

سهند نوعی ۹۹۲۳۰۸۷

تمرین سوم

۱. بخش اول: تحلیل بیان ژن تفاضلی با استفاده از ابزارهای GEO2R

وارد پایگاه داده GEO شوید و دیتاست GSE ۱۵۸۵۲ را پیدا کنید. این دیتاست شامل ۴۳ نمونه سرطان سینه و بافت های سالم متناظر آن است. با استفاده از ابزار GEO2R نمونه های سرطانی را با سالم مقایسه کنید.

۱. توضیح دهید در بیان ژن تفاضلی مشکل استفاده از **p-value** چیست؟ آزمایشی طراحی کنید که نشان دهد احتمال رخ دادن خطای نوع اول با افزایش تعداد آزمون ها چگونه تغییر می کند. درباره روش های اصلاح **p-value** تحقیق کنید و دو روش **Benjamini-Hochberg(BH)** و **Bonferroni** را توضیح دهید و تفاوت آنها را بیان کنید.

برای به دست آوردن **p-value** از روش های مختلفی میتوان استفاده کرد. مهم ترین آنها

Student's t-test, Standard score(z-score), chi-square test

است که با توجه به فرمول به دست آوردن هر یک از آنها(که در ادامه آورده شده است) مقادیر به دست آمده از آنها به اندازه جامعه آماری و تعداد نمونه ها بستگی دارد.

فرمول z-score:

$$Z = \frac{x - \mu}{\sigma}$$

Z = standard score

x = observed value

μ = mean of the sample

σ = standard deviation of the sample

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

σ = population standard deviation

N = the size of the population

x_i = each value from the population

μ = the population mean

فرمول student t:

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

t = Student's t-distribution

\bar{x} = sample mean

μ = population mean

s = sample standard deviation

n = sample size

فرمول chi-square:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

χ^2 = chi squared

O_i = observed value

E_i = expected value

$$E[X] = \sum_i x_i f(x_i)$$

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

با توجه به مقادیر به دست آمده از آنها به جدول مربوطه رجوع کرده و مقدار **p-value** را به دست می آوریم. بنابراین **p-value** به طور مستقیم به اندازه نمونه بستگی دارد و این مشکلات متعددی را ایجاد میکند.

برای بررسی مشکلات فرض میکنیم $\alpha = 0.05$ است:

۱. افزایش اندازه نمونه میتواند منجر به ایجاد خطای نوع اول شود یعنی فرض صفر میتواند به اشتباه رد شود. مثلاً اگر ۱۰ نمونه داشته باشیم و هیچ کدام از این ده نمونه تفاوت معناداری نداشته باشند هیچ کدام داده **significant** محسوب نمیشوند اما اگر از هر کدام از ۱۰ نمونه ۱۰۰ نمونه کپی کنیم و دوباره آزمایش را انجام بدهیم، انتظار میرود حدود ۵۰ داده **significant** گزارش شود در حالی که در عمل هیچ کدام **significant** نیستند.

آزمایش: اگر دو گروه داشته باشیم که تفاوت **significant**ی در آنها وجود نداشته باشند و ۱۰ داده با توزیع نرمال از آنها تولید کنیم داده **significant**ی با توجه به مقدار α احتمالاً وجود نخواهد داشت. حال اگر تعداد همین داده های مشابه را به ۲۰، ۱۰۰، ۱۰۰۰ و بیشتر افزایش دهیم، خواهیم دید که تعداد داده های **significant** بیشتر میشود چون در **z score**، مقدار n افزایش یافته، مقدار σ کاهش پیدا میکند، مقدار Z افزایش می یابد و در نهایت مقدار **p-value** کاهش می یابد و با توجه به ثابت ماندن مقدار α تعداد داده های **significant** افزایش می یابد در حالی که در عمل هیچ داده **significant**ی وجود نداشته است.

۲. برعکس این مورد هم صادق است. اگر تعداد داده ها کم باشد، تفاوت بین داده ها **significant** باشد و این را از قبل بدانیم اما مقدار α برای نشان دادن این تفاوت کافی نباشد، داده ای **significant** تشخیص داده نمیشود.

۳. در واقع انتخاب مقدار α به صورت سلیقه ای یا در حالت ایده آل بر اساس تجربه است و این میتواند برای بعضی آزمایش ها جواب بدهد اما برای بعضی نه. انتخاب نادرست مقدار **p-value** میتواند منجر به ایجاد

خطای اول (false positive) - بخصوص در تعداد داده های زیاد) یا خطای دوم (false negative) - بخصوص در داده های کم) بشود.


روش های مختلفی برای اصلاح مشکل p-value وجود دارد. از جمله معروف ترین آنها میتوان به موارد زیر اشاره کرد که در سایت NCBI هم ذکر شده است:

- Benjamini & Hochberg (False discovery rate)
- Benjamini & Yekutieli
- Bonferroni
- Holm

روش Bonferroni:

Family-wise error rate معیاری برای بررسی احتمال رخ دادن خطا با انجام m تست است و فرمول آن به شرح زیر است:

Number of tests


$$\text{FWER} = 1 - (1 - \alpha)^m$$

اگر نمودار آن را بررسی کنیم خواهیم دید که اگر یک تست داشته باشیم با مقدار α حدود ۰,۰۵ احتمال وجود false positive برابر ۰,۰۵ است که مطلوب ما هست اما اگر ۶۰ تست همزمان انجام شوند احتمال وجود false positive حدود ۹۵ درصد است. برای رفع این مشکل از روش bonferroni استفاده میشود که در همان حدود ۵ درصد ثابت نگه میدارد. در این روش مقدار α تقسیم بر تعداد تست میشود مثلاً اگر ۶۰ تست داشته باشیم ۰,۰۵/۶۰ میشود و این مقدار α جدید خواهد بود و مقدار p-value داده ها برای تشخیص differential significant بودن داده ها با این مقدار مقایسه میشود. همچنین میتوان به جای تغییر α مقدار p-value را مطابق سوال تغییر داد. برای این کار مقادیر p-value به دست آمده اولیه در تعداد تست ها ضرب میشود و این p-value ها با α قبلی مقایسه میشود.

روش Benjamin-Hochberg:

در این روش هم مجدد هم میتوان α را تغییر داد هم میتوان $p\text{-value}$ را تغییر داد. در اینجا من فقط روش تغییر $p\text{-value}$ را توضیح میدهم. در این روش ابتدا تست ها را به ترتیب $p\text{-value}$ بدست آمده به صورت صعودی مرتب میکنیم و با شروع از ۱ به هر کدام از این تست ها rank میدهم. حال $p\text{-value}$ جدید را از فرمول زیر بدست می آوریم:

$$P_{(k)} \frac{m}{k}$$

که در این رابطه $p\text{-value}$ هر کدام از تست ها در تعداد تست ها ضرب میشود و تقسیم بر رنک آن تست میشود. سپس از آخرین رنک به سمت اولین رنک به ترتیب زیر عمل میکنیم:

- در رنک m مقدار به دست آمده را $p\text{-value}$ جدید در نظر میگیریم.
- در رنک های دیگر در صورتی که $p\text{-value}$ به دست آمده در $k - 1$ کوچکتر مساوی $p\text{-value}$ رنک k بود، $p\text{-value}$ به دست آمده را به عنوان $p\text{-value}$ جدید در نظر میگیریم
- در صورتی که $p\text{-value}$ رنک $k - 1$ بزرگتر از k بود، $p\text{-value}$ به دست آمده برای رنک k را برای $k - 1$ در نظر میگیریم.

همانند شکل زیر:

Rank (k)	Sorted p-values	$p_{(k)} \frac{m}{k}$	BH adjusted p-value
1	0.005	$0.005 \times (8/1) = 0.040$	0.036
2	0.009	$0.009 \times (8/2) = 0.036$	0.036
3	0.019	$0.019 \times (8/3) = 0.051$	0.044
4	0.022	$0.022 \times (8/4) = 0.044$	0.044
5	0.051	$0.051 \times (8/5) = 0.082$	0.082
6	0.101	$0.101 \times (8/6) = 0.135$	0.135
7	0.361	$0.361 \times (8/7) = 0.413$	0.387
8	0.387	$0.387 \times (8/8) = 0.387$	0.387

برای انتخاب تست هایی که null hypothesis را رد میکنند از رنک آخر شروع میکنیم و به محض رسیدن به بزرگترین رنکی که مقدار p-value آن از مقدار α کمتر بود از آن رنک به قبل تست null hypothesis را رد میکنند و به عنوان داده differential significant محسوب میشوند.

تفاوت این دو روش آن است که روش bonferroni سخت گیرانه تر عمل میکند و تعداد تست های false positive در آن قطعا کوچکتر مساوی تعداد در BH است. بنابراین در مواردی که false positive نداشته بسیار اهمیت دارد روش Bonferroni مناسب تر است. در مواردی هم که false negative نداشته بسیار مهم است روش BH بهتر است.

۲. ژن هایی را معنادار در نظر بگیرید که $\text{adj p-value} > 0.001$ و $\log FC > 1$ باشد.

سپس نمودارهای Volcano plot و Mean Difference(MD) plot را رسم کنید و هر کدام را توضیح دهید. ژن های معنادار را ذخیره کنید. از آنها برای بخش دوم تمرین استفاده خواهیم کرد.

GEO2R
Options
Profile graph
R script

Apply adjustment to the P-values. [More...](#)
☒ Benjamini & Hochberg (False discovery rate)
☐ Benjamini & Yekutieli
☐ Bonferroni
☐ Holm

Apply log transformation to the data. [More...](#)
☒ Auto-detect
☐ Yes
☐ No

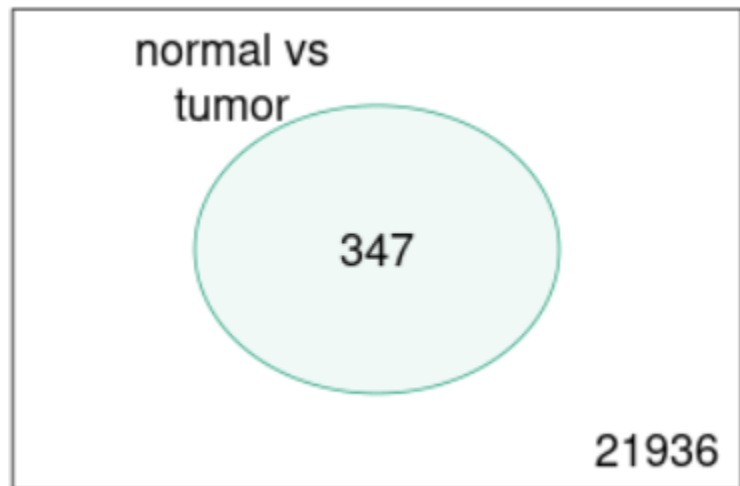
Category of Platform annotation to display on results.
☐ Submitter supplied
☒ NCBI generated

Plot displays. [More...](#)
Significance level cut-off (enter number between 0 and 1)
Log 2 fold change threshold
Volcano and Mean-difference plot contrasts (select up to 5)
1 selected (clear)
☒ normal vs tumor

Apply limma precision weights (vooma). [More...](#)
☐ Yes
☒ No

Force normalization. [More...](#)
☐ Yes
☒ No

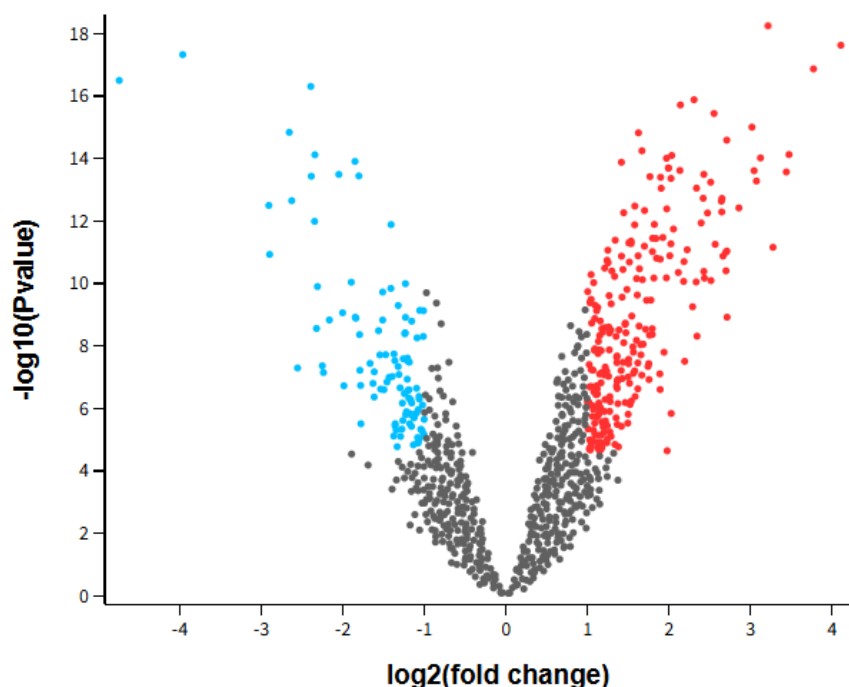
با توجه به این تنظیمات ۳۴۷ ژن معنادار پیدا شدند که فایل CSV آنها ضمیمه شده است. به دو گروه تقسیم شدند. گروه tumor و گروه normal



تحلیل نمودار Volcano plot:

Volcano plot
GSE15852: Expression data from human breast
tumors and their paired...
normal vs tumor, Padj<0.001

[Help](#)



