



## فاز دوم

پروژه مقدمه‌ای بر بیوانفورماتیک  
دکتر علی شریفی‌زارچی و دکتر سمیه کوهی

دانشکده مهندسی کامپیوتر

دانشگاه صنعتی شریف

نیم‌سال اول ۰۱-۰۲

امیرحسین باقری - ۹۸۱۰۵۶۲۱

مهدی مستانی - ۹۷۱۰۰۵۱۳

محمد رضا مفیضی - ۹۸۱۰۶۰۵۹

## فهرست مطالب

۲	۱	بیان ژن‌ها
۲	۱.۱	انتخاب تمام نمونه‌های سالم . . . . .
۲	۲.۱	انتخاب نمونه‌های سالم با بیشترین هم‌بستگی . . . . .
۳	۲	بررسی gene ontology و pathway
۳	۱.۲	ژن‌هایی که بیشتر بیان شده‌اند . . . . .
۳	۱.۱.۲	انتخاب تمام نمونه‌های سالم . . . . .
۴	۲.۱.۲	انتخاب نمونه‌های سالم با بیشترین هم‌بستگی . . . . .
۴	۲.۲	ژن‌هایی که کمتر بیان شده‌اند . . . . .
۷	۱.۲.۲	انتخاب تمام نمونه‌های سالم . . . . .
۷	۲.۲.۲	انتخاب نمونه‌های سالم با بیشترین هم‌بستگی . . . . .
۸	۳	جست‌وجو در مقالات



## ۱ بیان ژن‌ها

### ۱.۱ انتخاب تمام نمونه‌های سالم

در جدول ۱ ژن‌هایی که بیان آن‌ها به طرز معناداری بین گروه سالم و گروه بیماران متفاوت است را نمایش می‌دهیم.<sup>۱</sup>

Gene Symbol	Gene ID	Adj. P-Val	logFC
MPO	4353	3.617e-19	-5.563
FLT3	2322	4.835e-19	-5.250
KIAA0101	9768	6.308e-19	-4.559
BUB1B	701	1.664e-18	-2.756
SUCNR1	56670	1.938e-18	-2.996

جدول ۱: ۵ ژنی که بیان آن‌ها به طرز معناداری بین گروه سالم و گروه بیماران متفاوت است.

حالا ژن‌هایی که در بیماران از بقیه بیشتر بیان شده‌اند و همچنین ژن‌هایی که کمتر از بقیه بیان شده‌اند را در جدول ۲ نمایش می‌دهیم.<sup>۲</sup>

Most Expressed	Least Expressed
STK38	MPO
CBX7	FLT3
PLCL2	KIAA0101
PECR	BUB1B
HLA-F	SUCNR1

جدول ۲: ۵ ژنی که در بیماران بیشتر/کمتر از بقیه بیان شده‌اند.

### ۲.۱ انتخاب نمونه‌های سالم با بیشترین هم‌بستگی

در جدول ۳ ژن‌هایی که بیان آن‌ها به طرز معناداری بین گروه سالم و گروه بیماران متفاوت است را نمایش می‌دهیم.

Gene Symbol	Gene ID	Adj. P-Val	logFC
MPO	4353	7.714e-16	-5.73
PECR	55825	7.714e-16	2.314
KIT	3815	7.126e-15	-5.262
KIAA0101	9768	7.969e-15	-4.308
FLT3	2322	7.969e-15	-5.155

جدول ۳: ۵ ژنی که بیان آن‌ها به طرز معناداری بین گروه سالم و گروه بیماران متفاوت است.

حالا ژن‌هایی که در بیماران از بقیه بیشتر بیان شده‌اند و همچنین ژن‌هایی که کمتر از بقیه بیان شده‌اند را در جدول ۴ نمایش می‌دهیم.

<sup>۱</sup>جدول کامل به همراه این گزارش ضمیمه شده است.  
<sup>۲</sup>جدول کامل این بخش هم به همراه گزارش ضمیمه شده است.



Most Expressed	Least Expressed
PECR	MPO
DDX58	KIT
CBX7	KIAA0101
STK38	FLT3
NR1D2	SUCNR1

جدول ۴: ۵ ژنی که در بیماران بیشتر/کمتر از بقیه بیان شده‌اند.

