

## تمرین سوم درس مبانی بیوانفورماتیک

### بخش اول

مشکل استفاده از  $p$ -value :

#### مسئله‌ی آزمون‌های متعدد Problem Testing Multiple

توی بیان ژن، احتمال وقوع نتایج مثبت کاذب Positives False به طور قابل توجهی افزایش پیدا میکنه. مثلاً، اگه سطح معناداری  $\alpha$  را ۰.۰۵ در نظر بگیریم، به طور متوسط ۵٪ از ژن‌ها فقط به خاطر شانس به عنوان تفاضلی شناسایی میشن، حتی اگه تفاوت واقعی وجود نداشته باشه.

#### عدم بیان شدت تغییرات Size Effect

$p$ -value فقط نشون میده که آیا تفاوت بین شرایط از نظر آماری معنادار هست یا نه، اما:

- اطلاعاتی درباره اندازه یا شدت تغییرات بیان ژن مانند Change Fold ارائه نمیده.
- ممکن هست یک تغییر خیلی کوچک اما معنادار از نظر آماری، در عمل اهمیت بیولوژیکی نداشته باشه.

#### وابستگی به حجم نمونه Dependency Size Sample

$p$ -value به شدت به اندازه نمونه Size Sample حساس هست:

- تو نمونه‌های بزرگ، حتی تفاوت‌های کوچک ممکن هست از نظر آماری معنادار شن.
- تو نمونه‌های کوچیک، تفاوت‌های واقعی ممکن هست به خاطر ناکافی بودن قدرت آزمون شناسایی نشن.

## طراحی آزمایش

بررسی این که چگونه احتمال رخ دادن خطای نوع اول (رد شدن اشتباه فرض صفر) با افزایش تعداد آزمون‌ها تغییر می‌کند.

## مراحل آزمایش

۱. فرض اولیه: ابتدا داده‌ها را تحت فرض صفر  $H_0$  شبیه‌سازی می‌کنیم. به این معنی که هیچ تفاوت واقعی بین گروه‌ها وجود ندارد و هر تفاوتی که مشاهده می‌شود، صرفاً به خاطر شانس هست.

۲. تعریف سطح معناداری: سطح معناداری  $\alpha$  را ۰.۰۵ در نظر می‌گیریم.

۳. تعداد آزمون‌ها: شبیه‌سازی را برای مقادیر مختلف تعداد آزمون‌ها (مانند ۱۰، ۵۰، ۱۰۰، ۱۰۰۰، و غیره) انجام می‌دیم.

۴. اجرای آزمون‌های آماری: برای هر آزمون، فرض صفر رو تست می‌کنیم (مثلاً با آزمون  $t$  یا آزمون‌های مشابه). نتایج معنادار ( $p < \alpha$ ) را ثبت می‌کنیم.

۵. محاسبه نرخ خطای نوع اول: درصد آزمون‌هایی که به اشتباه به عنوان معنادار تشخیص داده شدن (رد اشتباه  $H_0$ ) را محاسبه می‌کنیم.

۶. تکرار آزمایش: برای افزایش دقت، شبیه‌سازی را چندین بار (مثلاً ۱۰۰۰ بار) تکرار می‌کنیم و میانگین نرخ خطای نوع اول رو برای هر تعداد آزمون محاسبه می‌کنیم.

## انتظار نتایج

- برای تعداد آزمون‌های کم (مثلاً ۱۰)، احتمال خطای نوع اول نزدیک به سطح معناداری  $\alpha$  خواهد بود (حدود ۰.۰۵).
- با افزایش تعداد آزمون‌ها، تعداد خطاهای نوع اول هم افزایش پیدا می‌کند، حتی اگر نرخ خطا برای هر آزمون ثابت باشد.
- مثال: اگر ۱۰۰۰ آزمون انجام بدیم و  $\alpha = 0.05$ ، به طور متوسط ۵٪ (یعنی ۵۰) از آزمون‌ها به اشتباه به عنوان معنادار شناسایی میشن.

## روش‌های اصلاح p-value

### روش بونفرونی

روش بونفرونی یکی از ساده‌ترین و سخت‌گیرانه‌ترین روش‌هاست که برای جلوگیری از خطای نوع اول استفاده می‌شود. ایده‌اش اینست که سطح معناداری ( $\alpha$ ) را تقسیم بر تعداد آزمون‌ها ( $m$ ) کنیم تا به سطح جدید برای معناداری به دست بیاریم:

$$\alpha_{\text{اصلاح شده}} = \frac{\alpha}{m}$$

اینجوری هر  $p$ -مقدار را با سطح معناداری اصلاح‌شده مقایسه می‌کنیم. اگر  $p$ -مقدار از این سطح پایین‌تر باشد، اون نتیجه را معنادار در نظر می‌گیریم.

#### خوبیش:

- ساده‌ست و احتمال خطای نوع اول را خیلی پایین میاره.

#### بدیش:

- سخت‌گیرانه‌ست و ممکنه باعث بشه نتایج درست رو هم از دست بدی، مخصوصاً وقتی تعداد آزمون‌ها زیاد باشه.

### روش بنجامینی-هوچبرگ BH

این یکی روش منعطف‌تریه. به جای اینکه احتمال به خطای نوع اول رو سخت بگیره، میاد درصد نتایج اشتباهی که به عنوان مثبت شناسایی شدن رو کنترل می‌کنه (بهش میگن FDR).  $p$ -value‌ها رو از کوچک به بزرگ مرتب می‌کنیم. بعد برای هر کدوم به آستانه حساب می‌کنیم که از فرمول  $\frac{i}{m} \times \alpha$  میاد. حالا بزرگ‌ترین  $p$ -مقداری که از آستانه‌اش کوچیک‌تره، و هر چی پایین‌تر از اون باشه، معنادار می‌شه.

#### خوبیش:

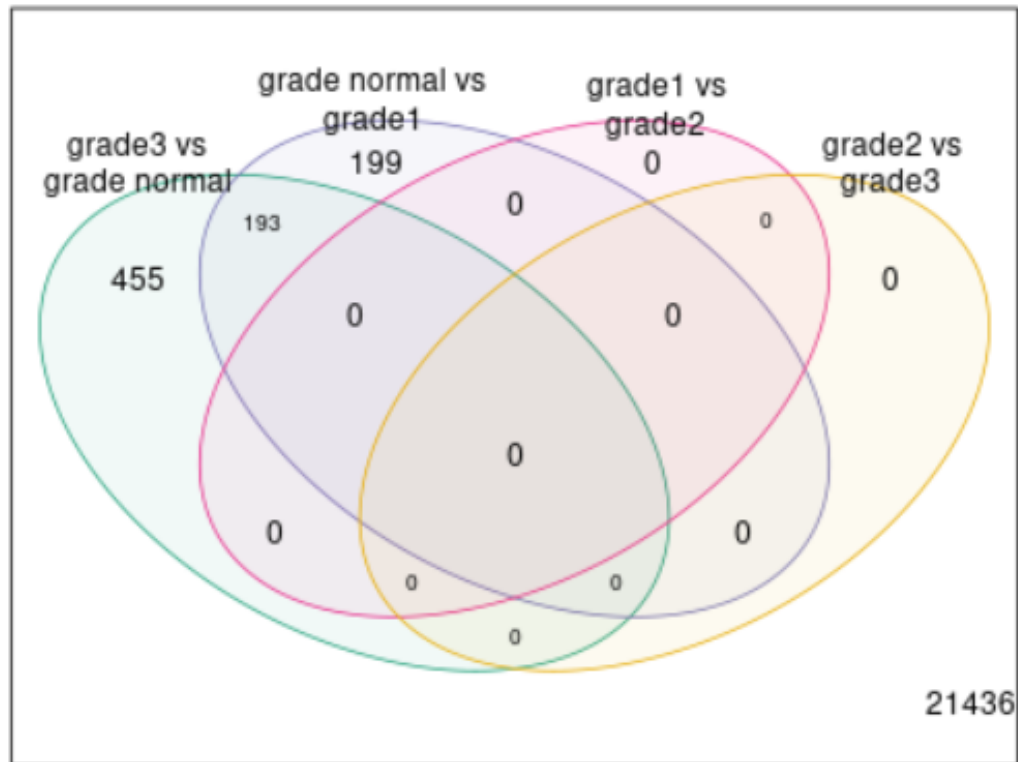
- خیلی بهتره برای وقتی که تعداد آزمون‌ها زیاده، چون حساسیت بهتری داره.

#### بدیش:

- اگر آزمون‌ها خیلی به هم مرتبط باشن، ممکنه دقیق کار نکنه.

## تقسیم بندی براساس گریدها و تحلیل نمودارهاش

**GSE15852: limma, Padj<0.05**



\*تحلیل نمودار ون

این نمودار ون نشون می‌ده که کدوم ژن‌ها توی مقایسه‌های مختلف بین گروه‌های (نرمال، درجه ۱، درجه ۲، درجه ۳) بیان متفاوتی دارن و معنادار هستن (با معیار  $P_{adj} < 0.05$ ). چهار گروه داریم:

• Normal Grade: بافت نرمال.

• Grade1: تومورهای درجه ۱.

• Grade2: تومورهای درجه ۲.

• Grade3: تومورهای درجه ۳.

مقایسه‌های زیر انجام شده:

۱. Grade1 vs Normal: نرمال در مقابل تومورهای درجه ۱.

۲. Grade3 vs Normal: نرمال در مقابل تومورهای درجه ۳.

۳. Grade2 vs Grade1: تومورهای درجه ۱ در مقابل درجه ۲.

۴. Grade3 vs Grade2: تومورهای درجه ۲ در مقابل درجه ۳.

هر بیضی توی نمودار تعداد ژن‌های معنادار (بر اساس  $P_{adj} < 0.05$ ) رو برای یک مقایسه خاص نشون می‌ده.

