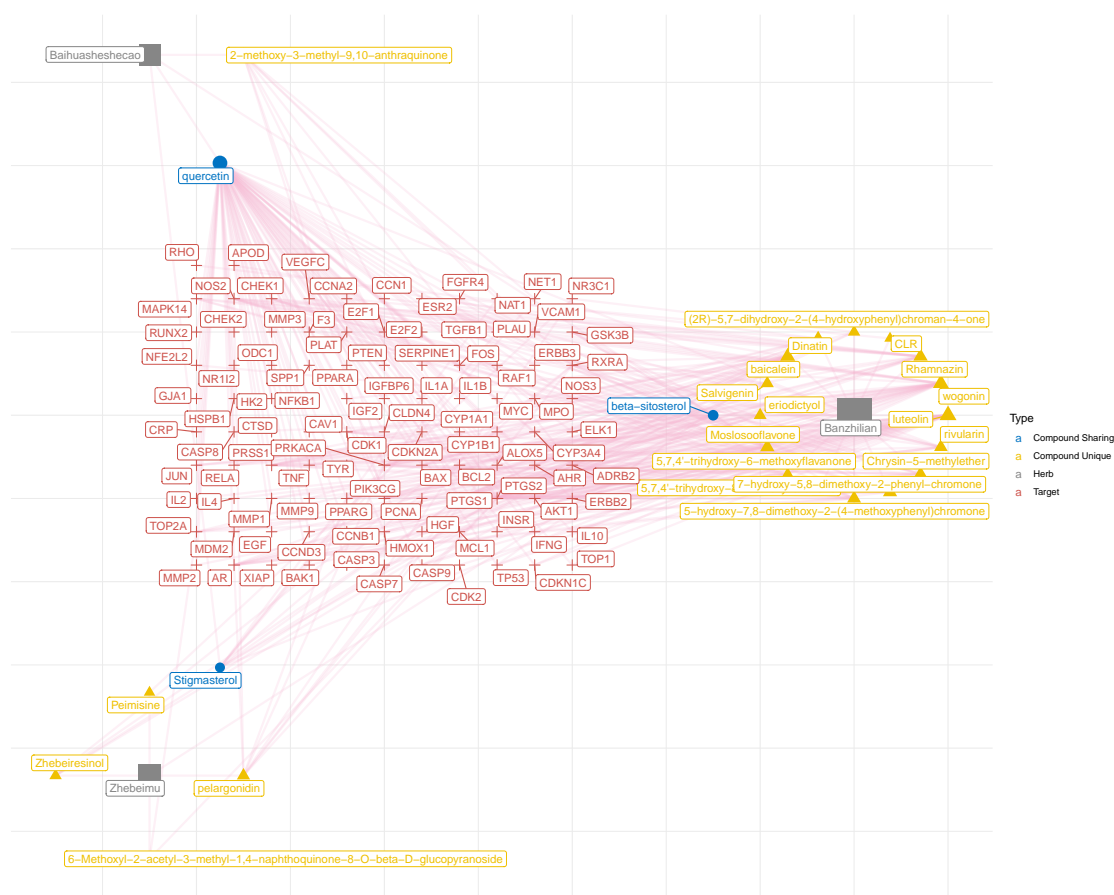


Figure 2: Network pharmacology with disease

Figure 3 (下方图) 为图 Targets intersect with targets of diseases 概览。

(对应文件为 Figure+Table/Targets-intersect-with-targets-of-diseases.pdf)



Figure 3: Targets intersect with targets of diseases

Intersection :

CHEK2, TP53, PTEN, ERBB2, CDKN2A, AKT1, AR, CASP8, ERBB3, JUN, MYC, IL2, MDM2, CDK2, IL1B, FGFR4, BCL2, BAX, TGFB1, ESR2, IGF2, NFE2L2, PPARG, EGF, PTGS2, TNF, MMP2, MMP9, RAF1, CASP3, CYP1A1, NFKB1, CTSD, PCNA, PLAU, TOP2A, CDK1, MMP1, E2F1, VEGFC, IFNG, CYP1B1, CHEK1, PIK3CG, IL10, CASP9, CAV1,...