## ۲ بررسی pathway و gene ontology

### ۱.۲ ژن‌هایی که بیشتر بیان شده‌اند

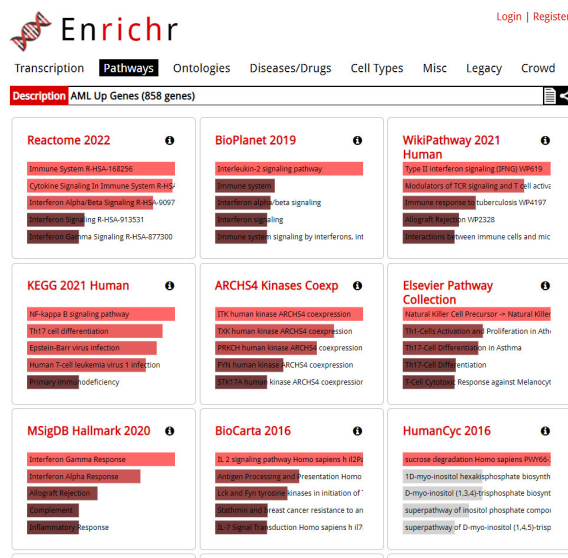
با رفتن به سایت enrichr و قرار دادن لیست ژن‌های بیشتر بیان‌شده (شکل ۱) آن‌ها را بررسی می‌کنیم.

شکل ۱: وارد کردن داده‌ها برای شروع آنالیز

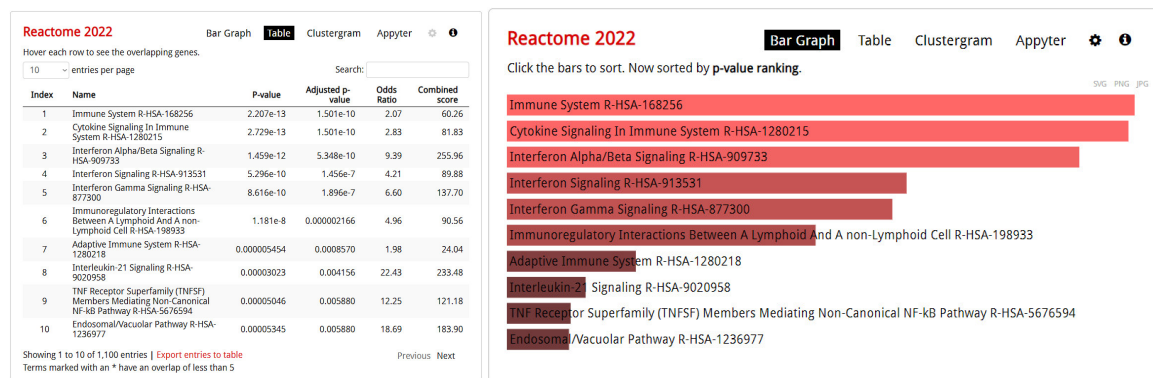
### ۱.۱.۲ انتخاب تمام نمونه‌های سالم

در بخش pathways (شکل ۲) می‌توان پایگاه‌داده‌های pathway مختلف را بررسی کرد. به‌عنوان مثال در پایگاه‌داده Reactome 2022 (شکل ۳) تعداد زیادی از ژن‌هایی که میزان بیان بیشتری داشتند مربوط به pathway سیستم ایمنی<sup>۳</sup> هستند. و همچنین پایگاه‌داده KEGG 2021 در شکل ۴ نمایش داده شده است. در بخش ontologies می‌توان biological processes (شکل ۶) و cellular components (شکل ۷) و molecular functions (شکل ۸) را بررسی کرد (شکل ۵).