- نرمال و درجه ۱:

۱۹۹ تا ژن معنادار داریم. یعنی این ژن‌ها بین این دو گروه بیان متفاوتی دارن.

- نرمال و درجه ۳:

۴۵۵ تا ژن معنادار پیدا شده. تعدادش بیشتر از مقایسه با درجه ۱ هست، که نشون می‌ده تفاوت بین نرمال و تومورهای پیشرفته‌تر Grade3 بیشتره.

- درجه ۱ و درجه ۲:

اینجا هیچ ژنی معنادار نیست (۰ تا). یعنی بیان ژن‌ها بین این دو گروه خیلی شبیه به هم بوده.

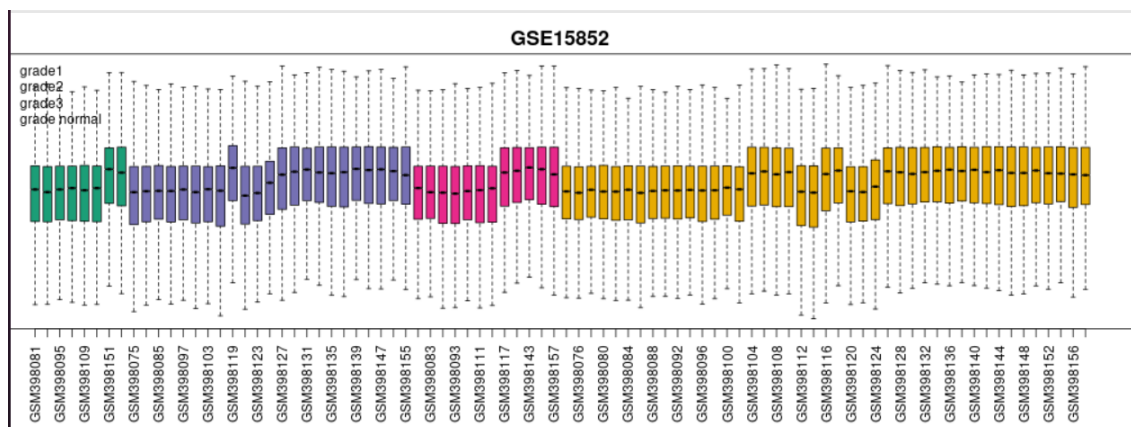
- درجه ۲ و درجه ۳:

اینجا هم ۰ تا ژن معنادار داریم. یعنی تومورهای درجه ۲ و ۳ از نظر بیان ژن شبیه به هم هستن.

- اشتراک بین نرمال، درجه ۱ و درجه ۳ (۱۹۳ ژن): این ژن‌ها بین نرمال و هر دو گروه درجه ۱ و درجه ۳ مشترکاً معنادار هستن. این یعنی بیان این ژن‌ها در همان مراحل اولیه سرطان (درجه ۱) تا مراحل پیشرفته‌تر (درجه ۳) از نرمال متفاوت شده و احتمالاً با فرآیندهای کلی سرطان مرتبط هستن. این ژن‌ها می‌تونن برای تشخیص سرطان در مراحل اولیه و پیشرفته مفید باشن.

- بیان ژن بین بافت نرمال و تومورهای درجه ۱ و ۳ خیلی متفاوته.

- اما بین درجات مختلف سرطان (مثل درجه ۱ و ۲ یا درجه ۲ و ۳) تفاوت معناداری توی بیان ژن پیدا نشده.



## تحلیل نمودار جعبه ای

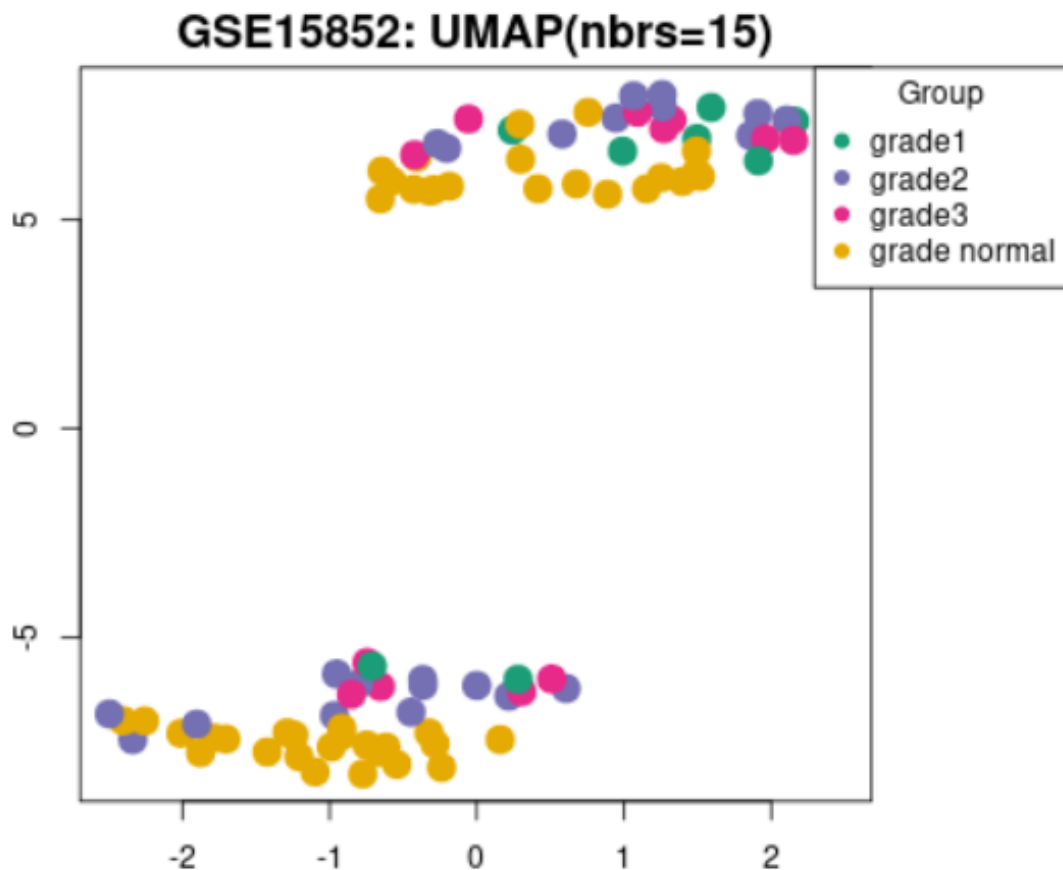
این نمودار پراکندگی داده‌های بیان ژن Values Expression توی هر نمونه رو به صورت جعبه‌ای box نشون می‌ده:

- خط وسط جعبه: مقدار میانه Median بیان ژن برای اون نمونه.
- طول جعبه: فاصله بین چارک اول Q1 و چارک سوم Q3 هست (۵۰ درصد میانی داده‌ها).
- خط‌های بیرونی Whiskers: حدود داده‌های مینیمم و ماکزیمم رو نشون می‌ده (غیر از داده‌های پرت).

۱. نرمال بودن داده‌ها (نرمال سازی): نمودار به‌طور کلی نشون می‌ده که داده‌ها به‌خوبی نرمال سازی شدن. خط میانه (وسط جعبه‌ها) در بیشتر نمونه‌ها نزدیک به هم قرار داره. این یعنی تفاوت‌هایی که بعداً بررسی می‌کنیم، واقعی‌تر هستن و ناشی از نرمال نبودن داده‌ها نیست.

۲. پراکندگی داده‌ها در گروه‌ها: داده‌های گروه نرمال (سمت چپ نمودار، رنگ سبز) کمی متراکم‌تر هستن. در گروه‌های سرطانی (grade1 grade2 grade3) پراکندگی داده‌ها بیشتر شده، که نشون می‌ده بیان ژن توی این گروه‌ها متنوع‌تره.

۳. تفاوت‌های احتمالی بین گروه‌ها: تفاوت‌های کمی بین گروه‌های سرطان (grade1 grade2 grade3) و نرمال توی خط میانه دیده می‌شه. این تفاوت‌ها احتمالاً نشون‌دهنده تغییرات بیان ژن در سرطان هستن که باید با تحلیل‌های آماری مثل تحلیل بیان ژن تفاضلی (Differential Analysis Expression Gene) بررسی بشن.



\*تحلیل نمودار umap این نمودار Projection and Approximation Manifold Uniform UMAP داده‌های بیان ژن رو توی دیتاست GSE15852 نشون می‌ده. این روش کمک می‌کنه داده‌های با ابعاد بالا (مثل داده‌های بیان ژن) رو به دو بعد کاهش بدیم تا خوشه‌بندی و ارتباط گروه‌ها رو بهتر ببینیم.

#### • چهار گروه داریم:

- normal : با رنگ زرد.
- grade1 : با رنگ سبز.
- grade2 : با رنگ بنفش.
- grade3 : با رنگ صورتی.

• **محورها:** محوره‌های  $x$  و  $y$  مقادیر کاهش‌یافته هستن و مفهوم خاصی ندارن، فقط برای نشون دادن موقعیت نسبی نمونه‌ها در فضا هستن.

• **چیدمان نقاط:** هر نقطه توی نمودار، یه نمونه هست و گروه‌ها به صورت خوشه‌های جدا یا نزدیک به هم دیده می‌شن.