با توجه به نمودار نقاط قرمز و آبی ژن های significant هستند. هر چه به سمت بالای محور yها میرویم نشان دهنده بیشتر significant بودن ژن ها است. چون هر چه $-\log(Pvalue)$ بیشتر میشود یعنی مقدار p-value به دست آمده آن کمتر است و در واقع احتمال این که واقعا موثر بوده است بیشتر میشود و احتمال false positive بودن آن کمتر است. هر چه به سمت دو سر نمودار میرویم اندازه تفاوت بین دو گروه بیشتر است. رابطه Fold Change به صورت زیر است:

$$\text{Fold Change (FC)} = \frac{\text{Expression of gene X in tumor}}{\text{Expression of gene X in normal}}$$

اگر لگاریتم این رابطه در مبنای ۲ را بگیریم $\log_2 FC$ به دست می آید و هر چه این عدد از صفر فاصله بگیرد یعنی میزان تفاوت بیان ژن در دو گروه بیشتر است. حال هر چه بزرگ تر از صفر باشد یعنی میزان بیان ژن X در داده های

داده های تومور و ناسالم بیشتر است (upregulated) که با قرمز نشان داده شده است و برعکس هر چه منفی تر باشد (downregulated) یعنی در داده های سالم با میزان بیشتری تفاوت بیان بیشتری دارد که با آبی نشان داده شده است. بنابراین اگر بخواهیم ژن هایی که نسبت به بقیه با احتمال بیشتری true positive هستند را در بیاوریم میتوان به صورت زیر اشاره کرد:

RBP4, PDE3B, PPP1R1A

Significantترین با بیشترین تفاوت بیان ژن که در بیماران بیشترین بیان را دارند

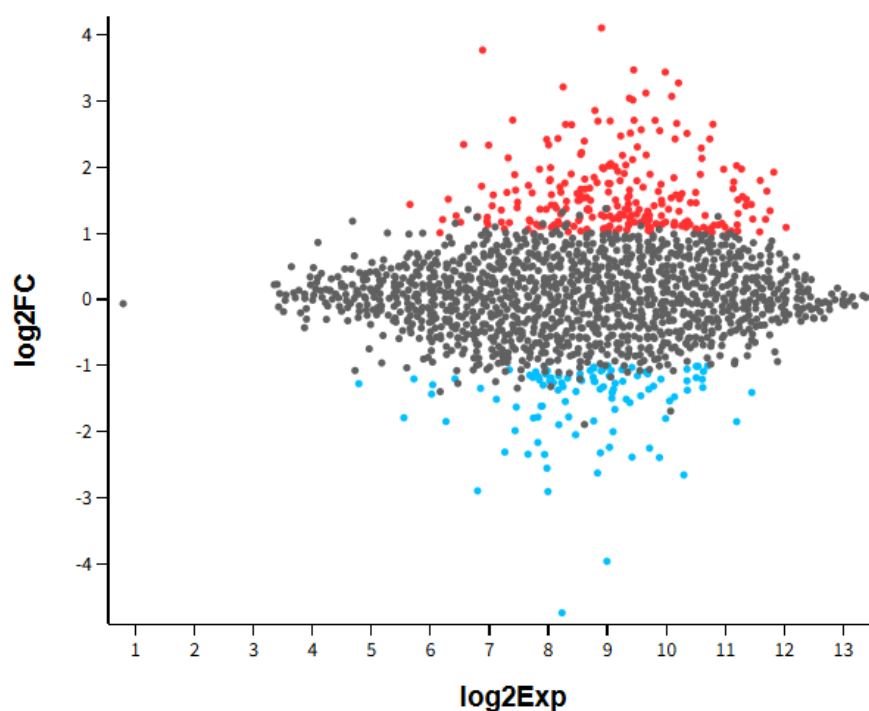
KRT19, CD24, EPCAM, TACSTD2

Significantترین با بیشترین تفاوت بیان ژن که در افراد سالم بیشترین بیان را دارند.

تحلیل نمودار Mean Difference plot:

Meandiff plot
GSE15852: Expression data from human breast
tumors and their paired...
normal vs tumor, Padj<0.001

[Help](#)



محور افقی لگاریتم میانگین بیان هر ژن در دو گروه را نشان میدهد و در واقع شدت سیگنال فلورسنت حاصل است. میبینیم که اکثر ژن ها اقل شدت بیان ژن ۸ را دارند. اکثرا ژن هایی که بیشترین و کمترین بیان ژن را داشتند جز ژن های significant نیستند. همچنین احتمالا دستگاه اندازه گیری هم سالم بوده است چون نسبت به شدت بیان خاصی بایاس نداشته است.

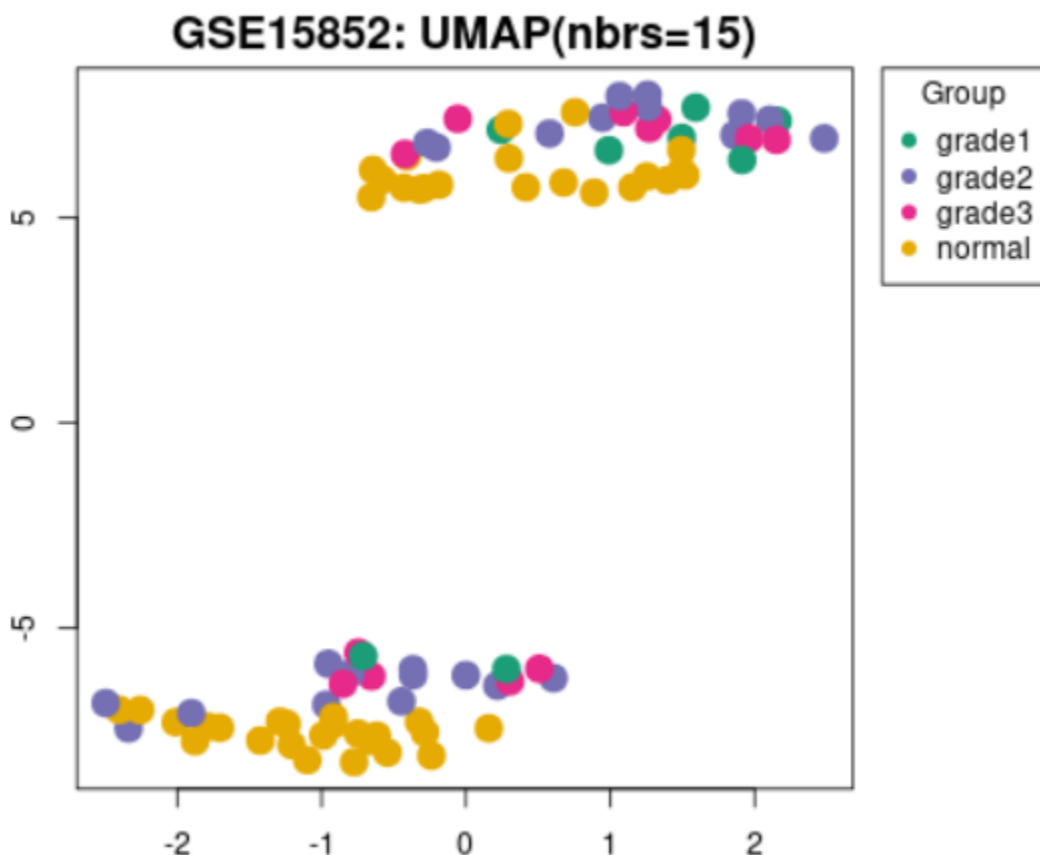
با توجه به نتایج بدست آمده:

ژن های ۴RBP و PDE3B و ۲SAA و CFD به میزان خوبی بیان شده اند و fold change زیادی دارند و در سلول های سرطانی احتمالا موثرترند. ژن ۱FHL ژن موثر در سرطان است که به بیشترین بیان را داشته است.

همچنین ژن های ۱۹KRT و ۲۴CD و ۲TACSTD بیشترین تفاوت بیان در سلول های سالم را دارند. ژن ۲۴CD بیشترین میزان بیان در سلول های سالم را در میان ژن های significant داشته است.

۳. در انتها، نمونه‌ها را براساس **Grade** به چهار گروه تقسیم کنید. سپس مجدد آنالیز را انجام دهید. نمودارهای **UMAP**، **Venn diagram**، **Boxplot** را رسم کنید و آنها را تحلیل کنید.

UMAP plot

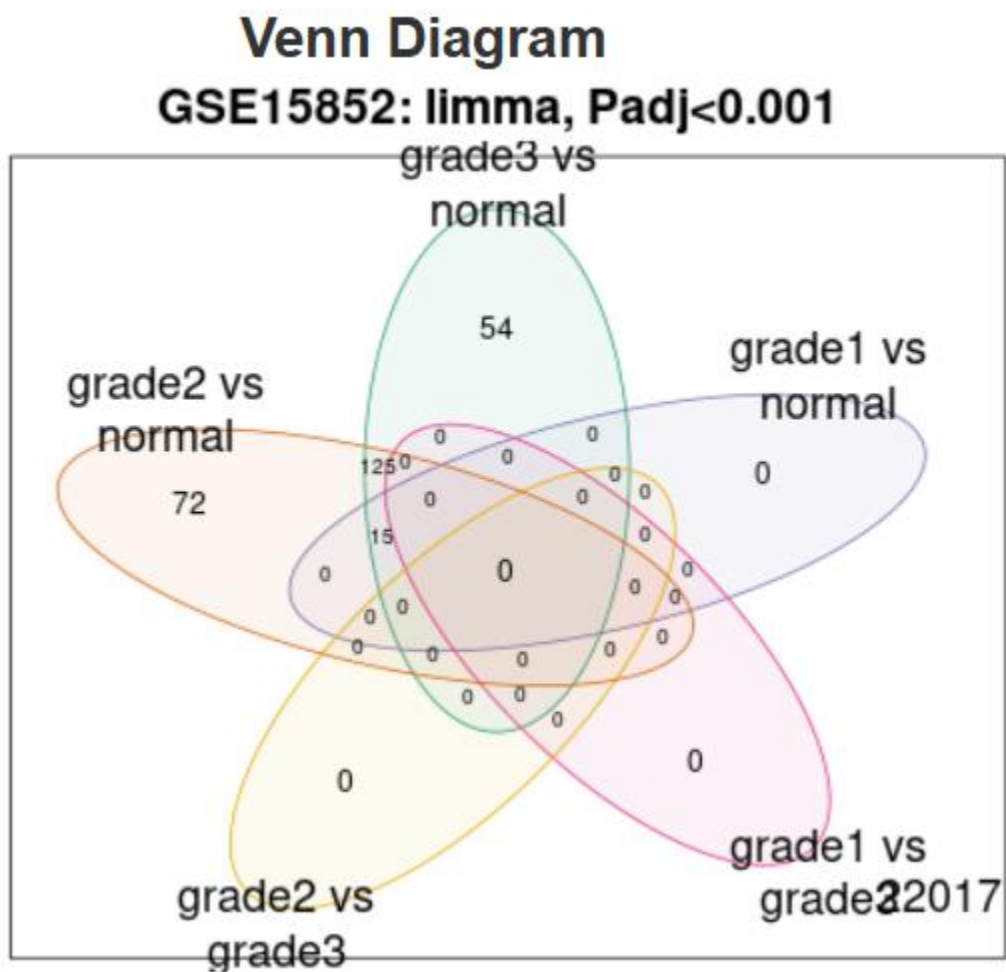


تحلیل نمودار UMAP:

UMAP یک روش برای dimensionality reduction است که در اینجا به دو بعد کاهش یافته است. با توجه به وجود نیامدن clusterهای مجزا، اطلاعات خاصی از این نمودار به دست نمی‌آید. این نشان دهنده

آن است که دو بعد برای دریافت اطلاعات significant کافی نیست یا تعداد neighborها به درستی انتخاب نشده است.

تحلیل Venn diagram:



با توجه به نمودار ۲۲۰۱۷ ژن بی تاثیر هستند.

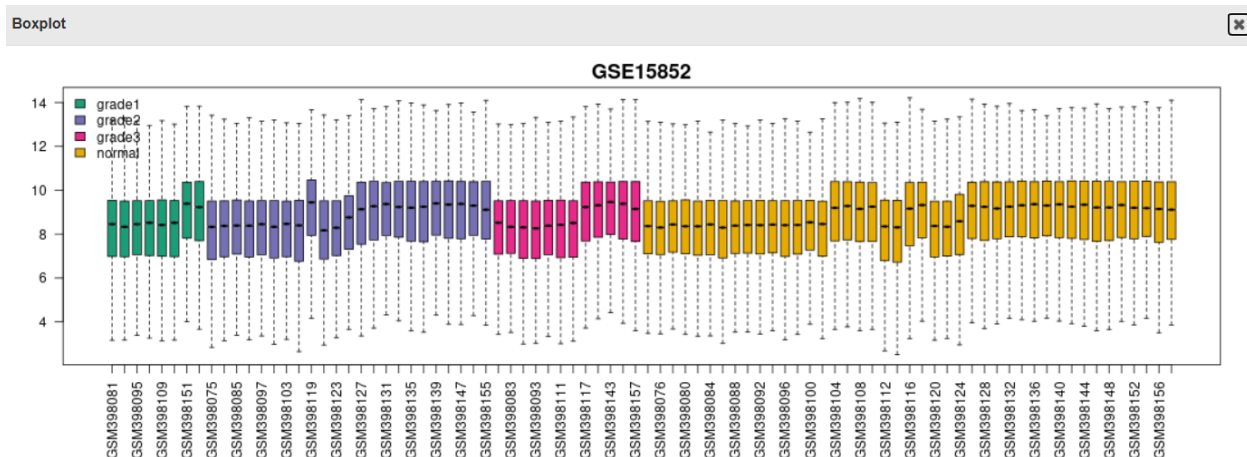
۷۲ ژن significant در مقایسه grade ۲ و normal پیدا شده است

۵۴ ژن significant در مقایسه grade ۳ و normal پیدا شده است

۱۵ ژن significant هم در مقایسه grade ۲ و normal و هم در مقایسه grade ۳ و normal و هم در مقایسه grade ۱ و normal پیدا شده است.