(上述信息框内容已保存至 `Figure+Table/Targets-intersect-with-targets-of-diseases-content`)

6.1.5 富集分析

Figure 4 (下方图) 为图 KEGG enrichment 概览。

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

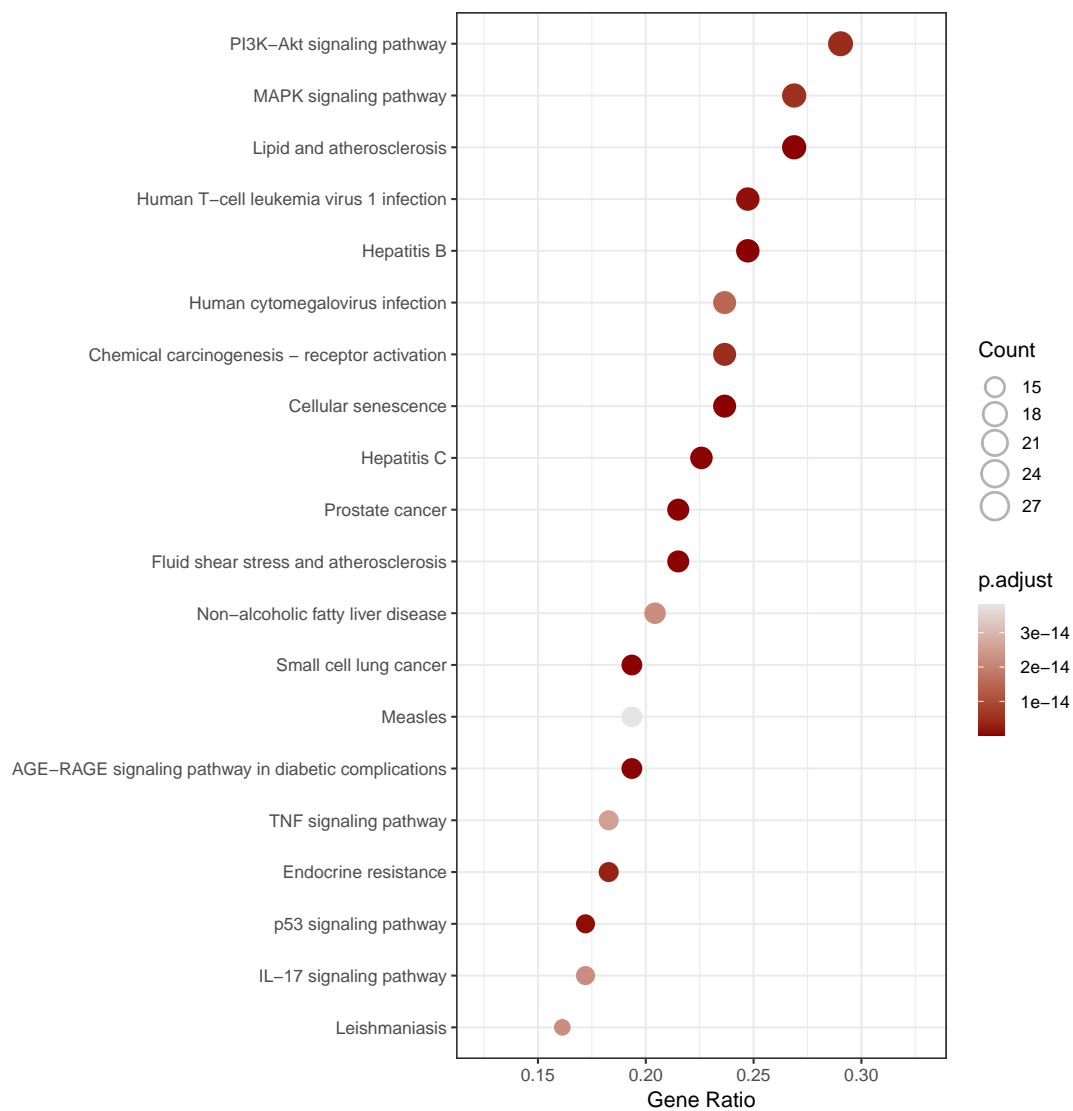


Figure 4: KEGG enrichment

Figure 5 (下方图) 为图 GO enrichment 概览。