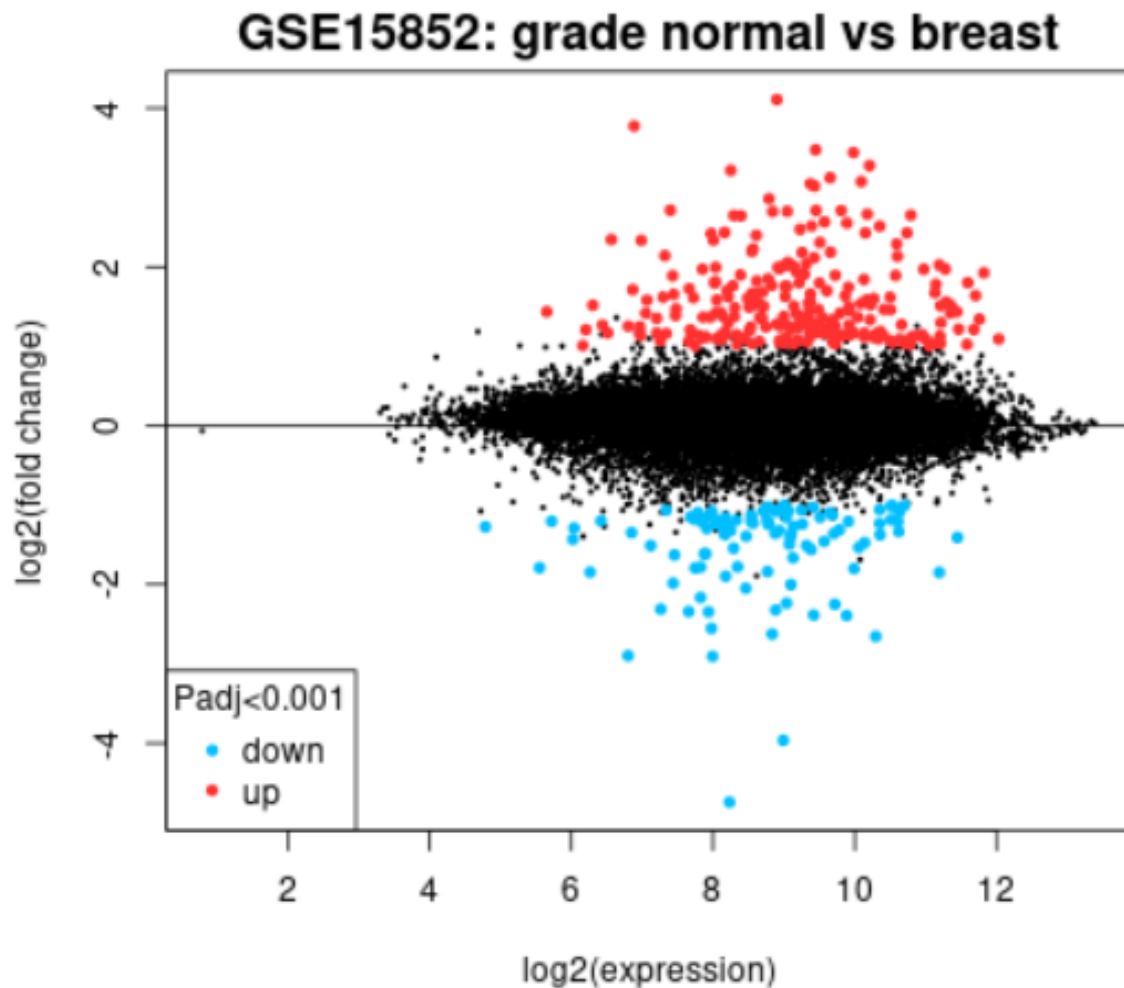
۱. **جدا شدن گروه نرمال (زرد)** نمونه‌های نرمال : normal توی پایین نمودار یه خوشه مشخص و جدا دارن. این نشون می‌ده که بیان ژن توی بافت‌های نرمال با بقیه گروه‌ها فرق زیادی داره.

۲. **تداخل گروه‌های سرطانی:** گروه‌های سرطان grade1 grade2 grade3 در بالای نمودار دیده می‌شن، اما خوشه‌های جداگانه‌ای ندارن و تا حدی با هم تداخل دارن. این یعنی بیان ژن توی درجات مختلف سرطان شباهت زیادی داره.

۳. **شباهت زیاد سرطان‌ها به هم:** نزدیکی گروه‌های grade1 grade2 grade3 نشون می‌ده که بیان ژن توی این سه گروه شباهت زیادی داره.

۴. **تفاوت بین نرمال و سرطان:** تفاوت بیان ژن بین گروه نرمال و گروه‌های سرطانی خیلی مشخصه. این می‌تونه برای پیدا کردن بیومارکرهای مرتبط با سرطان کمک کنه.





این نمودار تغییرات بیان ژن رو با استفاده از میانگین بیان ژن و تفاوت بیان ژن بین دو گروه (نرمال و تومور) بررسی می‌کند.

- محور  $x$ : این محور مقدار میانگین بیان ژن‌ها رو نشون می‌ده. ژن‌هایی که بیانشون در هر دو گروه زیاد باشه، به سمت راست نمودار می‌رن. ژن‌هایی که بیانشون کمتر باشه، به سمت چپ می‌رن.

- محور  $y$ : این محور اختلاف بیان ژن‌ها بین دو گروه رو نشون می‌ده:

$y > 0$ : بیان ژن در گروه تومور بیشتر از نرماله (upregulated).

$y < 0$ : بیان ژن در گروه نرمال بیشتر از توموره (downregulated).

- رنگ‌ها:

(upregulated) ژن‌هایی که بیان‌شون در تومور بیشتره و از نظر آماری معنادار هستن

( $P_{adj} < 0.001$ ).

(downregulated) ژن‌هایی که بیان‌شون در نرمال بیشتره و معنادار هستن

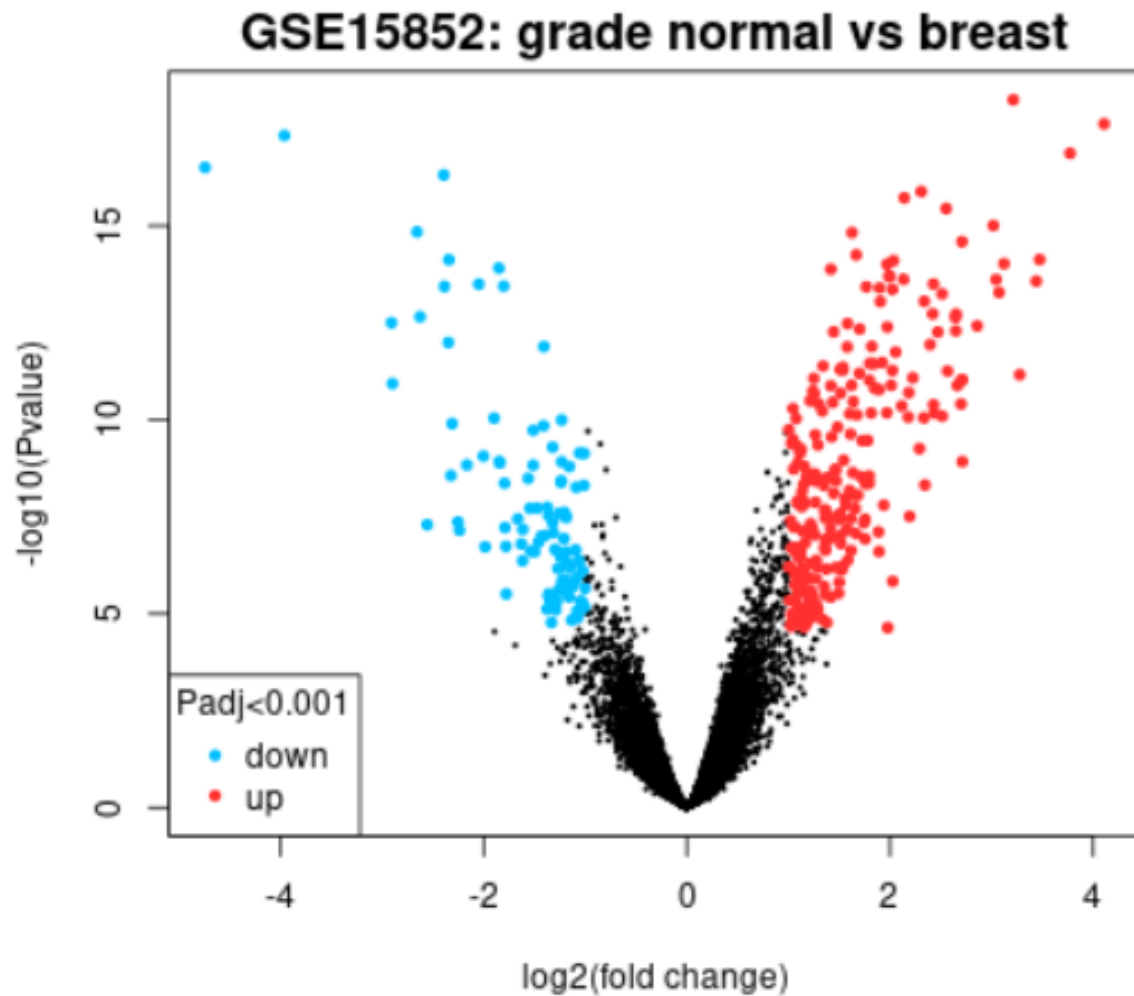
( $P_{adj} < 0.001$ ).

سیاه: ژن‌هایی که تغییر معناداری ندارن.

۱. **ژن‌های معنادار:** ژن‌های قرمز و آبی تغییرات معناداری دارن. قرمزها مرتبط با فرآیندهای سرطانی هستن، چون بیان‌شون در تومور بیشتر شده. آبی‌ها به فرآیندهای طبیعی مربوط می‌شن که توی سرطان سرکوب شدن.

۲. **اکثریت نقاط سیاه:** بیشتر ژن‌ها تغییر معناداری ندارن و حول خط  $y = 0$  متمرکز هستن.

۳. **پراکندگی:** ژن‌هایی که تغییرات بزرگ دارن (قرمز و آبی) بیشتر در انتهای بالا و پایین نمودار دیده می‌شن، که معمولاً مربوط به ژن‌هایی با بیان خیلی زیاد یا خیلی کم هستن.