۱۲۵ ژن significant هم در مقایسه grade ۲ و normal و هم در مقایسه grade ۳ و normal پیدا شده است.

تحلیل Box plot:



خط داخل جعبه نشان دهنده میانه هست

دو سر جعبه محدوده چارک اول و سوم را نشان می دهد

دو سر هر نمودار جعبه ای محدوده داده ها را مشخص میکند

در نهایت با توجه به نمودار داده outlier در این بررسی احتمالاً وجود نداشته و داده ها دارای کیفیت خوبی بوده اند و قبل از آزمایش داده های پرت حذف شده اند. بنابراین میتوان به نتایج آزمایش استناد کرد.

بخش دوم: غنی سازی مسیرهای زیستی

در این بخش، ژنهای معنادار شناسایی شده در مرحله قبل را با استفاده از ابزار **Enrichr** از نظر مسیرهای زیستی و عملکردهای ژنی تحلیل کنید. برای این منظور از پایگاه داده های **KEGG** و **GO** استفاده کنید.

۱. تحلیل مسیرهای زیستی در **KEGG** :

- لیست ژنهای معنادار را در ابزار **Enrichr** وارد کنید و مسیرهای زیستی معنادار در پایگاه داده **KEGG** را شناسایی کنید. توضیح دهید که این مسیرهای زیستی چه نقشی دارند و چگونه با پیشرفت یا شکل گیری سرطان مرتبط هستند. نتایج را با جداول و نمودارهای مناسب ارائه دهید و نکات برجسته هر مسیر زیستی را توضیح دهید.

ابتدا سمبل ژن های significant را وارد میکنیم که ۳۴۷ عدد هستند. ژن هایی که در خروجی **geo2r** تکراری بودن مثل **HBA2///HBA** در برخی موارد مورد اول را نگه داشتیم و در برخی موارد مورد دوم را. به نظرم این کار منطقی تر از نگه داری فقط یکی از آنها باشد هرچند در نتیجه کار ممکن است تاثیری نداشته باشد. نرم افزار بعد از حذف کردن تکراری ها ۲۸۵ ژن را در نظر میگیرد.

Description
BioinformaticHW3_SahandNoey (285 genes)

KEGG:

KEGG 2021 Human

Bar Graph

Table

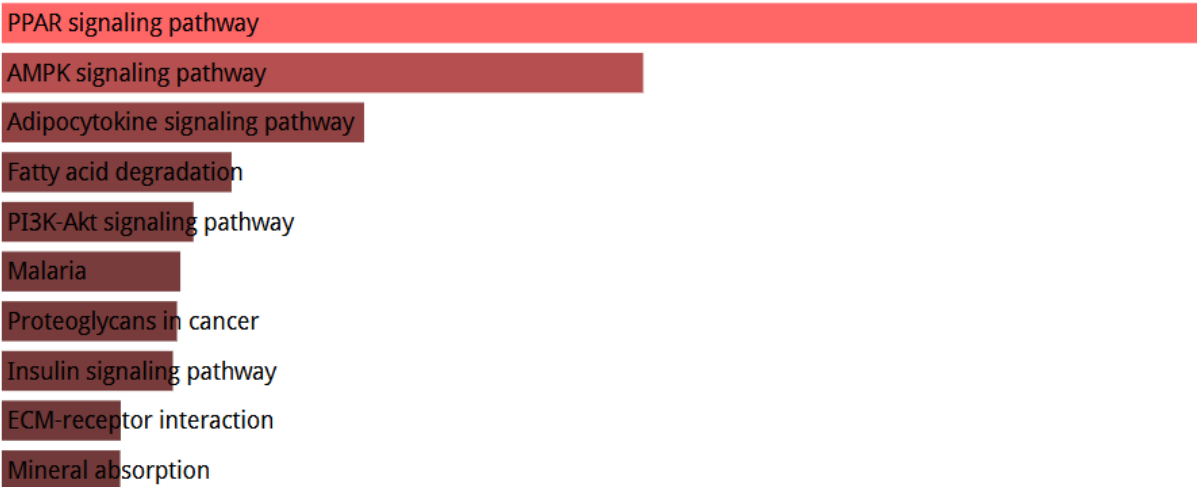
Clustergram

Appyter



Click the bars to sort. Now sorted by **p-value ranking**.

SVG PNG JPG



KEGG 2021 Human

Bar Graph

Table

Clustergram

Appyter



Hover each row to see the overlapping genes.

100

entries per page

Search:

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	PPAR signaling pathway	5.561e-10	1.424e-7	13.93	296.92
2	AMPK signaling pathway	0.000001191	0.0001525	7.22	98.50
3	Adipocytokine signaling pathway	0.00005622	0.004797	7.98	78.11
4	Fatty acid degradation	0.0003506	0.02244	9.25	73.57
5	PI3K-Akt signaling pathway	0.0005938	0.02526	2.94	21.87
6	Malaria	0.0007115	0.02526	7.81	56.58
7	Proteoglycans in cancer	0.0007459	0.02526	3.64	26.21
8	Insulin signaling pathway	0.0007893	0.02526	4.38	31.33
9	ECM-receptor interaction	0.001620	0.03865	5.15	33.08
10	Mineral absorption	0.001634	0.03865	6.38	40.96
11	Tyrosine metabolism	0.001661	0.03865	8.76	56.04
12	Focal adhesion	0.002439	0.04475	3.32	19.95
13	Fc gamma R-mediated phagocytosis	0.002657	0.04475	4.64	27.50
14	Ferroptosis	0.002701	0.04475	7.57	44.77
15	FoxO signaling pathway	0.002756	0.04475	3.98	23.45
16	Choline metabolism in cancer	0.002797	0.04475	4.59	26.97
17	Histidine metabolism	0.003608	0.05160	11.03	62.03
18	Melanoma	0.003646	0.05160	5.24	29.40
19	Fluid shear stress and atherosclerosis	0.003829	0.05160	3.74	20.79
20	Proximal tubule bicarbonate reclamation	0.004106	0.05255	10.48	57.57
21	Pyruvate metabolism	0.004450	0.05425	6.51	35.26
22	Ovarian steroidogenesis	0.005965	0.06930	5.96	30.51
23	Ras signaling pathway	0.006226	0.06930	2.85	14.48
24	Non-alcoholic fatty liver disease	0.006894	0.07354	3.33	16.57
25	Regulation of lipolysis in adipocytes	0.007789	0.07976	5.49	26.65

N
M

26	Glutathione metabolism	0.008824	0.08688	5.28	24.98
27	Pathways in cancer	0.009521	0.09027	2.07	9.62
28	Viral myocarditis	0.01054	0.09519	5.00	22.75
29	Tight junction	0.01082	0.09519	3.04	13.76
30	Steroid hormone biosynthesis	0.01115	0.09519	4.91	22.07
31	Glycerophospholipid metabolism	0.01315	0.1086	3.77	16.32
32	Estrogen signaling pathway	0.01384	0.1107	3.21	13.76
33	Longevity regulating pathway	0.01541	0.1134	3.61	15.07
34	African trypanosomiasis	0.01556	0.1134	6.16	25.63
35	Aldosterone-regulated sodium reabsorption	0.01556	0.1134	6.16	25.63
36	Amphetamine addiction	0.01692	0.1134	4.30	17.55
37	Renal cell carcinoma	0.01692	0.1134	4.30	17.55
38	Epithelial cell signaling in Helicobacter pylori infection	0.01776	0.1134	4.24	17.08
39	Glucagon signaling pathway	0.01858	0.1134	3.43	13.68
40	Adherens junction	0.01861	0.1134	4.17	16.63
41	Breast cancer	0.01894	0.1134	2.99	11.84
42	Drug metabolism	0.01927	0.1134	3.40	13.43
43	Insulin resistance	0.01927	0.1134	3.40	13.43
44	HIF-1 signaling pathway	0.01997	0.1134	3.37	13.18
45	Transcriptional misregulation in cancer	0.02045	0.1134	2.66	10.34
46	Bladder cancer	0.02049	0.1134	5.51	21.42
47	Human papillomavirus infection	0.02083	0.1134	2.20	8.51
48	Tryptophan metabolism	0.02184	0.1165	5.37	20.53
49	Glutamatergic synapse	0.02371	0.1218	3.21	12.02
50	Phenylalanine metabolism	0.02390	0.1218	9.28	34.66
51	Bacterial invasion of epithelial cells	0.02427	0.1218	3.83	14.24
52	MAPK signaling pathway	0.02561	0.1261	2.22	8.15
53	Fatty acid biosynthesis	0.02664	0.1287	8.70	31.54
54	JAK-STAT signaling pathway	0.02881	0.1366	2.70	9.56
55	One carbon pool by folate	0.03247	0.1464	7.73	26.51
56	Cell cycle	0.03249	0.1464	2.94	10.08

57	Cocaine addiction	0.03260	0.1464	4.55	15.57
58	Complement and coagulation cascades	0.03329	0.1469	3.45	11.74
59	Cholesterol metabolism	0.03433	0.1490	4.45	15.01
60	Osteoclast differentiation	0.03547	0.1494	2.87	9.58
61	cAMP signaling pathway	0.03561	0.1494	2.35	7.84
62	PD-L1 expression and PD-1 checkpoint pathway in cancer	0.03844	0.1587	3.29	10.71
63	Hypertrophic cardiomyopathy	0.03979	0.1617	3.25	10.48
64	Rheumatoid arthritis	0.04402	0.1761	3.14	9.80
65	Staphylococcus aureus infection	0.04697	0.1850	3.07	9.39
66	Dilated cardiomyopathy	0.04848	0.1881	3.04	9.19
67	Prostate cancer	0.05003	0.1911	3.00	9.00
68	Hematopoietic cell lineage	0.05319	0.2003	2.94	8.63
69	Long-term depression	0.05410	0.2003	3.67	10.70
70	Fatty acid elongation	0.05623	0.2003	5.57	16.02
71	Arachidonic acid metabolism	0.05632	0.2003	3.61	10.37
72	Glycerolipid metabolism	0.05632	0.2003	3.61	10.37
73	Protein digestion and absorption	0.05984	0.2081	2.82	7.94
74	Cell adhesion molecules	0.06096	0.2081	2.44	6.84
75	Phospholipase D signaling pathway	0.06096	0.2081	2.44	6.84
76	Pathogenic Escherichia coli infection	0.06342	0.2136	2.20	6.06
77	Citrate cycle (TCA cycle)	0.06779	0.2225	4.97	13.37
78	beta-Alanine metabolism	0.06779	0.2225	4.97	13.37
79	Endocytosis	0.06986	0.2257	2.00	5.33
80	Glycolysis / Gluconeogenesis	0.07054	0.2257	3.27	8.66
81	Serotonergic synapse	0.07829	0.2474	2.56	6.52
82	Leukocyte transendothelial migration	0.08028	0.2488	2.54	6.40
83	Rap1 signaling pathway	0.08065	0.2488	2.06	5.18
84	Hepatitis B	0.08250	0.2514	2.22	5.55
85	Human immunodeficiency virus 1 infection	0.08352	0.2515	2.04	5.06
86	p53 signaling pathway	0.08619	0.2566	2.99	7.32
87	Glioma	0.09170	0.2659	2.90	6.93
88	Thyroid hormone synthesis	0.09170	0.2659	2.90	6.93
89	Regulation of actin cytoskeleton	0.09244	0.2659	1.98	4.71
90	Chronic myeloid leukemia	0.09452	0.2688	2.86	6.75
91	Arrhythmogenic right ventricular cardiomyopathy	0.09737	0.2709	2.82	6.58
92	Thyroid cancer	0.09737	0.2709	3.97	9.26
93	Sulfur relay system	0.1085	0.2986	9.91	22.02
94	Glycine, serine and threonine metabolism	0.1110	0.3011	3.66	8.05
95	Peroxisome	0.1121	0.3011	2.64	5.79
96	Relaxin signaling pathway	0.1129	0.3011	2.23	4.87
97	Coronavirus disease	0.1152	0.3040	1.85	4.01
98	Dopaminergic synapse	0.1201	0.3136	2.18	4.62
99	Colorectal cancer	0.1245	0.3201	2.52	5.24
100	Fat digestion and absorption	0.1250	0.3201	3.39	7.05

اگر در اینجا هم p-value های کوچکتر از ۰,۰۰۱ را در نظر بگیریم ۸ مسیر زیستی معنادار به دست می آید

1	PPAR signaling pathway	5.561e-10	1.424e-7	13.9 3	296.9 2
2	AMPK signaling pathway	0.00000119 1	0.000152 5	7.22	98.50
3	Adipocytokine signaling pathway	0.00005622	0.004797	7.98	78.11
4	Fatty acid degradation	0.0003506	0.02244	9.25	73.57
5	PI3K-Akt signaling pathway	0.0005938	0.02526	2.94	21.87
6	Malaria	0.0007115	0.02526	7.81	56.58
7	Proteoglycans in cancer	0.0007459	0.02526	3.64	26.21
8	Insulin signaling pathway	0.0007893	0.02526	4.38	31.33

و ژن‌های مرتبط با این مسیرهای زیستی نشان داده شده‌اند.

KEGG 2021 Human

Bar Graph

Table

Clustergram

Appyter

Enriched Terms are the columns, input genes are the rows, and cells in the matrix indicate if a gene is associated with a term.