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

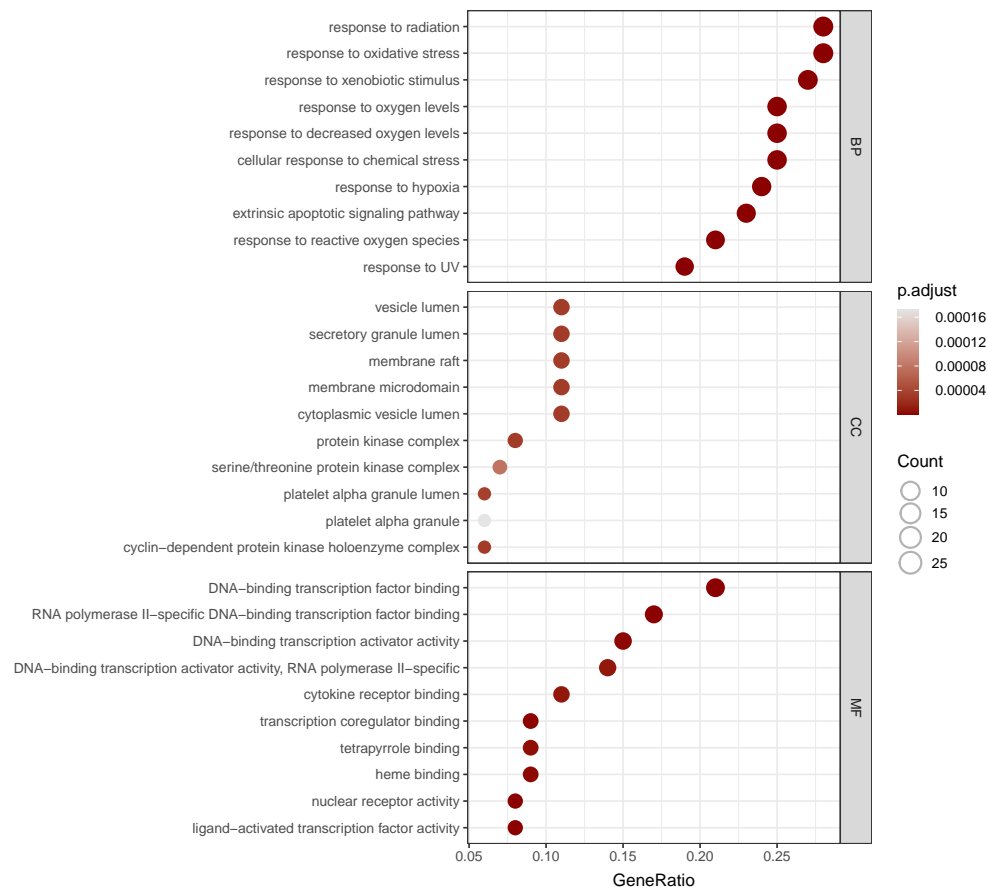


Figure 5: GO enrichment

6.1.6 疾病-成分-靶点-通路网络图

Figure 6 (下方图) 为图 Network pharmacology with disease and pathway 概览。

(对应文件为 Figure+Table/Network-pharmacology-with-disease-and-pathway.pdf)