این نمودار Plot Volcano تغییرات بیان ژن‌ها رو بین گروه بافت نرمال و بافت توموری نشون می‌ده. هدف این نمودار خلاصه کردن میزان Change Fold و معناداری تفاوت بیان ژن آماری (p-value) در یک نمودار هست.

•  $(\log_2(\text{fold change}))$ : این محور میزان تغییر بیان ژن رو نشون می‌ده:

$x > 0$ : ژن‌هایی که توی بافت توموری بیشتر بیان شدن (upregulated).

$x < 0$ : ژن‌هایی که توی بافت نرمال بیشتر بیان شدن (downregulated).

هر چی از  $x = 0$  فاصله بیشتر باشه، تغییر بیان بزرگ‌تره.

•  $(-\log_{10}(\text{p-value}))$ : این محور میزان معناداری آماری رو نشون می‌ده:

هر چی نقطه‌ای بالاتر باشه، معناداری آماری بیشتری داره.

نقاط پایین‌تر، تغییرات معناداری ندارند.

• رنگ‌ها:

قرمز: (upregulated) ژن‌هایی که بیان‌شون در تومور بیشتره و معنادار هستن ( $\text{Padj} < 0.001$  و  $\log_2\text{FC} > 1$ ).

آبی: (downregulated) ژن‌هایی که بیان‌شون در نرمال بیشتره و معنادار هستن ( $\text{Padj} < 0.001$  و  $\log_2\text{FC} < -1$ ).

سیاه: ژن‌هایی که تغییر معناداری ندارند.

ژن‌های قرمز: (upregulated) این ژن‌ها توی بافت توموری بیشتر بیان شدن و احتمالاً با فرآیندهای سرطانی مثل تکثیر سلولی یا متاستاز مرتبط هستن.

ژن‌های آبی: (downregulated) این ژن‌ها توی بافت نرمال بیشتر بیان شدن و معمولاً با عملکردهای طبیعی مثل آپوپتوز یا تنظیم چرخه سلولی مرتبط هستن.

نقاط سیاه: بیشتر ژن‌ها تغییر معناداری ندارند و نزدیک به خط  $x = 0$  هستن.

الگوی V شکل:

این الگو نشون می‌ده که ژن‌های معنادار (upregulated و downregulated) در دو طرف نمودار قرار دارن، جایی که تغییر بیان زیادی دارن.

• تعداد زیادی از ژن‌ها معنادار هستن (قرمز و آبی)، که نشون می‌ده سرطان تأثیر زیادی روی بیان ژن‌ها داره.

• ژن‌های قرمز و آبی برای شناسایی بیومارکرهای سرطان مناسب هستن.

• اکثر ژن‌ها تغییر معناداری ندارند (نقاط سیاه)، که به اهمیت ژن‌های خاص معنادار تأکید می‌کنه.

## بخش دوم تحلیل KEGG 2021 Human

### KEGG 2021 Human

Bar Graph

Table

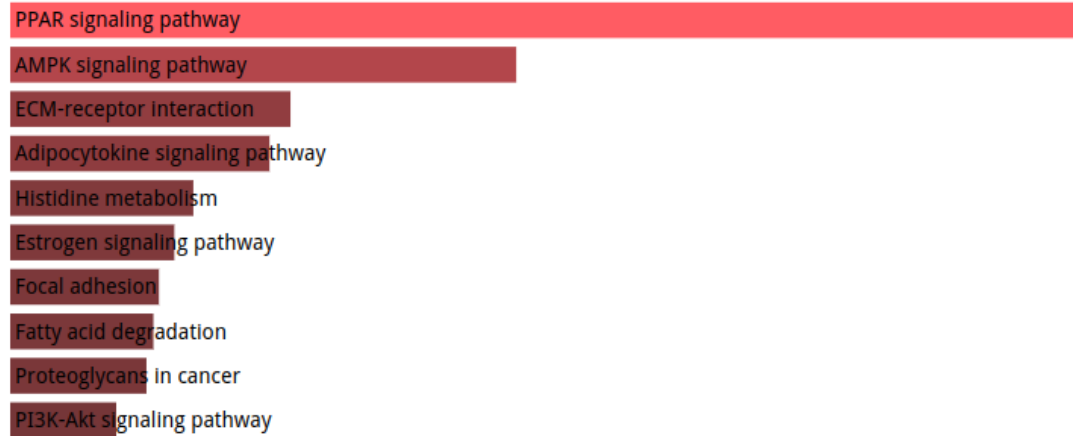
Clustergram

Appyter



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

SVG PNG JPG



:Graph Bar

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

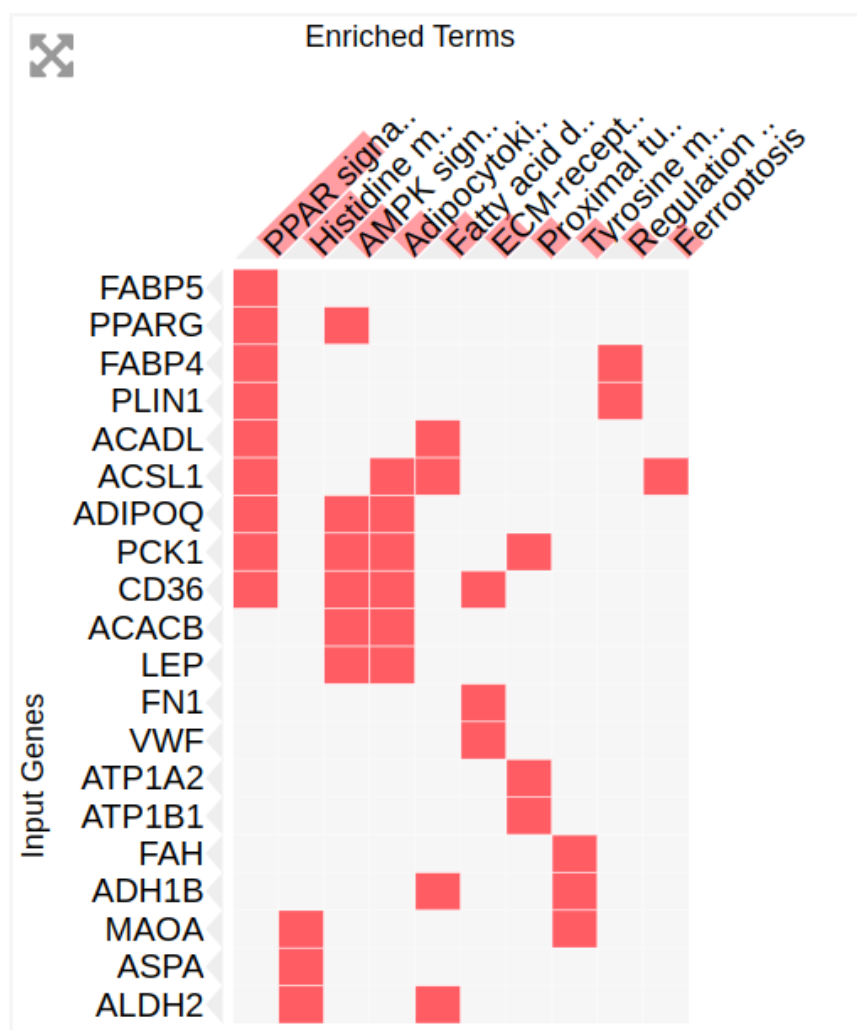
Combined Score

P-Value

Z-score

Top Enriched Terms: 10

Top rows sum: 20 rows



:Clustergram

چگونه از این نمودار استفاده کنیم؟

شناسایی ژن‌های کلیدی:

ژن‌هایی که در چندین مسیر نقش دارند ممکن است ژن‌های کلیدی باشند.

شناسایی مسیرهای مهم:

مسیرهایی که با چندین ژن مرتبط هستند ممکن است مسیرهای مهمی در داده‌ها باشند.

کشف الگوهای جدید:

خوشه‌بندی می‌تواند کمک کند الگوهای جدیدی در داده‌ها کشف کنیم، مثلاً ببینیم کدام ژن‌ها و مسیرها

با هم همکاری می‌کنند. برای مثال طبق نمودار داریم:

• ASPA: احتمالاً با metabolism Histidine مرتبط است.

• ALDH2: احتمالاً با metabolism Tyrosine و degradation acid Fatty مرتبط است.

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	PPAR signaling pathway	4.984e-9	0.000001316	11.35	217.04
2	AMPK signaling pathway	0.000007667	0.001012	5.89	69.35
3	ECM-receptor interaction	0.0001428	0.01238	5.79	51.27
4	Adipocytokine signaling pathway	0.0001876	0.01238	6.53	55.99
5	Histidine metabolism	0.0005032	0.02657	12.76	96.90
6	Estrogen signaling pathway	0.0006440	0.02709	4.07	29.93
7	Focal adhesion	0.0007857	0.02709	3.36	24.05
8	Fatty acid degradation	0.0008448	0.02709	7.57	53.56
9	Proteoglycans in cancer	0.0009234	0.02709	3.29	23.01
10	PI3K-Akt signaling pathway	0.001365	0.03457	2.58	17.03
11	Fc gamma R-mediated phagocytosis	0.001481	0.03457	4.49	29.25
12	Choline metabolism in cancer	0.001571	0.03457	4.44	28.66
13	FoxO signaling pathway	0.002028	0.04118	3.76	23.30
14	Regulation of lipolysis in adipocytes	0.002579	0.04687	5.75	34.27
15	Insulin signaling pathway	0.002681	0.04687	3.58	21.21
16	Fluid shear stress and atherosclerosis	0.002932	0.04687	3.53	20.57
17	Glutathione metabolism	0.003018	0.04687	5.53	32.08
18	Tyrosine metabolism	0.003350	0.04914	7.17	40.87
19	Mineral absorption	0.003775	0.05245	5.23	29.15
20	Ferroptosis	0.005384	0.07107	6.20	32.40
21	Staphylococcus aureus infection	0.006146	0.07726	3.88	19.75
22	Proximal tubule bicarbonate reclamation	0.007032	0.08002	8.59	42.56
23	Glycerophospholipid metabolism	0.007131	0.08002	3.75	18.55
24	Epithelial cell signaling in Helicobacter pylori infection	0.007275	0.08002	4.42	21.76
25	Melanoma	0.008179	0.08637	4.29	20.60

:Table

در جدول بالا ۲۵ سطر اول این دیتا آورده شده است.

p-value بسیار پایین نشان دهنده ارتباط بسیار قوی با داده‌های ست.

ratio odds بالا نشان دهنده احتمال قوی ارتباط این مسیر با ژن‌های ست.