Row Order

Cluster

Sum

Column Order

Cluster

Sum

Gene

Search

Combined Score

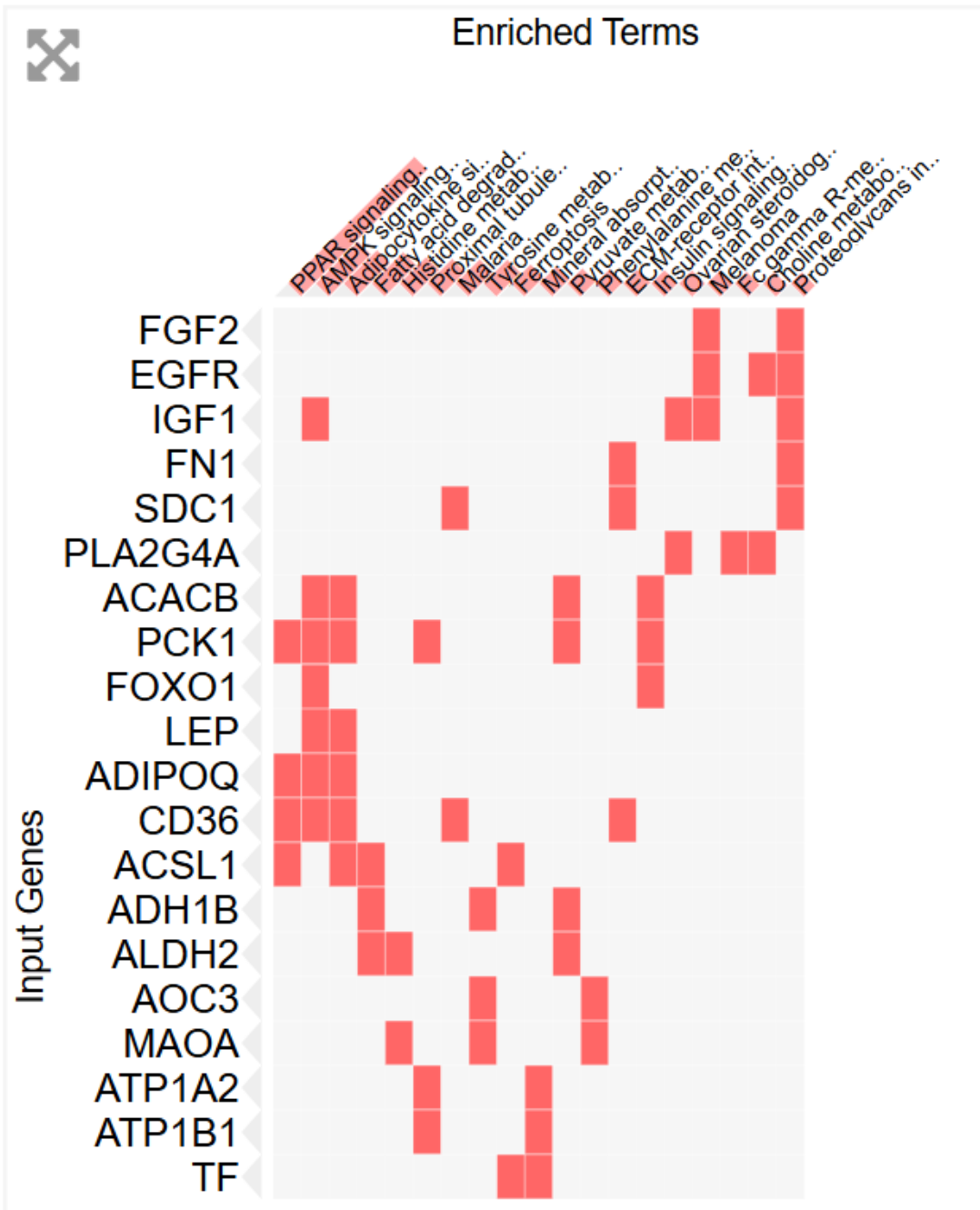
P-Value

Z-score

Top Enriched Terms: 20



Top rows sum: 20 rows



PPAR Signaling Pathway

نقش: این مسیر در تنظیم متابولیسم اسیدهای چرب و هومئوستاز گلوکز نقش دارد و همچنین باعث مهار التهاب و استرس اکسیداتیو می‌شود.

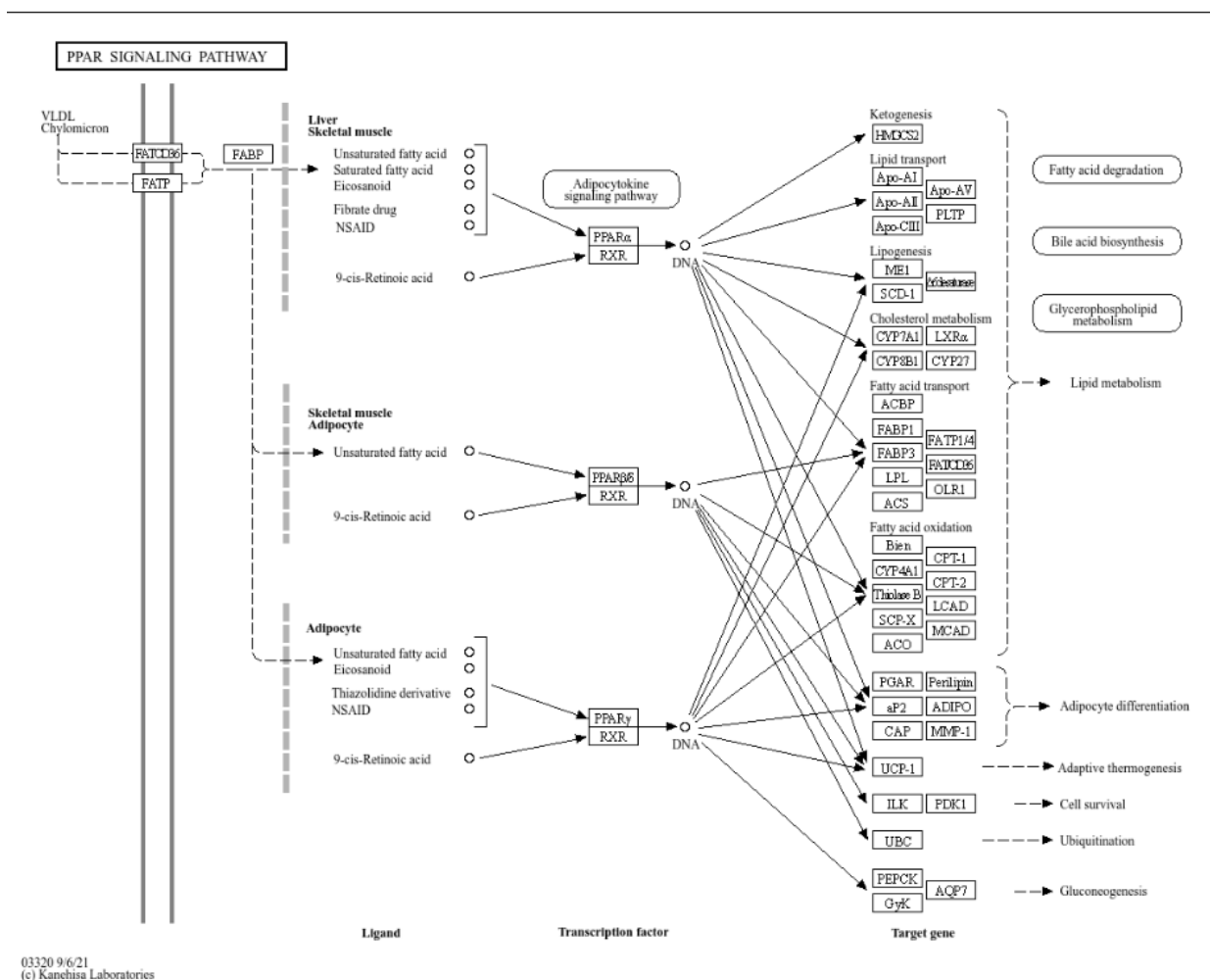
ارتباط با سرطان سینه: اختلال در عملکرد این مسیر می‌تواند باعث تغییر در متابولیسم انرژی و ایجاد شرایطی مانند التهاب مزمن شود که با رشد و بقای سلول‌های سرطانی در سرطان سینه مرتبط است.



PATHWAY: map03320

[Help](#)

Entry	map03320	Pathway
Name	PPAR signaling pathway	
Description	Peroxisome proliferator-activated receptors (PPARs) are nuclear hormone receptors that are activated by fatty acids and their derivatives. PPAR has three subtypes (PPARalpha, beta/delta, and gamma) showing different expression patterns in vertebrates. Each of them is encoded in a separate gene and binds fatty acids and eicosanoids. PPARalpha plays a role in the clearance of circulating or cellular lipids via the regulation of gene expression involved in lipid metabolism in liver and skeletal muscle. PPARbeta/delta is involved in lipid oxidation and cell proliferation. PPARGgamma promotes adipocyte differentiation to enhance blood glucose uptake.	
Class	Organismal Systems; Endocrine system BRITE hierarchy	



AMPK Signaling Pathway

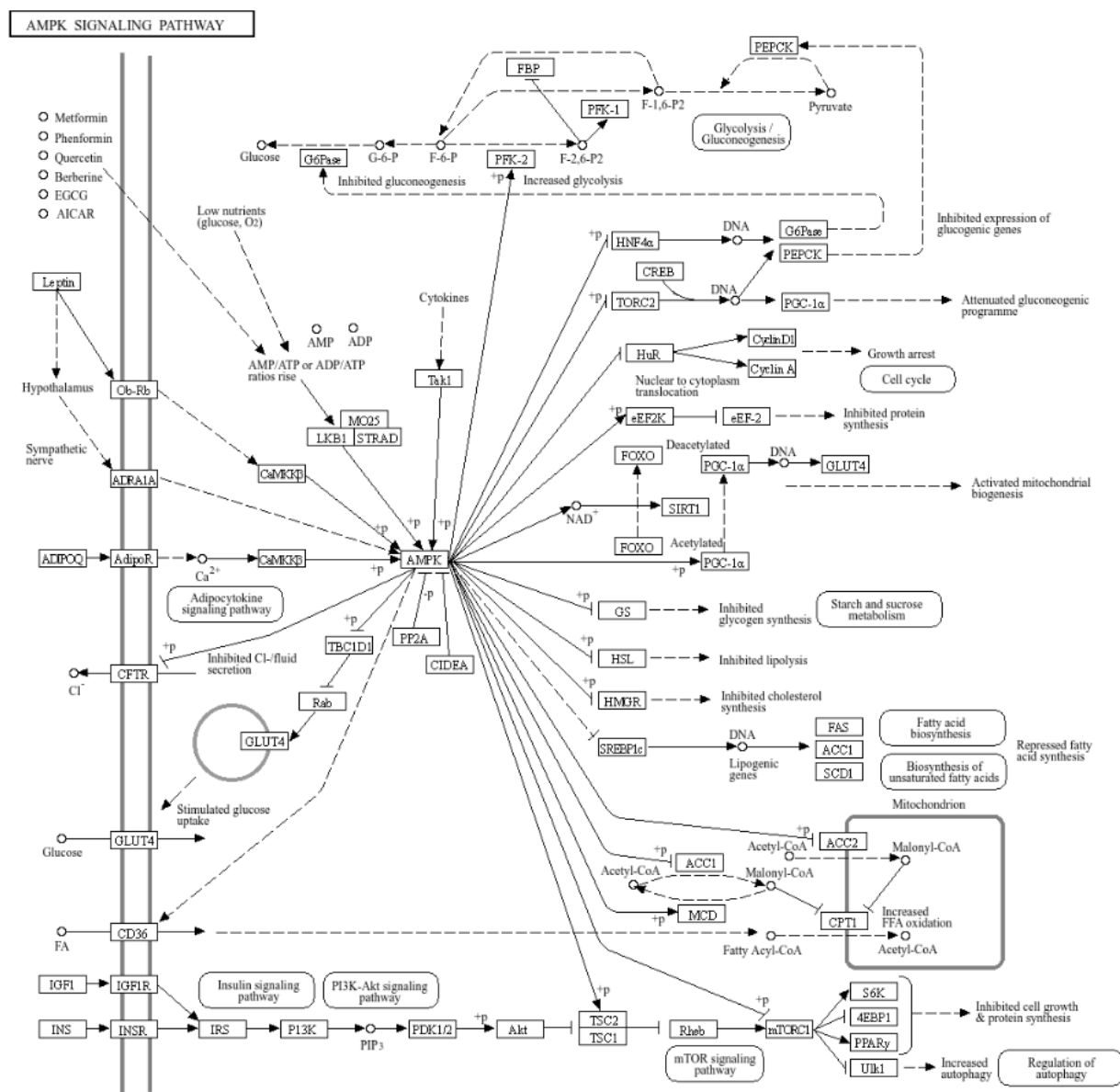
نقش: AMPK به عنوان حسگر انرژی سلولی، تعادل انرژی را تنظیم کرده و مسیرهای مرتبط با رشد سلولی را کنترل می‌کند.

ارتباط با سرطان سینه: سرکوب یا فعال‌سازی غیرطبیعی AMPK می‌تواند منجر به تنظیم غیرطبیعی رشد و تکثیر سلولی در سرطان سینه شود. AMPK معمولاً به عنوان یک مسیر محافظ در برابر رشد سلول‌های سرطانی عمل می‌کند.