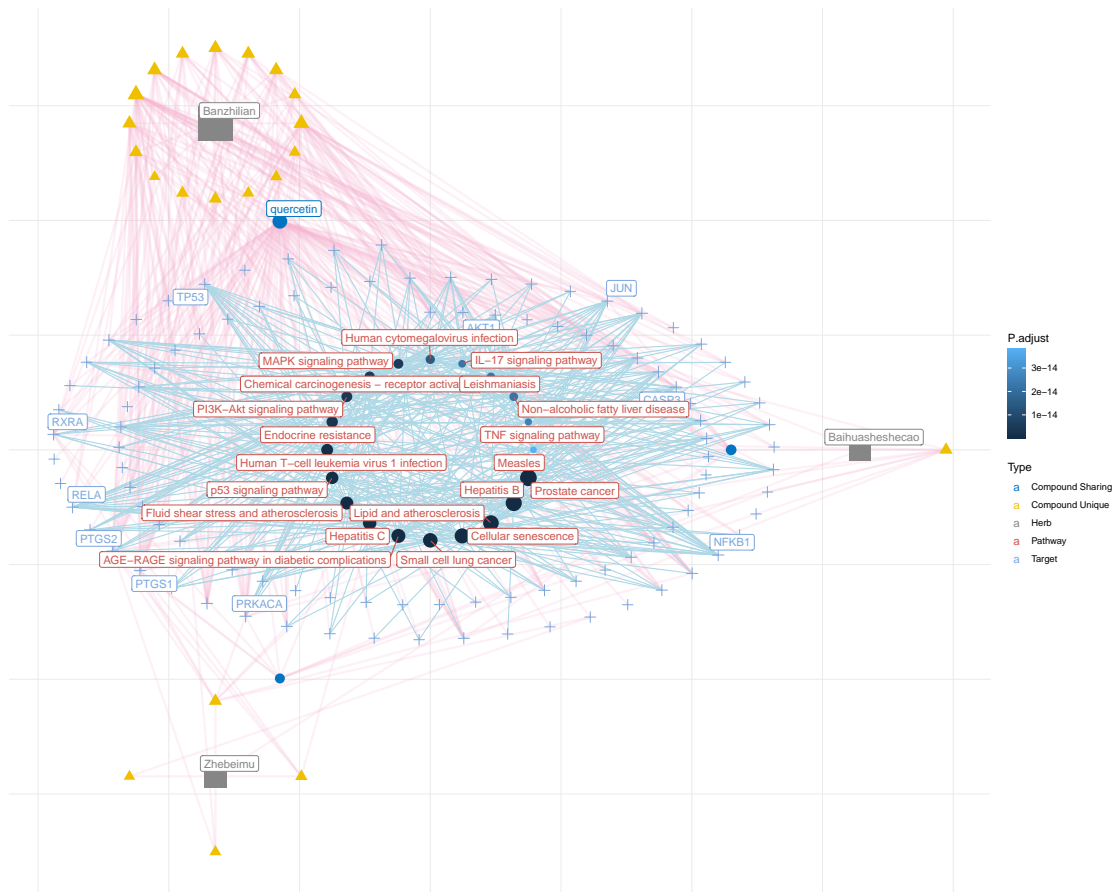


Figure 6: Network pharmacology with disease and pathway

Table 5 (下方表格) 为表格 Network pharmacology with disease and pathway data 概览。

(对应文件为 **Figure+Table/Network-pharmacology-with-disease-and-pathway-data.xlsx**)

注：表格共有 431 行 5 列，以下预览的表格可能省略部分数据；含有 3 个唯一 ‘Herb_pinyin_name’；含有 24 个唯一 ‘Ingredient.name’；含有 101 个唯一 ‘Target.name’。

Table 5: Network pharmacology with disease and pathway data

Herb_pinyin_name	Ingredient.name	Target.name	Hit_pathway_number	Enriched_pathways
Baihuasheshecao	quercetin	NFKB1	18	AGE-RAGE signalin...
Baihuasheshecao	quercetin	RELA	18	AGE-RAGE signalin...
Banzhilian	baicalein	RELA	18	AGE-RAGE signalin...
Banzhilian	luteolin	RELA	18	AGE-RAGE signalin...
Banzhilian	quercetin	NFKB1	18	AGE-RAGE signalin...
Banzhilian	quercetin	RELA	18	AGE-RAGE signalin...
Banzhilian	wogonin	RELA	18	AGE-RAGE signalin...
Baihuasheshecao	quercetin	AKT1	17	AGE-RAGE signalin...
Banzhilian	baicalein	AKT1	17	AGE-RAGE signalin...

Herb_pinyin_name	Ingredient.name	Target.name	Hit_pathway_number	Enriched_pathways
Banzhilian	luteolin	AKT1	17	AGE-RAGE signalin...
Banzhilian	quercetin	AKT1	17	AGE-RAGE signalin...
Banzhilian	wogonin	AKT1	17	AGE-RAGE signalin...
Baihuasheshecao	quercetin	TP53	14	Cellular senescen...
Banzhilian	baicalein	TP53	14	Cellular senescen...
Banzhilian	luteolin	TP53	14	Cellular senescen...
...

Reference

1. Wu, T. *et al.* ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* **2**, (2021).
2. 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).
3. Ru, J. *et al.* TCMSP: A database of systems pharmacology for drug discovery from herbal medicines. *Journal of cheminformatics* **6**, (2014).