**توضیح مسیرهای زیستی و نقش آن‌ها در سرطان:**

\*۱. مسیر PPAR

• **چی کار می‌کنه:** این مسیر کمک می‌کنه بدن چربی‌ها و قندها رو درست مدیریت کنه و جلوی التهاب رو هم بگیره.

• **رابطه با سرطان:** اگر این مسیر درست کار نکنه، سلول‌ها ممکنه بی‌رویه رشد کنن و سرطان ایجاد بشه.

• **نکته مهم:** این مسیر می‌تونه یه هدف خوب برای داروهای ضد سرطان باشه. نقش زیستی: مسیر PPAR Receptors Proliferator-Activated Peroxisome در تنظیم متابولیسم لیپیدها،

گلوکز و التهاب نقش دارد. این مسیر همچنین در تمایز سلولی و آپوپتوز (مرگ برنامه‌ریزی‌شده سلولی) دخیل است.

ارتباط با سرطان: اختلال در مسیر PPAR می‌تواند منجر به رشد غیرطبیعی سلول‌ها و مقاومت به آپوپتوز شود، که هر دو از ویژگی‌های سرطان هستند. فعال‌سازی نادرست این مسیر ممکن است باعث پیشرفت تومور و متاستاز شود

#### \*۲. مسیر AMPK

- **چی کار می‌کند:** مثل یه حسگر انرژی و وقتی بدن کم‌بود انرژی داره، فعال می‌شه.
- **رابطه با سرطان:** اگر این مسیر درست کار نکنه، سلول‌های سرطانی می‌تونن بیشتر رشد کنن.
- **نکته مهم:** فعال کردن این مسیر می‌تونه به کاهش رشد تومور کمک کنه. نقش زیستی: مسیر AMPK-activated AMP kinase protein در تنظیم انرژی سلولی و هموستاز متابولیک نقش دارد. این مسیر به‌عنوان یک حسگر انرژی عمل می‌کند و در شرایط کم‌بود انرژی فعال می‌شود.
- **ارتباط با سرطان:** در سرطان، اختلال در مسیر AMPK می‌تواند منجر به افزایش متابولیسم سلولی و رشد تومور شود. فعال‌سازی این مسیر ممکن است به مهار رشد تومور کمک کند.

#### \*۳. تعامل ECM-Receptor

- **چی کار می‌کند:** این مسیر کمک می‌کند سلول‌ها با محیط اطرافشون ارتباط برقرار کنن.
- **رابطه با سرطان:** اگر این مسیر به هم بریزه، سلول‌های سرطانی می‌تونن به بقیه قسمت‌های بدن پخش بشن.
- **نکته مهم:** این مسیر می‌تونه نشون بده که تومور چقدر خطرناکه.
- **نقش زیستی:** این مسیر به تعامل بین سلول‌ها و ماتریکس خارج سلولی (ECM) مربوط می‌شود و در انتقال سیگنال‌های مکانیکی و شیمیایی نقش دارد.
- **ارتباط با سرطان:** تغییرات در ECM و تعامل آن با سلول‌ها می‌تواند به مهاجرت سلول‌های سرطانی و متاستاز کمک کند. اختلال در این مسیر اغلب با تهاجمی‌تر شدن تومورها همراه است.

#### \*۴. مسیر آدیپوکاین

- **چی کار می‌کند:** این مسیر به هورمون‌هایی مربوطه که از چربی بدن ترشح می‌شن و التهاب رو کنترل می‌کنن.



- **رابطه با سرطان:** اگر این هورمون‌ها به هم بریزن، ممکنه به رشد سرطان کمک کنن.
- **نکته مهم:** کنترل این مسیر می‌تونه به درمان سرطان‌های مرتبط با چاقی کمک کنه. نقش زیستی: این مسیر به سیگنال‌دهی توسط آدیپوکاین‌ها (هورمون‌های ترشح‌شده از بافت چربی) مربوط می‌شود و در تنظیم التهاب و متابولیسم نقش دارد.
- ارتباط با سرطان: آدیپوکاین‌ها می‌توانند رشد تومور و التهاب مرتبط با سرطان را تحت تأثیر قرار دهند. سطوح غیرطبیعی آدیپوکاین‌ها ممکن است به پیشرفت سرطان کمک کند.

#### \*۵. متابولیسم هیستیدین

- **چی کار می‌کنه:** این مسیر به بدن کمک می‌کنه تا به اسید آمینه به نام هیستیدین رو پردازش کنه.
- **رابطه با سرطان:** اگر این مسیر به هم بریزه، سلول‌های سرطانی می‌تونن بیشتر رشد کنن.
- **نکته مهم:** این مسیر می‌تونه به درک بهتر نقش تغذیه در سرطان کمک کنه. نقش زیستی: این مسیر در متابولیسم هیستیدین، یک اسید آمینه ضروری، نقش دارد.
- ارتباط با سرطان: تغییرات در متابولیسم هیستیدین ممکن است بر رشد سلول‌های سرطانی تأثیر بگذارد، زیرا اسیدهای آمینه برای تکثیر سلولی ضروری هستن.

#### \*۶. مسیر استروژن

- **چی کار می‌کنه:** این مسیر به هورمون استروژن مربوطه و رشد سلول‌ها رو کنترل می‌کنه.
- **رابطه با سرطان:** اگر این مسیر زیاد فعال بشه، ممکنه باعث سرطان‌هایی مثل سرطان پستان بشه.
- **نکته مهم:** داروهایی که این مسیر رو بلوک می‌کنن، می‌تونن به درمان سرطان کمک کنن. نقش زیستی: این مسیر به سیگنال‌دهی استروژن و تنظیم رشد و تمایز سلولی مربوط می‌شود.
- ارتباط با سرطان: در سرطان‌های وابسته به هورمون (مانند سرطان پستان)، فعال‌سازی بیش از حد این مسیر می‌تواند به رشد تومور کمک کند.

#### \*۷. اتصال کانونی Adhesion Focal

- **چی کار می‌کنه:** این مسیر کمک می‌کنه سلول‌ها به بافت‌های اطرافشون بچسبن.
- **رابطه با سرطان:** اگر این مسیر به هم بریزه، سلول‌های سرطانی می‌تونن به بقیه جاهای بدن پخش بشن.

- **نکته مهم:** این مسیر می‌تونه به جلوگیری از پخش سرطان کمک کنه. نقش زیستی: این مسیر در اتصال سلول‌ها به ماتریکس خارج سلولی و انتقال سیگنال‌های مکانیکی نقش دارد.
- ارتباط با سرطان: اختلال در این مسیر می‌تواند به مهاجرت سلول‌های سرطانی و متاستاز منجر شود.

#### \* ۸. تجزیه اسیدهای چرب

- **چی کار می‌کنه:** این مسیر کمک می‌کنه بدن اسیدهای چرب رو برای تولید انرژی تجزیه کنه.
- **رابطه با سرطان:** سلول‌های سرطانی از این مسیر برای رشد خودشون استفاده می‌کنن.
- **نکته مهم:** مهار این مسیر می‌تونه به کاهش رشد تومور کمک کنه.
- نقش زیستی: این مسیر در تجزیه اسیدهای چرب برای تولید انرژی نقش دارد.
- ارتباط با سرطان: تغییرات در متابولیسم اسیدهای چرب می‌تواند به رشد و بقای سلول‌های سرطانی کمک کند.

#### \* ۹. پروتئوگلیکان‌ها در سرطان

- **چی کار می‌کنه:** این مسیر به سلول‌ها کمک می‌کنه تا با محیط اطرافشون ارتباط برقرار کنن.
- **رابطه با سرطان:** اگر این مسیر به هم بریزه، تومورها می‌تونن بیشتر رشد کنن و پخش بشن.
- **نکته مهم:** این مسیر می‌تونه یه هدف خوب برای داروهای ضد سرطان باشه. نقش زیستی: پروتئوگلیکان‌ها در تنظیم سیگنال‌دهی سلولی و تعامل با ماتریکس خارج سلولی نقش دارند.
- ارتباط با سرطان: تغییرات در پروتئوگلیکان‌ها می‌تواند به پیشرفت تومور و متاستاز کمک کند.

#### \* ۱۰. مسیر PI3K-AKT

- **چی کار می‌کنه:** این مسیر به رشد و بقای سلول‌ها کمک می‌کنه.
- **رابطه با سرطان:** اگر این مسیر زیاد فعال بشه، سلول‌های سرطانی می‌تونن بیشتر رشد کنن و در برابر درمان مقاوم بشن.
- **نکته مهم:** این مسیر یه هدف اصلی برای داروهای ضد سرطانه.
- نقش زیستی: این مسیر در رشد، بقا و متابولیسم سلولی نقش دارد.
- ارتباط با سرطان: فعال‌سازی بیش از حد این مسیر یکی از شایع‌ترین تغییرات در سرطان است و به رشد و مقاومت تومور در برابر درمان کمک می‌کند.