Entry	map04152	Pathway
Name	AMPK signaling pathway	
Description	AMP-activated protein kinase (AMPK) is a serine threonine kinase that is highly conserved through evolution. AMPK system acts as a sensor of cellular energy status. It is activated by increases in the cellular AMP:ATP ratio caused by metabolic stresses that either interfere with ATP production (eg, deprivation for glucose or oxygen) or that accelerate ATP consumption (eg, muscle contraction). Several upstream kinases, including liver kinase B1 (LKB1), calcium/calmodulin kinase kinase-beta (CaMKK beta), and TGF-beta-activated kinase-1 (TAK-1), can activate AMPK by phosphorylating a threonine residue on its catalytic alpha-subunit. Once activated, AMPK leads to a concomitant inhibition of energy-consuming biosynthetic pathways, such as protein, fatty acid and glycogen synthesis, and activation of ATP-producing catabolic pathways, such as fatty acid oxidation and glycolysis.	
Class	Environmental Information Processing; Signal transduction	
	BRITE hierarchy	

All links

[Drug \(1\)](#)
[KEGG DRUG \(1\)](#)
[Chemical substance \(22\)](#)
[KEGG COMPOUND \(22\)](#)
[Gene \(267140\)](#)
[KEGG ORTHOLOGY \(85\)](#)
[RefGene \(267055\)](#)
[Literature \(20\)](#)
[PubMed \(20\)](#)
[All databases \(267183\)](#)
[Download RDF](#)

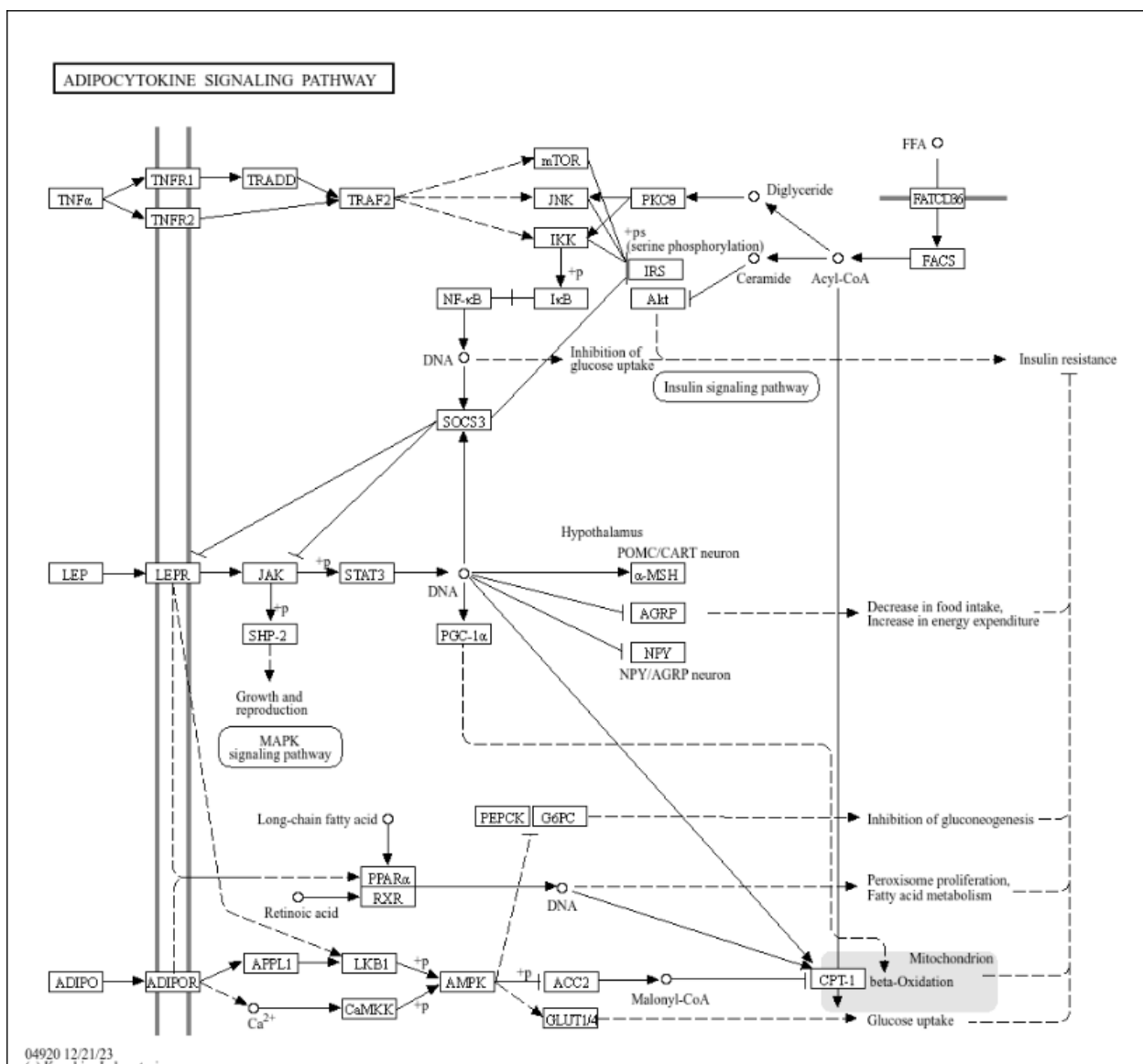


Adipocytokine Signaling Pathway

نقش: این مسیر شامل سیگنال‌دهی آدیپوکاین‌ها (مانند لپتین و آدیپونکتین) است که متابولیسم انرژی، التهاب و حساسیت به انسولین را تنظیم می‌کند.

ارتباط با سرطان سینه: لپتین که در این مسیر نقش دارد، می‌تواند رشد و مهاجرت سلول‌های سرطانی را تحریک کند. سطوح پایین آدیپونکتین نیز با افزایش خطر سرطان سینه همراه است.

Entry	map04920	Pathway
Name	Adipocytokine signaling pathway	
Description	<p>Increased adipocyte volume and number are positively correlated with leptin production, and negatively correlated with production of adiponectin.</p> <p>Leptin is an important regulator of energy intake and metabolic rate primarily by acting at hypothalamic nuclei. Leptin exerts its anorectic effects by modulating the levels of neuropeptides such as NPY, AGRP, and alpha-MSH. This leptin action is through the JAK kinase, STAT3 phosphorylation, and nuclear transcriptional effect.</p> <p>Adiponectin lowers plasma glucose and FFAs. These effects are partly accounted for by adiponectin-induced AMPK activation, which in turn stimulates skeletal muscle fatty acid oxidation and glucose uptake. Furthermore, activation of AMPK by adiponectin suppresses endogenous glucose production, concomitantly with inhibition of PEPCK and G6Pase expression.</p> <p>The proinflammatory cytokine TNFalpha has been implicated as a link between obesity and insulin resistance. TNFalpha interferes with early steps of insulin signaling. Several data have shown that TNFalpha inhibits IRS1 tyrosine phosphorylation by promoting its serine phosphorylation. Among the serine/threonine kinases activated by TNFalpha, JNK, mTOR and IKK have been shown to be involved in this phosphorylation.</p>	
Class	Organismal Systems; Endocrine system BRITe hierarchy	

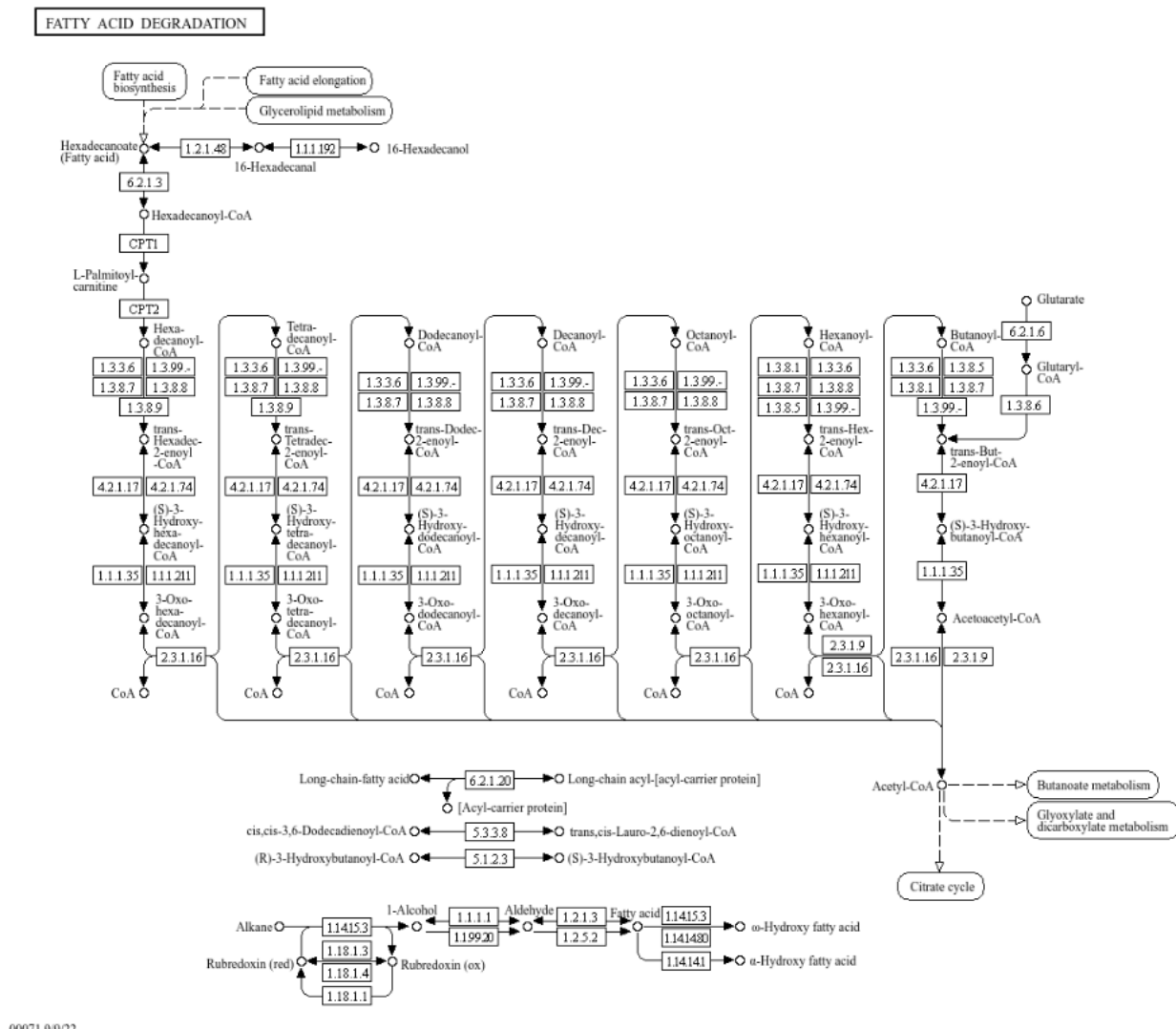


Fatty Acid Degradation

نقش: این مسیر مسئول تجزیه اسیدهای چرب و تولید انرژی است.

ارتباط با سرطان سینه: سلول‌های سرطانی می‌توانند برای تأمین انرژی خود به این مسیر وابسته باشند. تغییر در متابولیسم اسیدهای چرب معمولاً در سلول‌های سرطانی مشاهده می‌شود.

Entry	map00071	Pathway
Name	Fatty acid degradation	
Class	Metabolism; Lipid metabolism	
	BRITE hierarchy	
Pathway map	map00071 Fatty acid degradation	
	<p>98071: 6/10/22 103: Kanaoka Laboratories</p>	
	Ortholog table	
Module	M00086 beta-Oxidation, acyl-CoA synthesis [PATH:map00071] M00087 beta-Oxidation [PATH:map00071]	
Other DBs	COI: 0000000	



PI3K-Akt Signaling Pathway

نقش: این مسیر در تنظیم رشد، بقا و متابولیسم سلولی نقش دارد.

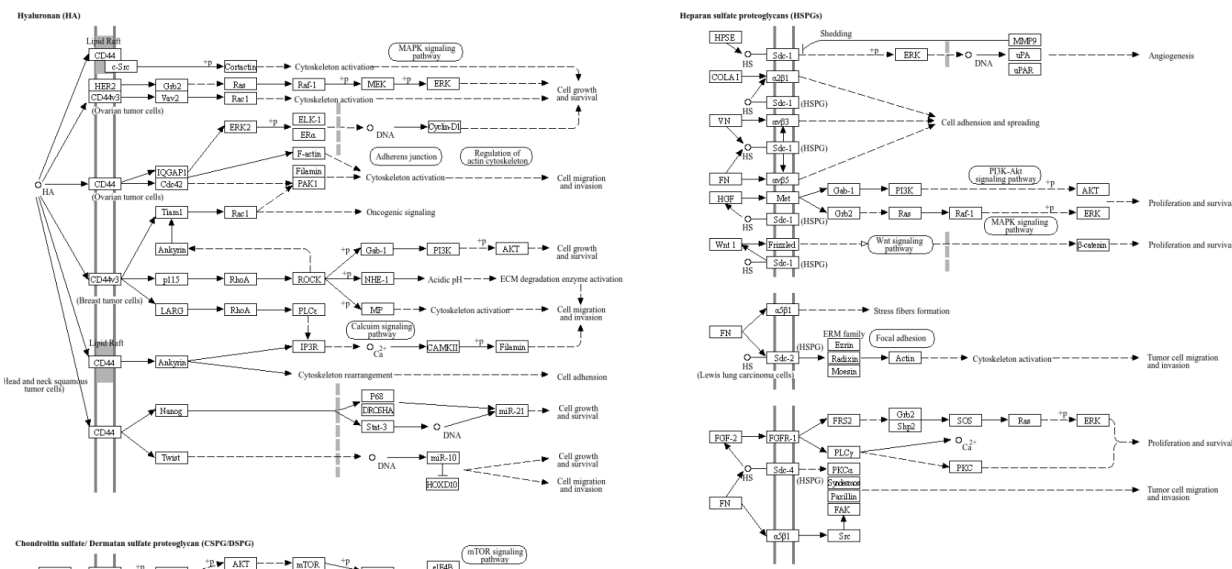
ارتباط با سرطان سینه: PI3K-Akt یک مسیر شناخته شده در سرطان سینه است. جهش‌ها یا فعال‌سازی بیش‌ازحد این مسیر می‌تواند باعث رشد غیرطبیعی سلولی و مقاومت به درمان شود.