## بخش دوم تحلیل GO

### تحلیل فرایندهای زیستی:

#### GO Biological Process 2023

Bar Graph

Table

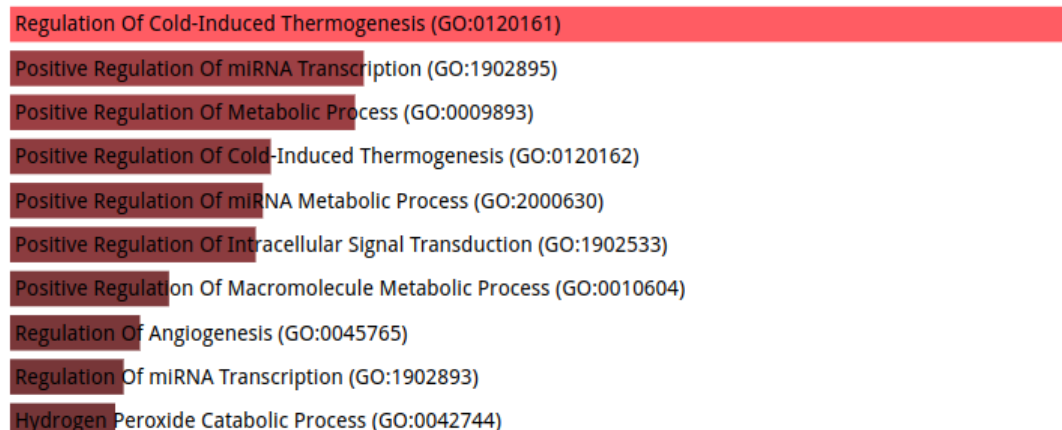
Clustergram

Appyter

⚙️ ⓘ

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

SVG PNG JPG



#### GO Biological Process 2023

Bar Graph

Table

Clustergram

Appyter

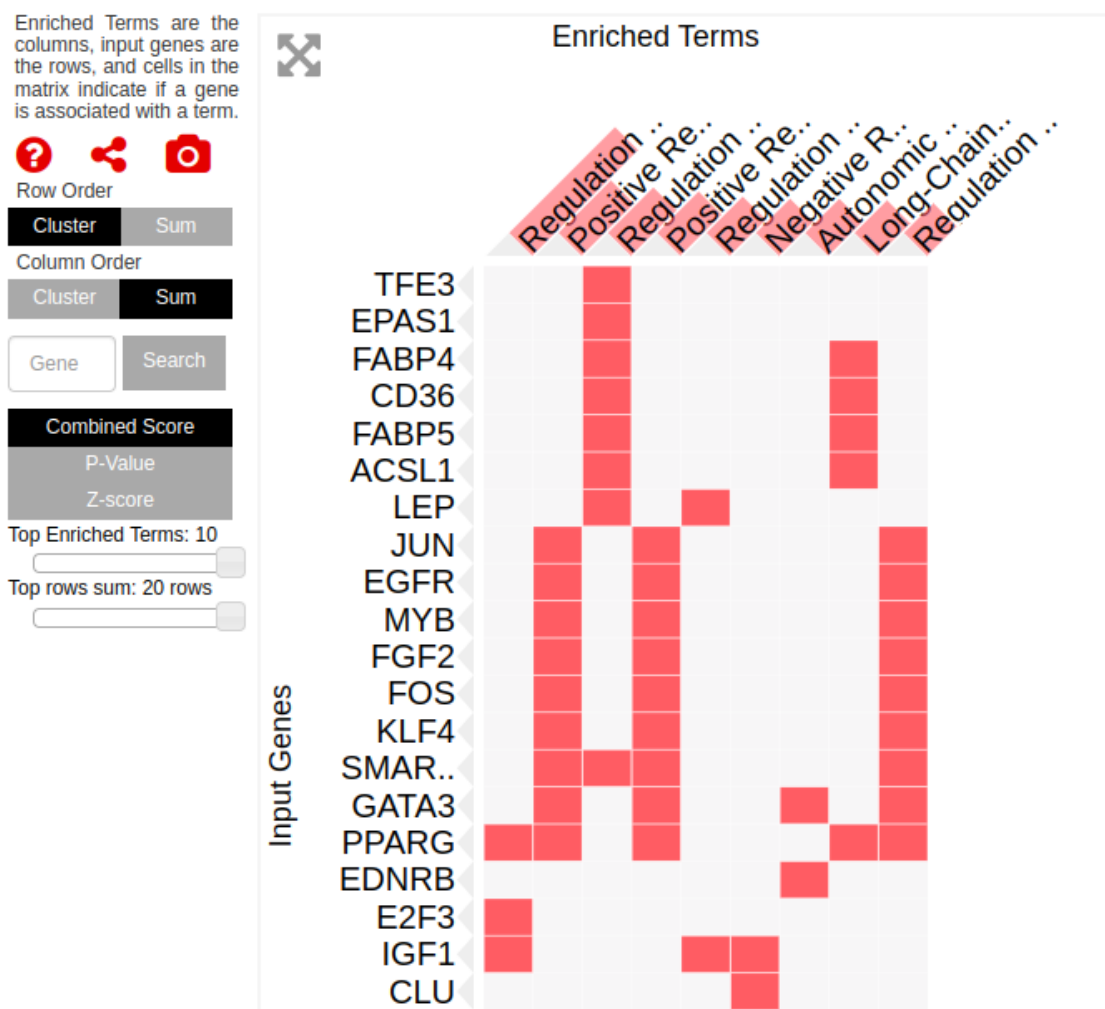
⚙️ ⓘ

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	Regulation Of Cold-Induced Thermogenesis (GO:0120161)	8.899e-12	2.119e-8	8.93	227.32
2	Positive Regulation Of miRNA Transcription (GO:1902895)	4.202e-8	0.00003702	15.41	261.76
3	Positive Regulation Of Metabolic Process (GO:0009893)	4.664e-8	0.00003702	8.04	135.68
4	Positive Regulation Of Cold-Induced Thermogenesis (GO:0120162)	1.272e-7	0.00006038	8.17	129.79
5	Positive Regulation Of miRNA Metabolic Process (GO:2000630)	1.397e-7	0.00006038	13.10	206.69
6	Positive Regulation Of Intracellular Signal Transduction (GO:1902533)	1.522e-7	0.00006038	3.39	53.28
7	Positive Regulation Of Macromolecule Metabolic Process (GO:0010604)	4.274e-7	0.0001454	3.83	56.23
8	Regulation Of Angiogenesis (GO:0045765)	6.039e-7	0.0001797	4.99	71.51
9	Regulation Of miRNA Transcription (GO:1902893)	7.341e-7	0.0001930	10.47	147.90
10	Hydrogen Peroxide Catabolic Process (GO:0042744)	8.108e-7	0.0001930	24.76	347.21



- **تنظیم ترموزنز ناشی از سرما:** - این فرآیند ممکنه با متابولیسم انرژی در سلولهای سرطانی مرتبط باشه، مخصوصاً تو تومورهایی که تو محیطهای کم اکسیژن رشد می کنند.  
- p-value خیلی پایینه و ratio odds بالاست که نشون می ده این فرآیند خیلی با ژن ها مرتبطه.
- **تنظیم مثبت رونویسی miRNA - miRNA** نقش مهمی تو تنظیم بیان ژن ها دارن و اگر به هم بریزن، می تونن به رشد تومور و متاستاز کمک کنن.  
- p-value پایینه و ratio odds خیلی بالاست که نشون می ده این فرآیند خیلی قوی با ژن ها مرتبطه.
- **تنظیم مثبت فرآیندهای متابولیک:** - سلولهای سرطانی برای رشد سریع و زنده موندن، نیاز دارن متابولیسمشون رو خوب تنظیم کنن.  
- p-value پایینه و ratio odds بالاست ، که نشون می ده این فرآیند با ژن ها مرتبطه.
- **تنظیم مثبت ترموزنز ناشی از سرما:** - این فرآیند هم مثل اولی ممکنه با متابولیسم انرژی در

سلول‌های سرطانی مرتبط باشد.

- p-value پایین و ratio odds بالاست که نشان می‌دهد این فرآیند با ژن‌ها مرتبط است.

• **تنظیم مثبت فرآیندهای متابولیک miRNA:** - این فرآیند هم مثل رونویسی miRNA می‌تونه تو

رشد تومور و متاستاز نقش داشته باشه.

- p-value پایین و ratio odds بالاست خیلی بالاست، که نشان می‌دهد این فرآیند خیلی

قوی با ژن‌ها مرتبط است.

• **تنظیم مثبت انتقال سیگنال درون سلولی:** - انتقال سیگنال‌های درون سلولی نقش کلیدی تو

رشد و بقای سلول‌های سرطانی داره.

- p-value پایین و ratio odds بالاست متوسطه که نشان می‌دهد این فرآیند با ژن‌ها مرتبط است.

• **تنظیم مثبت فرآیندهای متابولیک ماکرومولکول‌ها:** - این فرآیند تو سنتز و تجزیه ماکرومولکول‌ها

(مثل پروتئین‌ها و اسیدهای نوکلئیک) نقش داره که برای رشد تومور ضروریه.

- p-value پایین و ratio odds بالاست متوسطه، که نشان می‌دهد این فرآیند با ژن‌ها

مرتبط است.

• **تنظیم رگزایی:** - رگزایی فرآیندی که تومورها برای تأمین خون و مواد مغذی بهش نیاز دارن.

- p-value پایین و ratio odds بالاست که نشان می‌دهد این فرآیند با ژن‌ها مرتبط است.

• **تنظیم رونویسی miRNA:** - این فرآیند هم مثل بقیه فرآیندهای miRNA می‌تونه تو رشد تومور و

متاستاز نقش داشته باشه.

- p-value پایین و ratio odds بالاست، که نشان می‌دهد این فرآیند با ژن‌ها مرتبط است.

• **فرآیند کاتابولیک پراکسید هیدروژن:** - پراکسید هیدروژن یه مولکول سیگنالی که می‌تونه

تو استرس اکسیداتیو و بقای سلول‌های سرطانی نقش داشته باشه.

- p-value پایین و ratio odds خیلی بالاست، که نشان می‌دهد این فرآیند خیلی قوی با ژن‌ها

مرتبط است.

## تحليل اجزای سلولی:

### GO Cellular Component 2023

Bar Graph

Table

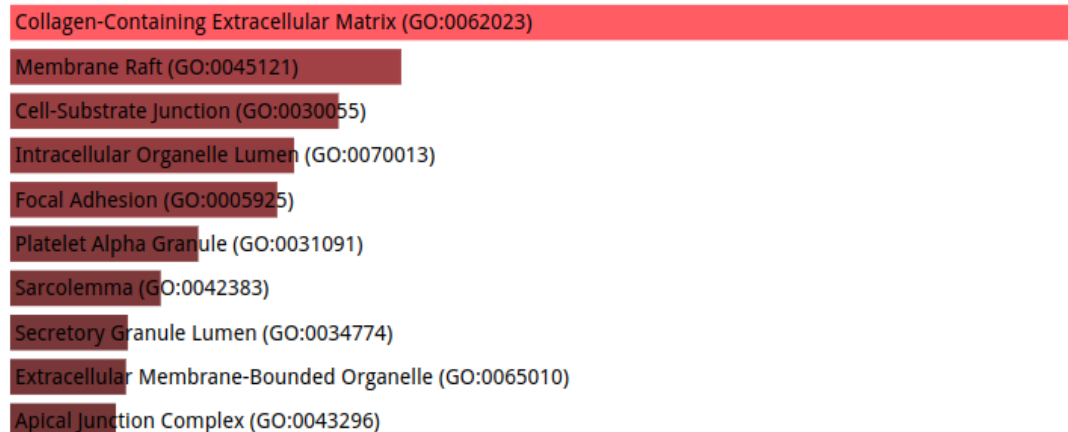
Clustergram

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.

10 entries per page

Search:

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	Collagen-Containing Extracellular Matrix (GO:0062023)	4.172e-10	8.678e-8	4.72	102.01
2	Membrane Raft (GO:0045121)	0.000008073	0.0008396	4.88	57.22
3	Cell-Substrate Junction (GO:0030055)	0.00002008	0.001392	3.15	34.11
4	Intracellular Organelle Lumen (GO:0070013)	0.00003837	0.001995	2.33	23.68
5	Focal Adhesion (GO:0005925)	0.00004904	0.002040	3.05	30.22
6	Platelet Alpha Granule (GO:0031091)	0.0001546	0.005360	5.72	50.18
7	Sarcolemma (GO:0042383)	0.0002673	0.007941	7.52	61.89
8	Secretory Granule Lumen (GO:0034774)	0.0004312	0.01024	2.91	22.58
9	Extracellular Membrane-Bounded Organelle (GO:0065010)	0.0004431	0.01024	6.78	52.38
10	Apical Junction Complex (GO:0043296)	0.0005124	0.01066	4.72	35.78

Showing 1 to 10 of 208 entries | [Export entries to table](#)

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

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

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

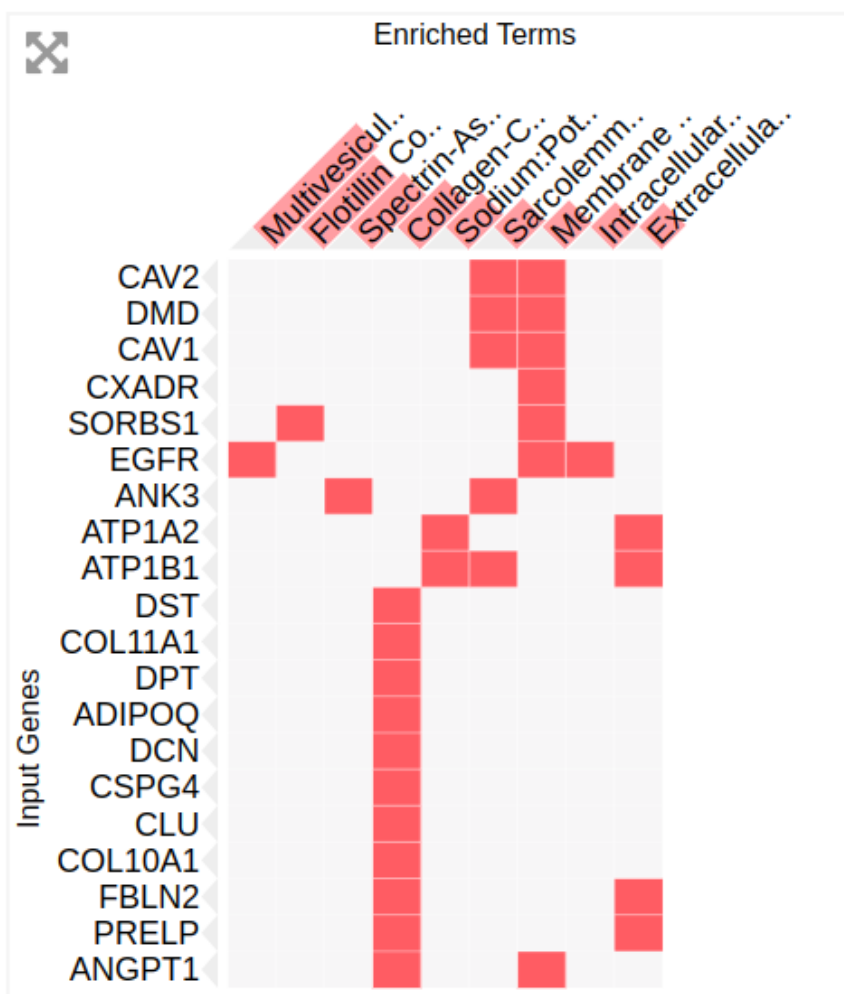
Combined Score

P-Value

Z-score

Top Enriched Terms: 10

Top rows sum: 20 rows



### • ماتریس خارج سلولی حاوی کلاژن:

ژن‌های معنادار: ژن‌هایی مثل COL11A1، COL10A1 و DCN تو این بخش فعال هستن.

اهمیت: این ماتریس برای حمایت ساختاری و سیگنال‌دهی سلولی مهمه. تو سرطان، تغییرات تو این بخش می‌تونه به متاستاز و تهاجم تومور کمک کنه.

### • رافت غشایی:

ژن‌های معنادار: ژن‌هایی مثل CAV1 و EGFR تو این بخش فعال هستن.

اهمیت: رافت‌های غشایی تو سیگنال‌دهی و انتقال مواد نقش دارن. این بخش‌ها می‌تونن تو رشد سلولی و پاسخ به سیگنال‌های خارجی مؤثر باشن.

### • اتصال سلول به زیرلایه:

**ژن‌های معنادار:** ژن‌هایی مثل SORBS1 و DST تو این بخش فعال هستن.

**اهمیت:** این اتصالات برای چسبندگی سلول به ماتریکس خارج سلولی و انتقال سیگنال‌های مکانیکی مهم هستن. تو سرطان، اختلال تو این اتصالات می‌تونه به متاستاز کمک کنه.

• **فضای درون اندامک‌های سلولی:**

**ژن‌های معنادار:** ژن‌هایی مثل ATP1A2 و ATP1B1 تو این بخش فعال هستن.

**اهمیت:** این فضاها تو پردازش و انتقال پروتئین‌ها و لیپیدها نقش دارن. تو سرطان، تغییرات تو این فرآیندها می‌تونه به رشد تومور کمک کنه.

• **چسبندگی کانونی:**

**ژن‌های معنادار:** ژن‌هایی مثل EGFR و ANK3 تو این بخش فعال هستن.

**اهمیت:** چسبندگی‌های کانونی برای اتصال سلول به ماتریکس خارج سلولی و انتقال سیگنال‌های مکانیکی و شیمیایی مهم هستن. تو سرطان، این بخش‌ها می‌تونن تو مهاجرت و متاستاز سلول‌های سرطانی نقش داشته باشن.

• **گرانول آلفای پلاکت:**

**ژن‌های معنادار:** ژن‌هایی مثل ADIPOQ و CLU تو این بخش فعال هستن.

**اهمیت:** این گرانول‌ها تو انعقاد خون و ترمیم بافت نقش دارن. تو سرطان، این بخش‌ها می‌تونن تو تشکیل رگ‌های خونی تومور نقش داشته باشن.

• **سارکولما:**

**ژن‌های معنادار:** ژن‌هایی مثل DMD و CAV1 تو این بخش فعال هستن.

**اهمیت:** سارکولما غشای سلول‌های عضلانی و تو انتقال سیگنال‌های الکتریکی و مکانیکی نقش داره. تو سرطان، تغییرات تو این بخش می‌تونه بر عملکرد عضلات تأثیر بذاره.

• **فضای درون گرانول‌های ترشحی:**

**ژن‌های معنادار:** ژن‌هایی مثل CLU و ANGPT1 تو این بخش فعال هستن.

**اهمیت:** این فضاها تو ذخیره و ترشح مواد نقش دارن. تو سرطان، این بخش‌ها می‌تونن تو ترشح فاکتورهای رشد و سیتوکین‌ها نقش داشته باشن.



## • اندامک‌های غشایی خارج سلولی:

ژن‌های معنادار: ژن‌هایی مثل CSPG4 و FBLN2 تو این بخش فعال هستن.