نقش: پروتئوگلیکان‌ها در ماتریکس خارج سلولی قرار دارند و در فرآیندهای چسبندگی سلولی و انتقال سیگنال نقش دارند.

ارتباط با سرطان سینه: پروتئوگلیکان‌ها می‌توانند از طریق تعامل با گیرنده‌های سطح سلولی، مسیرهای مرتبط با رشد و مهاجرت سلولی را فعال کنند و متاستاز سرطان سینه را تسهیل نمایند.

Entry	map05205	Pathway
Name	Proteoglycans in cancer	
Description	Many proteoglycans (PGs) in the tumor microenvironment have been shown to be key macromolecules that contribute to biology of various types of cancer including proliferation, adhesion, angiogenesis and metastasis, affecting tumor progress. The four main types of proteoglycans include hyaluronan (HA), which does not occur as a PG but in free form, heparan sulfate proteoglycans (HSPGs), chondroitin sulfate proteoglycans (CSPGs), dermatan sulfate proteoglycans (DSPG) and keratan sulfate proteoglycans (KSPGs) [BR:00535]. Among these proteoglycans such as HA, acting with CD44, promotes tumor cell growth and migration, whereas other proteoglycans such as syndecans (-1~4), glypican (-1, -3) and perlecan may interact with growth factors, cytokines, morphogens and enzymes through HS chains [BR:00536], also leading to tumor growth and invasion. In contrast, some of the small leucine-rich proteoglycans, such as decorin and lumican, can function as tumor repressors, and modulate the signaling pathways by the interaction of their core proteins and multiple receptors.	

PROTEOGLYCANS IN CANCER

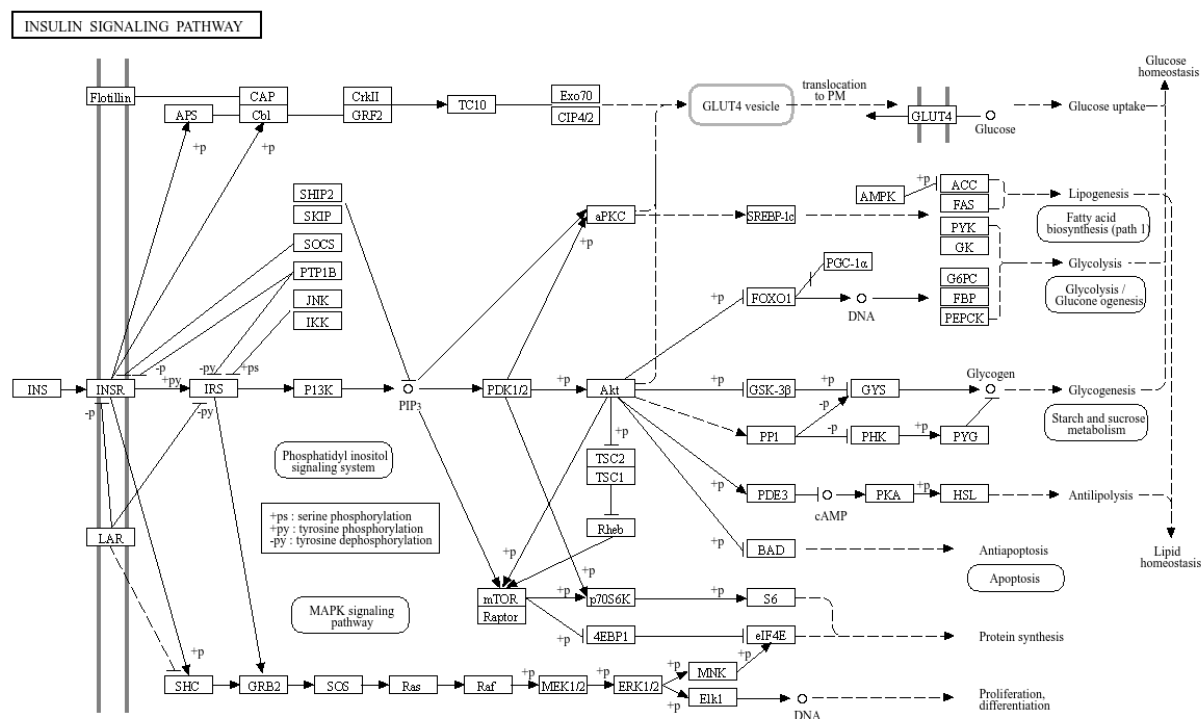


Insulin Signaling Pathway

نقش: این مسیر متابولیسم گلوکز و لیپیدها را تنظیم می‌کند و نقش مهمی در رشد سلولی دارد.

ارتباط با سرطان سینه: مقاومت به انسولین و افزایش سطح انسولین می‌تواند مسیرهای رشد سلولی مرتبط با سرطان سینه را فعال کند. این موضوع به‌ویژه در افراد دارای اضافه‌وزن یا چاقی اهمیت دارد.

Entry	map04910	Pathway
Name	Insulin signaling pathway	
Description	<p>Insulin binding to its receptor results in the tyrosine phosphorylation of insulin receptor substrates (IRS) by the insulin receptor tyrosine kinase (INSR). This allows association of IRSs with the regulatory subunit of phosphoinositide 3-kinase (PI3K). PI3K activates 3-phosphoinositide-dependent protein kinase 1 (PDK1), which activates Akt, a serine kinase. Akt in turn deactivates glycogen synthase kinase 3 (GSK-3), leading to activation of glycogen synthase (GYS) and thus glycogen synthesis. Activation of Akt also results in the translocation of GLUT4 vesicles from their intracellular pool to the plasma membrane, where they allow uptake of glucose into the cell. Akt also leads to mTOR-mediated activation of protein synthesis by eIF4 and p70S6K. The translocation of GLUT4 protein is also elicited through the CAP/Cbl/TC10 pathway, once Cbl is phosphorylated by INSR.</p> <p>Other signal transduction proteins interact with IRS including GRB2. GRB2 is part of the cascade including SOS, RAS, RAF and MEK that leads to activation of mitogen-activated protein kinase (MAPK) and mitogenic responses in the form of gene transcription. SHC is another substrate of INSR. When tyrosine phosphorylated, SHC associates with GRB2 and can thus activate the RAS/MAPK pathway independently of IRS-1.</p>	



Malaria Pathway

نقش: این مسیر به طور خاص به واکنش‌های ایمنی بدن در برابر انگل Plasmodium اشاره دارد که عامل بیماری مالاریا است. این مسیر شامل تنظیمات ایمنی، تولید سیتوکین‌ها و عوامل التهابی است.

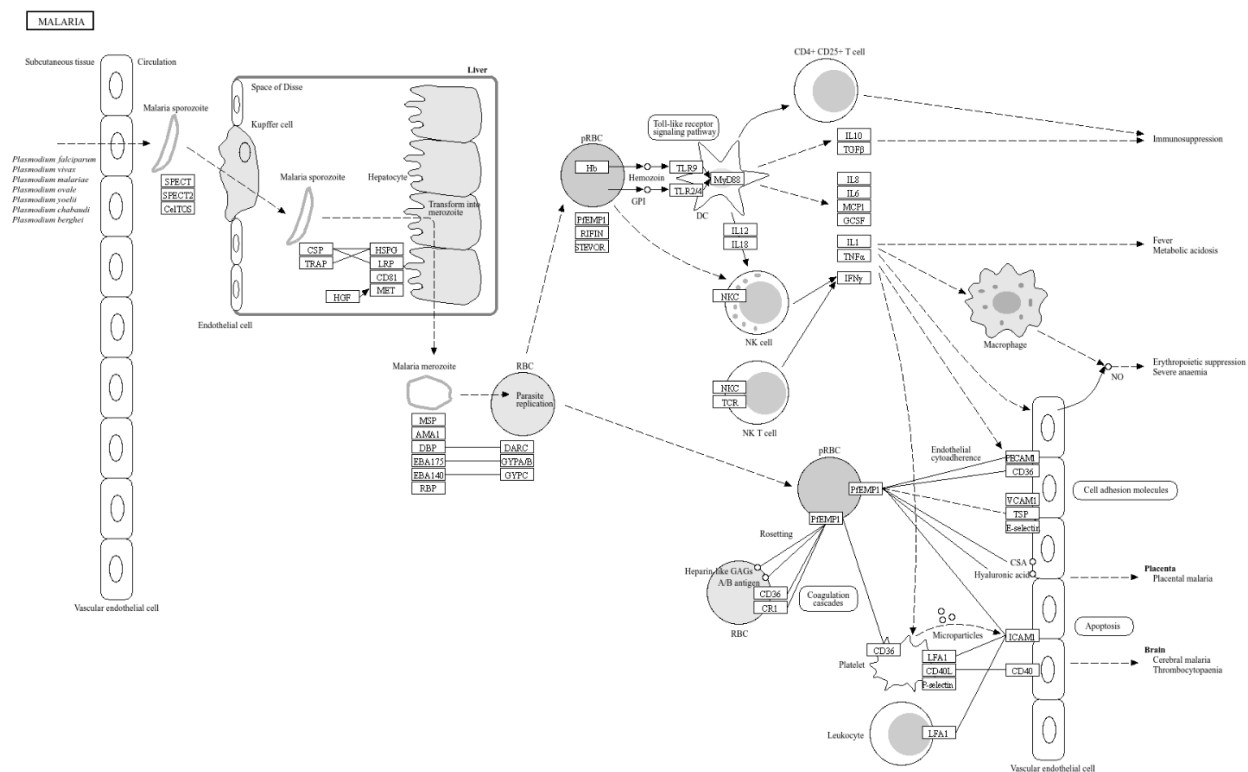
التهاب مزمن: مالاریا با ایجاد التهاب سیستمیک، ریسک شکل‌گیری سرطان، از جمله سرطان سینه، را افزایش می‌دهد.

سرکوب ایمنی: ضعف سیستم ایمنی ناشی از عفونت‌های طولانی‌مدت مالاریا به سلول‌های سرطانی کمک می‌کند از شناسایی فرار کنند.

استرس اکسیداتیو: افزایش استرس اکسیداتیو در مالاریا باعث آسیب به DNA و جهش‌های مرتبط با سرطان می‌شود.

ر

Entry	map05144	Pathway
Name	Malaria	
Description	Plasmodium protozoa are parasites that account for malaria infection. Sporozoite forms of the parasite are injected by mosquito bites under the skin and are carried to the liver where they develop into the merozoite form. Sporozoite invasion of hepatocytes is mediated by parasite surface protein like CSP. Subsequent infection into red blood cells (RBCs) by merozoites causes malaria disease via aberrant cytokine production and sequestration of parasite-infected red blood cells (pRBCs) to host endothelium. Microvasculature sequestration in the brain brings about cerebral malaria that can results in death or persisting neurological impairment. PfEMP1 has been suggested as the key adhesive molecule of pRBCs.	
Class	Human Diseases; Infectious disease: parasitic BRITE hierarchy	



۲. تحلیل عملکرد ژنها در GO :

دیتاست GO را بررسی کنید و ژنهای معنادار را در سه دسته زیر تحلیل کنید. نتایج را با استفاده از نمودار و جداول مناسب نشان دهید.

- فرایندهای زیستی (Process Biological): بررسی کنید کدام فرایندهای زیستی معنادار هستند و نقش آنها در سرطان را توضیح دهید.

GO Biological Process 2023

Bar Graph

Table

Clustergram

Appyter



Click the bars to sort. Now sorted by **p-value ranking**.

SVG PNG JPG

Regulation Of Cold-Induced Thermogenesis (GO:0120161)

Hydrogen Peroxide Catabolic Process (GO:0042744)

Positive Regulation Of miRNA Transcription (GO:1902895)

Positive Regulation Of Cold-Induced Thermogenesis (GO:0120162)

Positive Regulation Of miRNA Metabolic Process (GO:2000630)

Positive Regulation Of Metabolic Process (GO:0009893)

Positive Regulation Of Cellular Process (GO:0048522)

Regulation Of miRNA Transcription (GO:1902893)

Positive Regulation Of Multicellular Organismal Process (GO:0051240)

Positive Regulation Of Macromolecule Metabolic Process (GO:0010604)

GO Biological Process 2023

Bar Graph

Table

Clustergram

Appyter



Hover each row to see the overlapping genes.