اهمیت: این اندامک‌ها تو انتقال مواد و سیگنال‌دهی بین سلولی نقش دارن. تو سرطان، این بخش‌ها می‌تونن تو متاستاز و تهاجم تومور نقش داشته باشن.

## • کمپلکس اتصال آپیکال:

ژن‌های معنادار: ژن‌هایی مثل CXADR و EGFR تو این بخش فعال هستن.

اهمیت: این کمپلکس‌ها تو اتصال سلول‌های اپیتلیال به هم نقش دارن و تو حفظ یکپارچگی بافت‌ها مهم هستن. تو سرطان، اختلال تو این اتصالات می‌تونه به متاستاز کمک کنه.

## تحلیل عملکردهای مولکولی :

### GO Molecular Function 2023

Bar Graph

Table

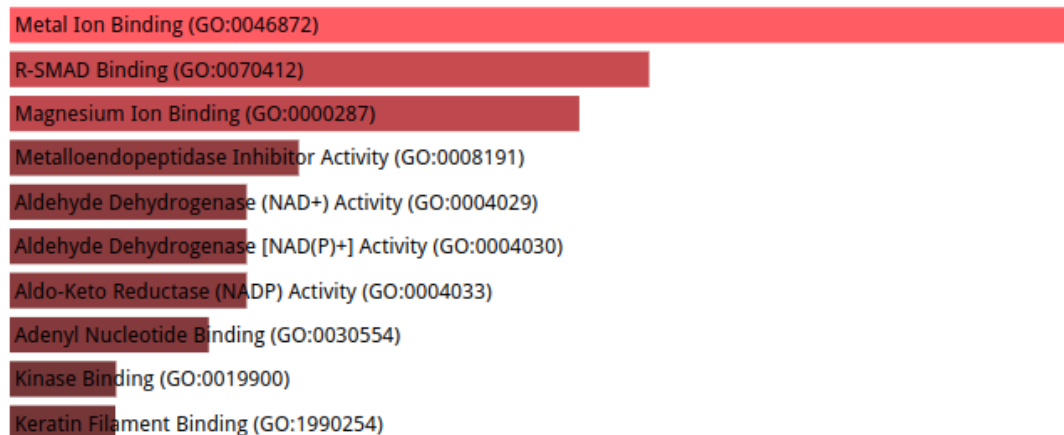
Clustergram

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	Metal Ion Binding (GO:0046872)	0.00004247	0.01741	2.74	27.57
2	R-SMAD Binding (GO:0070412)	0.0002778	0.05155	15.31	125.39
3	Magnesium Ion Binding (GO:0000287)	0.0003772	0.05155	4.00	31.56
4	Metalloendopeptidase Inhibitor Activity (GO:0008191)	0.001291	0.09498	17.18	114.30
5	Aldehyde Dehydrogenase (NAD+) Activity (GO:0004029)	0.001622	0.09498	15.62	100.34
6	Aldehyde Dehydrogenase [NAD(P)+] Activity (GO:0004030)	0.001622	0.09498	15.62	100.34
7	Aldo-Keto Reductase (NADP) Activity (GO:0004033)	0.001622	0.09498	15.62	100.34
8	Adenyl Nucleotide Binding (GO:0030554)	0.001915	0.09815	8.50	53.21
9	Kinase Binding (GO:0019900)	0.002875	0.1071	2.24	13.11
10	Keratin Filament Binding (GO:1990254)	0.002883	0.1071	38.08	222.75

## 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

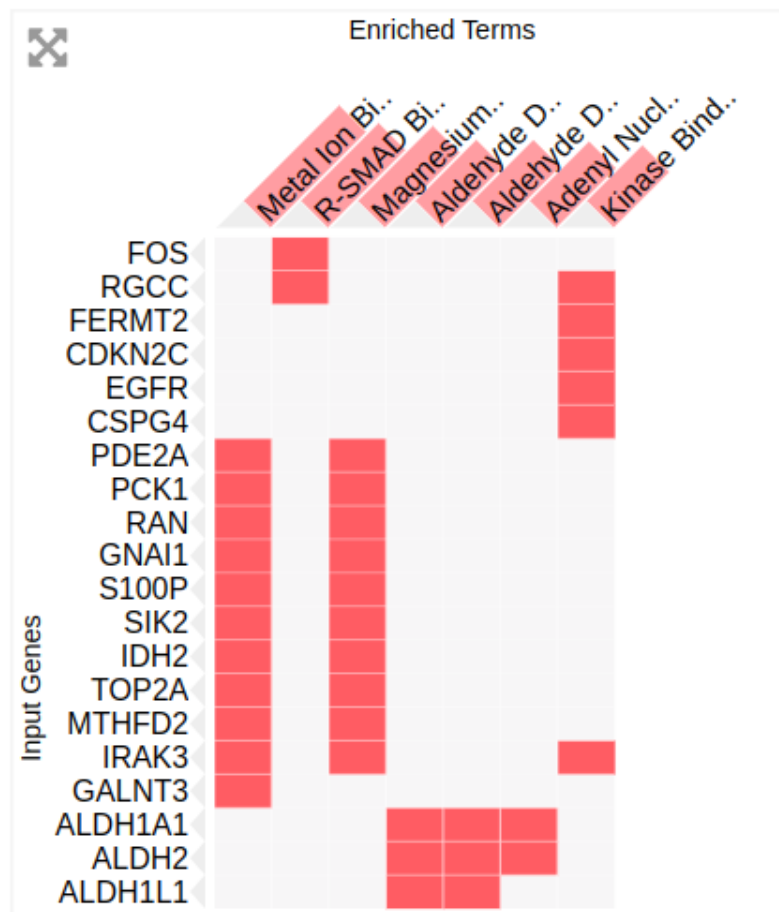
Combined Score

P-Value

Z-score

Top Enriched Terms: 10

Top rows sum: 20 rows



• اتصال به یون فلزی (Binding Ion Metal):

**ژن‌های معنادار:** ژن‌هایی مثل FOS، EGFR، و IDH2 تو این بخش فعال هستند.

**اهمیت:** این عملکرد در اتصال به یون‌های فلزی مثل آهن، روی و مس نقش دارد. این یون‌ها در فعالیت آنزیم‌ها و سیگنال‌دهی سلولی مهم هستند. تو سرطان، اختلال در این عملکرد می‌تونه به رشد تومور و مقاومت به درمان کمک کنه.

• **اتصال به R-SMAD (R-SMAD Binding):**

**ژن‌های معنادار:** ژن‌هایی مثل FERMT2 و EGFR تو این بخش فعال هستند.

**اهمیت:** R-SMAD‌ها در مسیر سیگنالینگ TGF- نقش دارن که در رشد سلولی و تمایز مهمه. تو سرطان، تغییرات در این مسیر می‌تونه به متاستاز و تهاجم تومور کمک کنه.

• **اتصال به یون منیزیم (Binding Ion Magnesium):**

**ژن‌های معنادار:** ژن‌هایی مثل PCK1 و IDH2 تو این بخش فعال هستند.

**اهمیت:** منیزیم در فعالیت آنزیم‌ها و متابولیسم انرژی نقش داره. تو سرطان، تغییرات در این عملکرد می‌تونه بر متابولیسم سلول‌های سرطانی تأثیر بذاره.

• **فعالیت مهارکننده متالوآندوپپتیداز (Activity Inhibitor Metalloendopeptidase):**

**ژن‌های معنادار:** ژن‌هایی مثل CSPG4 و ALDH1A1 تو این بخش فعال هستند.

**اهمیت:** این مهارکننده‌ها در تنظیم تجزیه پروتئین‌ها نقش دارن. تو سرطان، اختلال در این عملکرد می‌تونه به تهاجم تومور و متاستاز کمک کنه.

• **فعالیت آلدئید دهیدروژناز (Activity Dehydrogenase Aldehyde):**

**ژن‌های معنادار:** ژن‌هایی مثل ALDH1A1، ALDH2، و ALDH1L1 تو این بخش فعال هستند.

**اهمیت:** این آنزیم‌ها در متابولیسم الکل و سم‌زدایی نقش دارن. تو سرطان، فعالیت بیش از حد این آنزیم‌ها می‌تونه به مقاومت به درمان کمک کنه.

• **فعالیت آلدو-کتو ردوکتاز (Activity Reductase Aldo-Keto):**

**ژن‌های معنادار:** ژن‌هایی مثل ALDH1A1 و ALDH2 تو این بخش فعال هستند.

**اهمیت:** این آنزیم‌ها در متابولیسم قندها و سم‌زدایی نقش دارن. تو سرطان، تغییرات در این عملکرد می‌تونه بر متابولیسم سلول‌های سرطانی تأثیر بذاره.

• اتصال به نوکلئوتید آدنوزین (Binding Nucleotide Adenyl):

ژن‌های معنادار: ژن‌هایی مثل EGFR و TOP2A تو این بخش فعال هستن.

اهمیت: این عملکرد در متابولیسم انرژی و سیگنال‌دهی سلولی نقش داره. تو سرطان، اختلال در این عملکرد می‌تونه به رشد تومور کمک کنه.

• اتصال به کیناز (Binding Kinase):

ژن‌های معنادار: ژن‌هایی مثل EGFR و SIK2 تو این بخش فعال هستن.

اهمیت: کینازها در انتقال سیگنال‌های رشد و بقای سلولی نقش دارن. تو سرطان، تغییرات در این عملکرد می‌تونه به رشد تومور و مقاومت به درمان کمک کنه.

• اتصال به رشته‌های کراتین (Binding Filament Keratin):

ژن‌های معنادار: ژن‌هایی مثل CSPG4 و FERMT2 تو این بخش فعال هستن.

اهمیت: این عملکرد در حفظ ساختار سلولی و انتقال سیگنال‌ها نقش داره. تو سرطان، تغییرات در این عملکرد می‌تونه به تهاجم تومور کمک کنه.