100

entries per page

Search:

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	Regulation Of Cold-Induced Thermogenesis (GO:0120161)	3.267e-11	7.073e-8	9.63	232.54
2	Hydrogen Peroxide Catabolic Process (GO:0042744)	7.332e-9	0.000005686	38.16	714.80
3	Positive Regulation Of miRNA Transcription (GO:1902895)	7.879e-9	0.000005686	18.88	352.20
4	Positive Regulation Of Cold-Induced Thermogenesis (GO:0120162)	1.520e-8	0.000008225	10.03	180.61
5	Positive Regulation Of miRNA Metabolic Process (GO:2000630)	2.664e-8	0.00001153	16.04	279.74
6	Positive Regulation Of Metabolic Process (GO:0009893)	4.597e-8	0.00001659	8.98	151.77
7	Positive Regulation Of Cellular Process (GO:0048522)	1.233e-7	0.00003814	3.53	56.22
8	Regulation Of miRNA Transcription (GO:1902893)	1.438e-7	0.00003893	12.82	202.05
9	Positive Regulation Of Multicellular Organismal Process (GO:0051240)	1.889e-7	0.00004545	4.21	65.11
10	Positive Regulation Of Macromolecule Metabolic Process (GO:0010604)	3.057e-7	0.00006618	4.25	63.75
11	Positive Regulation Of Cell Population Proliferation (GO:0008284)	4.952e-7	0.00009747	3.67	53.35
12	Positive Regulation Of Intracellular Signal Transduction (GO:1902533)	5.773e-7	0.00009761	3.53	50.66
13	Regulation Of Cell Population Proliferation (GO:0042127)	5.861e-7	0.00009761	3.03	43.53
14	Regulation Of Glucose Import (GO:0046324)	6.495e-7	0.0001004	17.09	243.52
15	Positive Regulation Of Protein Kinase B Signaling (GO:0051897)	8.408e-7	0.0001214	8.60	120.32
16	Regulation Of Angiogenesis (GO:0045765)	0.000001537	0.0002080	5.28	70.69
17	Positive Regulation Of Gene Expression (GO:0010628)	0.000001661	0.0002116	3.52	46.81

در اینجا ۶ فرایند اول را به عنوان significant در نظر میگیریم

1	Regulation Of Cold-Induced Thermogenesis (GO:0120161)	3.267e-11	7.073e-8	9.63	232.54
2	Hydrogen Peroxide Catabolic Process (GO:0042744)	7.332e-9	0.000005686	38.16	714.80
3	Positive Regulation Of miRNA Transcription (GO:1902895)	7.879e-9	0.000005686	18.88	352.20

4	Positive Regulation Of Cold-Induced Thermogenesis (GO:0120162)	1.520 e-8	0.00000 8225	10. 03	180 .61
5	Positive Regulation Of miRNA Metabolic Process (GO:2000630)	2.664 e-8	0.00001 153	16. 04	279 .74
6	Positive Regulation Of Metabolic Process (GO:0009893)	4.597 e-8	0.00001 659	8.9 8	151 .77

Regulation Of Cold-Induced Thermogenesis

این فرآیند مربوط به تولید گرما در واکنش به سرما از طریق فعالیت میتوکندری ها است. در سرطان سینه، تغییرات متابولیسم انرژی و فعالیت میتوکندریایی ممکن است بقای سلول های سرطانی را در شرایط استرسزا، مانند کمبود اکسیژن یا مواد مغذی، تسهیل کند.

Hydrogen Peroxide Catabolic Process

این فرآیند شامل تجزیه هیدروژن پراکسید (H_2O_2)، یکی از گونه های فعال اکسیژن (ROS)، به ترکیبات بی ضرر است. ROS می تواند نقش تنظیمی داشته باشد، اما تولید بیش از حد آن باعث آسیب DNA و جهش های سرطانی می شود. در سرطان سینه، سلول های سرطانی این مسیر را فعال نگه می دارند تا تعادل اکسیداتیو خود را حفظ کرده و از مرگ سلولی جلوگیری کنند.

Positive Regulation Of miRNA Transcription

این فرآیند به افزایش بیان miRNA ها می پردازد که نقش های تنظیمی مهمی در بیان ژن ها دارند. در سرطان سینه، تغییرات در تنظیم miRNA ها می تواند به عدم تعادل در بیان ژن های مرتبط با رشد سلولی، متاستاز و بقا منجر شود و نقش کلیدی در پیشرفت بیماری داشته باشد.

Positive Regulation Of Cold-Induced Thermogenesis

این مسیر مشابه مورد اول است و به تحریک ترموژنز ناشی از سرما مرتبط است. در سرطان سینه، تنظیم انرژی و متابولیسم سلولی که از ویژگی های ترموژنز است، می تواند به تطابق سلول های سرطانی با محیط های کم انرژی کمک کرده و بقا و رشد آنها را تسریع کند.

Positive Regulation Of miRNA Metabolic Process

این فرآیند به افزایش متابولیسم miRNAها، شامل پردازش و فعالیت آنها، می‌پردازد. در سرطان سینه، تغییرات در متابولیسم miRNAها می‌تواند بر بیان ژن‌های دخیل در تکثیر، مهاجرت و متاستاز سلولی اثر گذاشته و به پیشرفت سرطان کمک کند.

Positive Regulation Of Metabolic Process

این فرآیند شامل تحریک متابولیسم سلولی است که در سرطان‌ها به‌طور قابل‌توجهی بازبرنامه‌ریزی می‌شود. در سرطان سینه، سلول‌های سرطانی از متابولیسم گلیکولیتیک (اثر واربورگ) برای تولید سریع انرژی و زیست‌مولکول‌های مورد نیاز برای تکثیر و رشد استفاده می‌کنند.

GO Biological Process 2023

Bar Graph

Table

Clustergram

Appyter



Enriched Terms are the columns, input genes are the rows, and cells in the matrix indicate if a gene is associated with a term.



Row Order

Cluster

Sum

Column Order

Cluster

Sum

Gene

Search

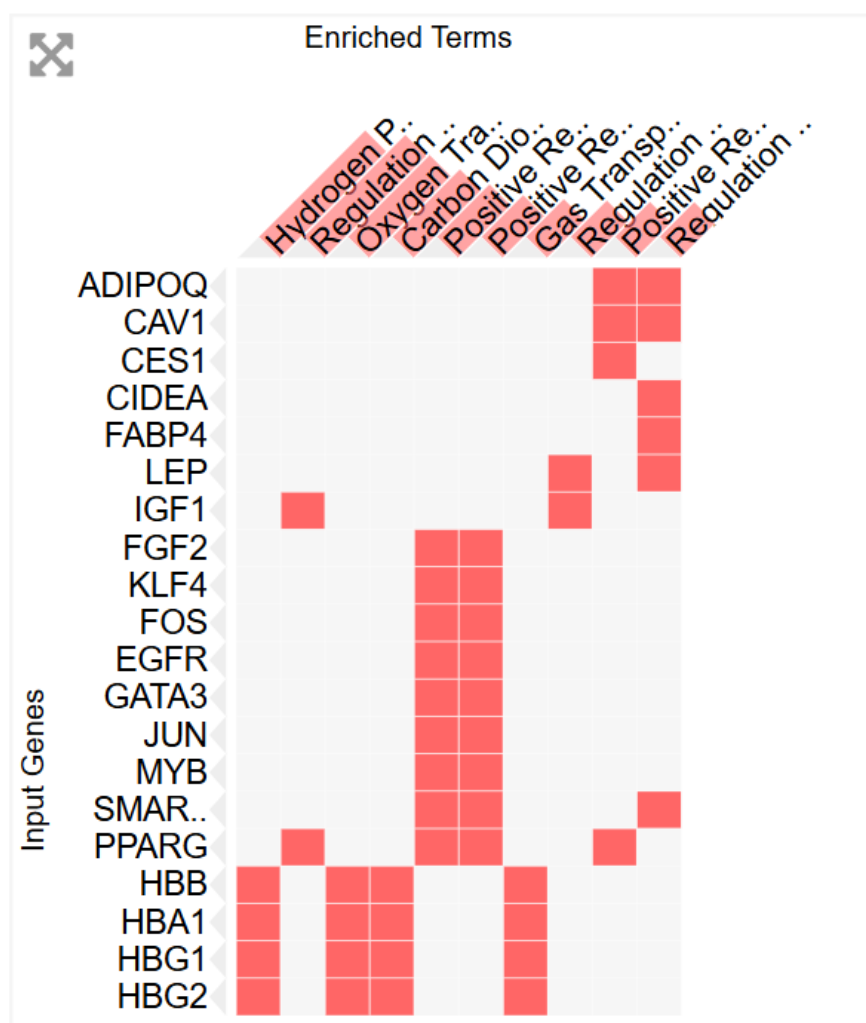
Combined Score

P-Value

Z-score

Top Enriched Terms: 10

Top rows sum: 20 rows



ژن‌های مربوط به هر کدام از این فرایندهای زیستی هم در clustergram بالا مشخص است.

- اجزای سلولی (Component Cellular): مشخص کنید ژنهای معنادار در کدام بخش های سلول فعال هستند و اهمیت این بخش ها را توضیح دهید.

GO Cellular Component 2023

Bar Graph

Table

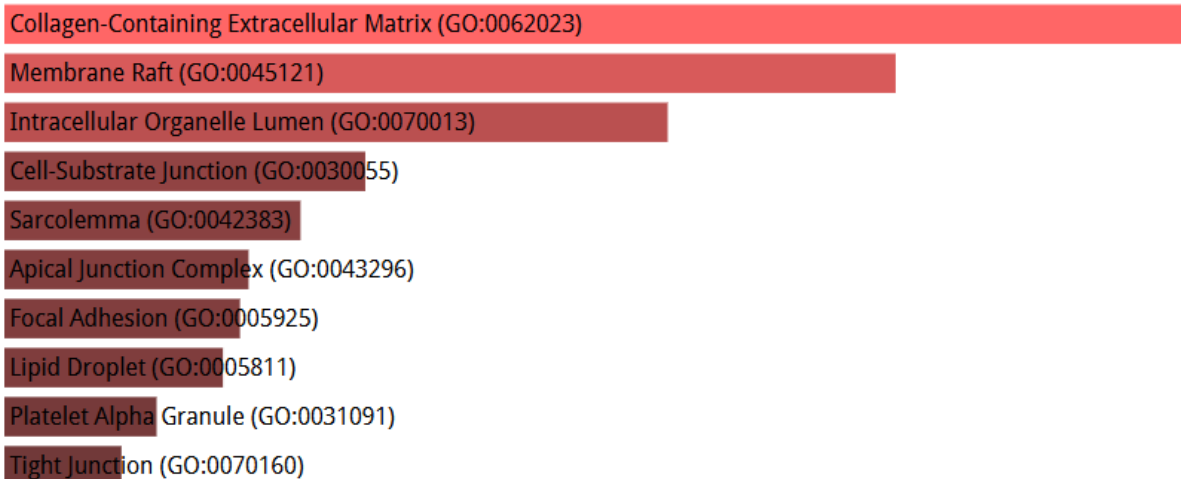
Clustergram

Appyter



Click the bars to sort. Now sorted by p-value ranking.

SVG PNG JPG



GO Cellular Component 2023

Bar Graph

Table

Clustergram

Appyter



Hover each row to see the overlapping genes.

100 entries per page

Search:

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	Collagen-Containing Extracellular Matrix (GO:0062023)	1.020e-7	0.00002020	4.38	70.44
2	Membrane Raft (GO:0045121)	9.624e-7	0.00009528	5.99	83.02
3	Intracellular Organelle Lumen (GO:0070013)	0.000005552	0.0003665	2.69	32.56
4	Cell-Substrate Junction (GO:0030055)	0.00005668	0.002806	3.24	31.73
5	Sarcolemma (GO:0042383)	0.00009327	0.003693	9.20	85.33
6	Apical Junction Complex (GO:0043296)	0.0001391	0.004189	5.78	51.34
7	Focal Adhesion (GO:0005925)	0.0001486	0.004189	3.10	27.34
8	Lipid Droplet (GO:0005811)	0.0001693	0.004189	6.59	57.26
9	Platelet Alpha Granule (GO:0031091)	0.0002819	0.006202	6.03	49.28
10	Tight Junction (GO:0070160)	0.0003693	0.006885	5.75	45.42
11	Cell-Cell Junction (GO:0005911)	0.0003825	0.006885	3.25	25.55
12	Vesicle (GO:0031982)	0.0005800	0.009570	3.49	26.03
13	Bicellular Tight Junction (GO:0005923)	0.001054	0.01605	5.63	38.61

در اینجا هم در صورتی که **p-value** های کمتر از ۰,۰۰۱ را در نظر بگیریم ۱۲ بخش سلولی فعال سلولی پیدا میشود. به عنوان نمونه سه مورد از این بخش های سلولی را توضیح میدهیم:

تحلیل اجزای سلولی (**Cellular Component**) به شناسایی بخش‌هایی از سلول می‌پردازد که ژن‌های موردنظر در آن فعال هستند. در تصویر ارائه‌شده، داده‌ها از طریق آنالیز **GO (Gene Ontology)** مشخص کرده‌اند که برخی از اجزای سلولی به طور معناداری با ژن‌های شما مرتبط هستند. در ادامه برخی از مهم‌ترین بخش‌ها و اهمیت آن‌ها آورده شده است:

Collagen-Containing Extracellular Matrix

ماتریکس خارج سلولی حاوی کلاژن یکی از اجزای مهم محیط خارج سلولی است که نقش اصلی آن فراهم کردن حمایت ساختاری برای سلول‌ها و تنظیم سیگنال‌دهی سلولی است. این بخش در فرایندهای حیاتی مانند بازسازی بافت، حفظ پایداری سلول‌ها، و مهاجرت سلولی نقش اساسی دارد. تغییرات در این ساختار معمولاً به بیماری‌هایی نظیر سرطان و فیبروز مرتبط است و مطالعه آن می‌تواند به درک عمیق‌تر بیماری‌های بافتی کمک کند.

Membrane Raft

رفت‌های غشایی، میکرو دامنه‌هایی در غشای سلولی هستند که غنی از لیپیدها و پروتئین‌های خاص می‌باشند. این نواحی به دلیل نقششان در انتقال سیگنال‌های سلولی، سازمان‌دهی پروتئین‌های غشایی، و تقسیم سلولی بسیار مهم هستند. رفت‌های غشایی در بسیاری از فرایندهای فیزیولوژیکی و بیماری‌های مرتبط با اختلالات سیگنال‌دهی، از جمله سرطان و بیماری‌های ایمنی، نقشی کلیدی ایفا می‌کنند.

Intracellular Organelle Lumen

حفره اندامک‌های داخل سلولی شامل فضای داخلی اندامک‌هایی مانند میتوکندری، هسته و شبکه اندوپلاسمی است که محل انجام فرایندهای ضروری سلولی نظیر ترجمه، تاخوردگی پروتئین، و متابولیسم انرژی می‌باشد. این فضاها به دلیل نقشی که در عملکرد صحیح سلولی و تنظیم پاسخ‌های استرسی دارند، از اهمیت بالایی برخوردارند. اختلال در عملکرد آن‌ها ممکن است منجر به بیماری‌های متابولیکی و ژنتیکی شود.

Enriched Terms are the columns, input genes are the rows, and cells in the matrix indicate if a gene is associated with a term.



Row Order

Cluster

Sum

Column Order

Cluster

Sum

Gene

Search

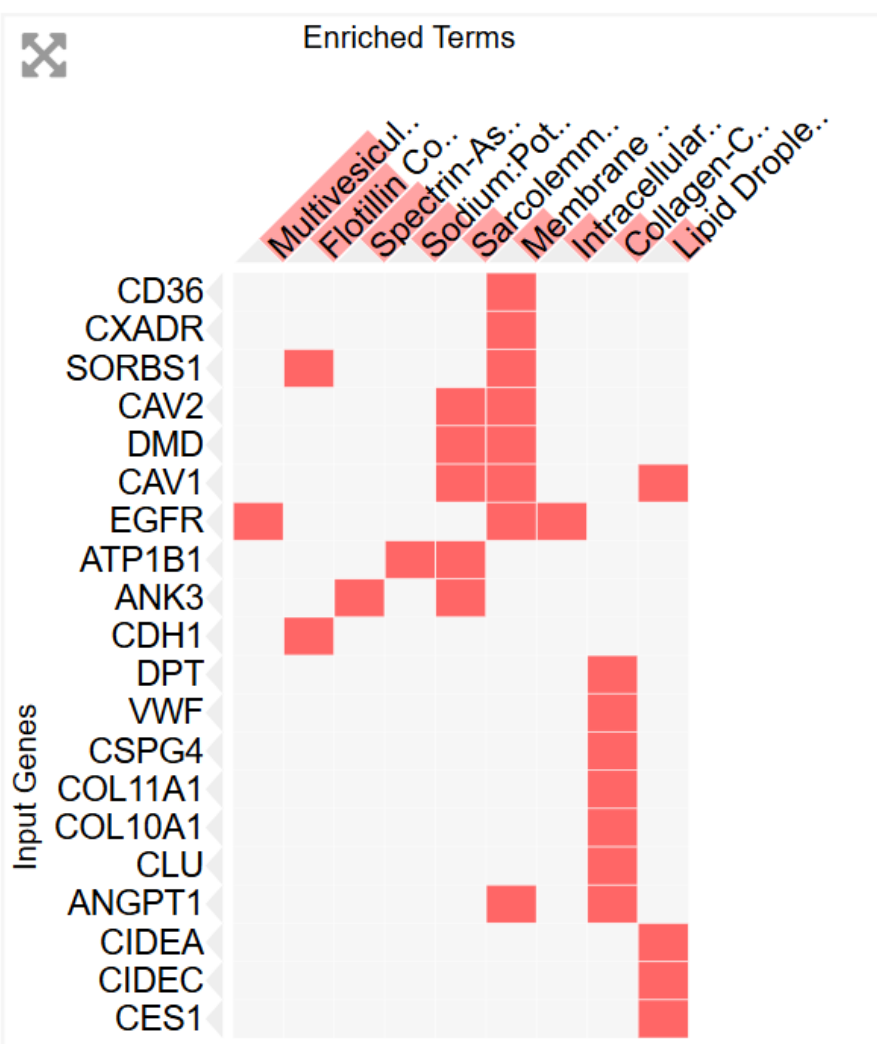
Combined Score

P-Value

Z-score

Top Enriched Terms: 10

Top rows sum: 20 rows



مجدداً ژن‌های مرتبط هم در نمودار بالا مشخص شده‌اند.

- عملکردهای مولکولی (Function Molecular): عملکردهای مولکولی معنادار را

بررسی کرده و ارتباط آن‌ها با فرایندهای سرطان را توضیح دهید.

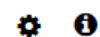
GO Molecular Function 2023

Bar Graph

Table

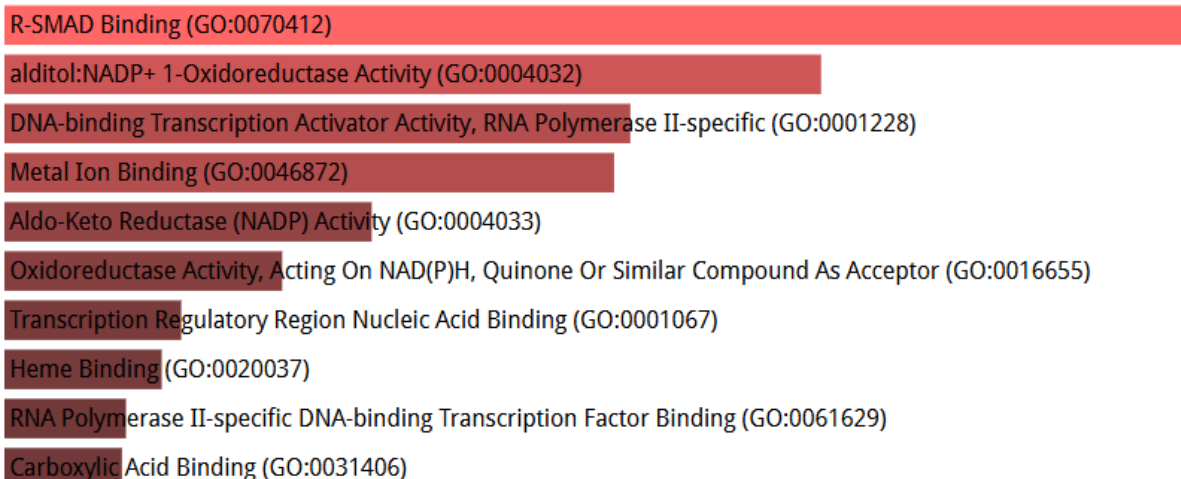
Clustergram

Appyter



Click the bars to sort. Now sorted by **p-value ranking**.

SVG PNG JPG



GO Molecular Function 2023

Bar Graph

Table

Clustergram

Appyter



Hover each row to see the overlapping genes.

10 entries per page

Search:

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	R-SMAD Binding (GO:0070412)	0.0001322	0.05073	18.70	166.97
2	alditol:NADP+ 1-Oxidoreductase Activity (GO:0004032)	0.0003190	0.05073	29.95	241.12
3	DNA-binding Transcription Activator Activity, RNA Polymerase II-specific (GO:0001228)	0.0005019	0.05073	3.00	22.77
4	Metal Ion Binding (GO:0046872)	0.0005216	0.05073	2.57	19.46
5	Aldo-Keto Reductase (NADP) Activity (GO:0004033)	0.0009277	0.06531	19.06	133.07
6	Oxidoreductase Activity, Acting On NAD(P)H, Quinone Or Similar Compound As Acceptor (GO:0016655)	0.001147	0.06531	17.47	118.26
7	Transcription Regulatory Region Nucleic Acid Binding (GO:0001067)	0.001458	0.06531	3.31	21.64
8	Heme Binding (GO:0020037)	0.001528	0.06531	5.21	33.80
9	RNA Polymerase II-specific DNA-binding Transcription Factor Binding (GO:0061629)	0.001663	0.06531	3.25	20.81
10	Carboxylic Acid Binding (GO:0031406)	0.001679	0.06531	14.97	95.65

Showing 1 to 10 of 389 entries | [Export entries to table](#)

[Previous](#) [Next](#)

Terms marked with an * have an overlap of less than 5

اگر p-value های کمتر از ۰,۰۰۱ را در نظر بگیریم ۵ molecular function معنادار پیدا میشود.

در این جدول، عملکردهای مولکولی معنادار (GO Molecular Functions) شناسایی شده‌اند و ارتباط هر عملکرد با فرایندهای بیولوژیکی، از جمله سرطان‌زایی، قابل بررسی است. در ادامه هر عملکرد به همراه ارتباط آن با سرطان توضیح داده شده است:

R-SMAD Binding

این عملکرد شامل اتصال به پروتئین‌های SMAD است که بخشی از مسیر سیگنال‌دهی $TGF-\beta$ محسوب می‌شوند. مسیر $TGF-\beta$ در تنظیم رشد سلول و تمایز نقش دارد و تغییرات آن به رشد کنترل‌نشده سلول‌ها در سرطان منجر می‌شود. افزایش یا کاهش اتصال R-SMAD می‌تواند در فرایندهای متاستاز و سرکوب ایمنی در سرطان تأثیرگذار باشد.

alditol:NADP+ 1-Oxidoreductase Activity

این عملکرد مرتبط با مسیرهای متابولیکی است که در تبدیل قندها نقش دارند. تغییرات متابولیکی یکی از ویژگی‌های اصلی سلول‌های سرطانی است که انرژی موردنیاز برای رشد سریع آن‌ها را تأمین می‌کند. فعالیت غیرطبیعی این آنزیم‌ها می‌تواند در مقاومت سلولی نسبت به درمان‌های متابولیکی دخیل باشد.

DNA-binding Transcription Activator Activity, RNA Polymerase II-specific

این عملکرد شامل تنظیم بیان ژن‌ها از طریق RNA Polymerase II است. در سرطان، این فرایند ممکن است منجر به افزایش بیان ژن‌های مرتبط با بقا، تکثیر، و مقاومت به دارو شود. تنظیم نابجای عوامل رونویسی نقش مهمی در سرطان‌زایی دارد.

Metal Ion Binding

اتصال به یون‌های فلزی در تنظیم فعالیت‌های آنزیمی و انتقال الکترون‌ها ضروری است. برخی یون‌ها مانند آهن و مس در سرطان می‌توانند نقش دوگانه‌ای داشته باشند: تسهیل استرس اکسیداتیو که موجب تخریب سلول‌های سالم می‌شود، یا حمایت از رشد سلول‌های سرطانی.

Aldo-Keto Reductase (NADP) Activity

این آنزیم‌ها در متابولیسم لیپیدها و کربوهیدرات‌ها نقش دارند. در سرطان، آن‌ها می‌توانند با تنظیم مقاومت سلول به استرس و حمایت از رشد تومور در محیط‌های پرتنش نقش داشته باشند. افزایش فعالیت این آنزیم‌ها ممکن است مقاومت به شیمی‌درمانی را تقویت کند.

GO Molecular Function 2023

Bar Graph

Table

Clustergram

Appyter



Enriched Terms are the columns, input genes are the rows, and cells in the matrix indicate if a gene is associated with a term.



Row Order

Cluster

Sum

Column Order

Cluster

Sum

Gene

Search

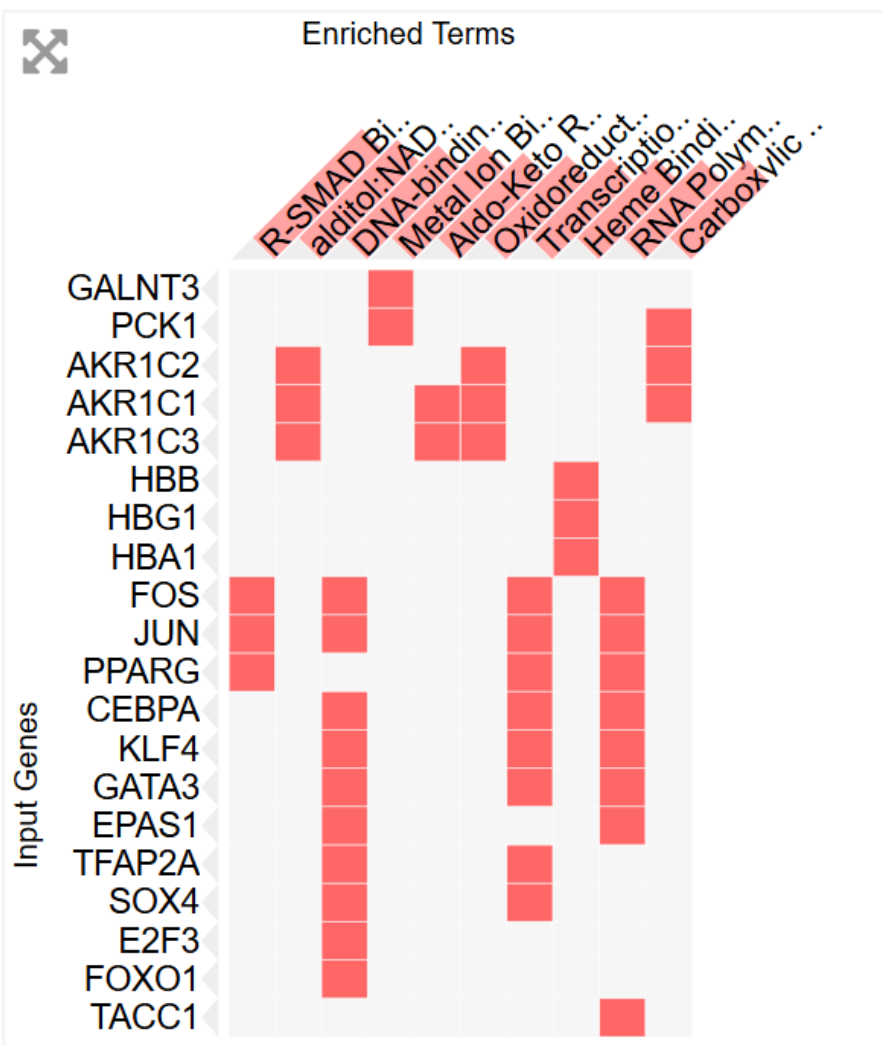
Combined Score

P-Value

Z-score

Top Enriched Terms: 10

Top rows sum: 20 rows



در انتها نیز ژن‌های مرتبط با این عملکردهای مولکولی مشخص شده‌اند.