<sup>3</sup>Immune System



شکل ۲: بخش pathways در enrichr



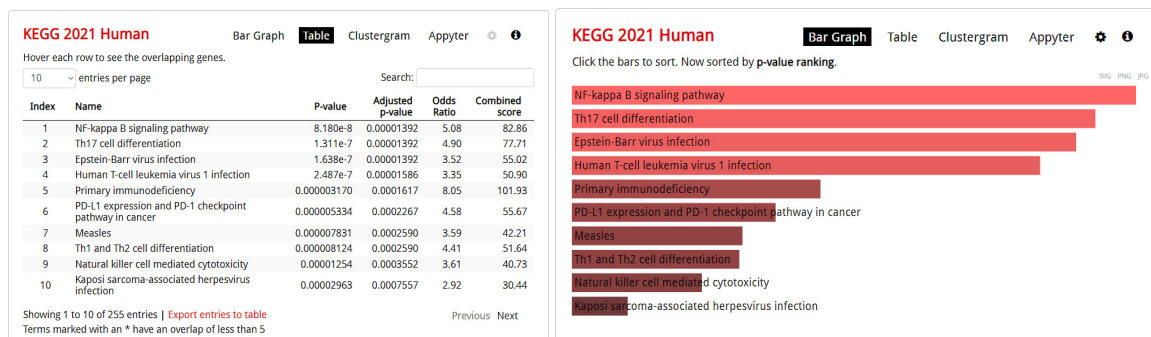
شکل ۳: pathway های پایگاه داده Reactome

## ۲.۱.۲ انتخاب نمونه‌های سالم با بیشترین هم‌بستگی

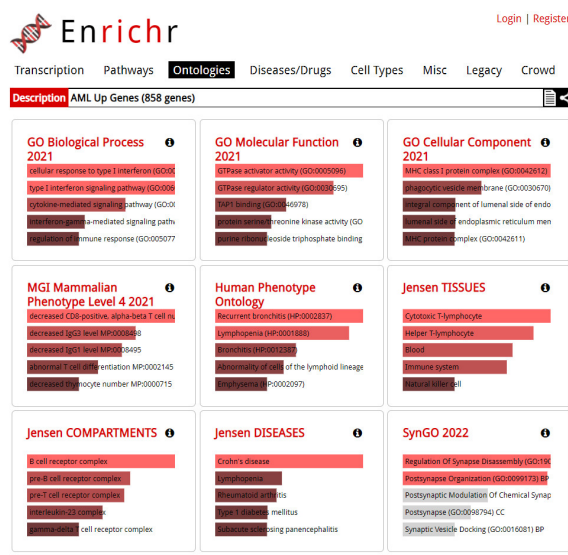
دوباره با رفتن به بخش pathways (شکل ۹) می‌توان پایگاه‌داده‌های pathway مختلف را این‌بار برای نمونه‌های جدید بررسی کرد. پایگاه‌داده Reactome 2022 و همچنین پایگاه‌داده KEGG 2021 در شکل ۱۰ نمایش داده شده است. در بخش ontologies می‌توان biological processes و cellular components و molecular functions را بررسی کرد (شکل ۱۱).

## ۲.۲ ژن‌هایی که کمتر بیان شده‌اند

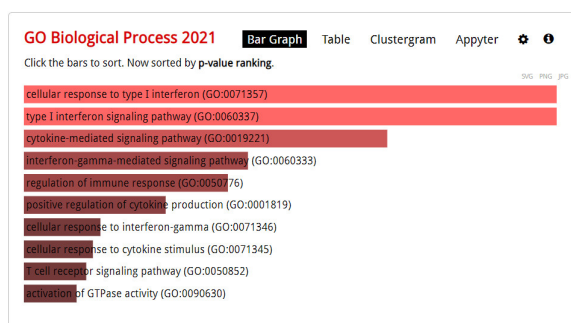
حالا لیست ژن‌های کمتر بیان شده را بررسی می‌کنیم.



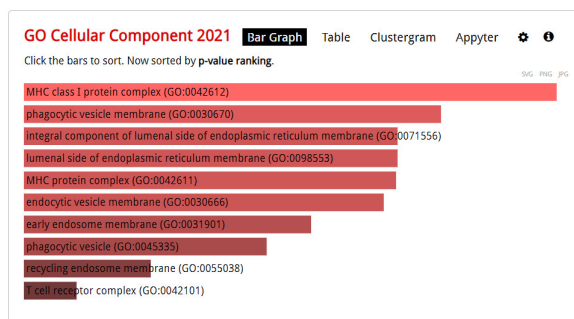
شکل ۴: pathway های پایگاه داده KEGG



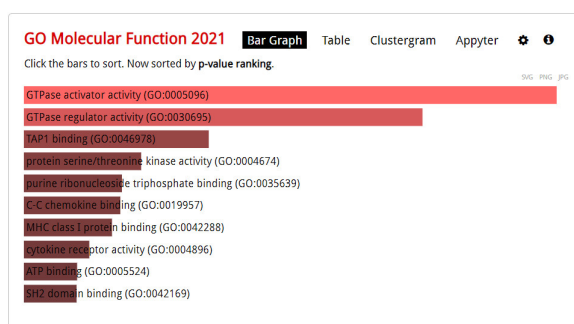
شکل ۵: بخش ontologies در enrichr



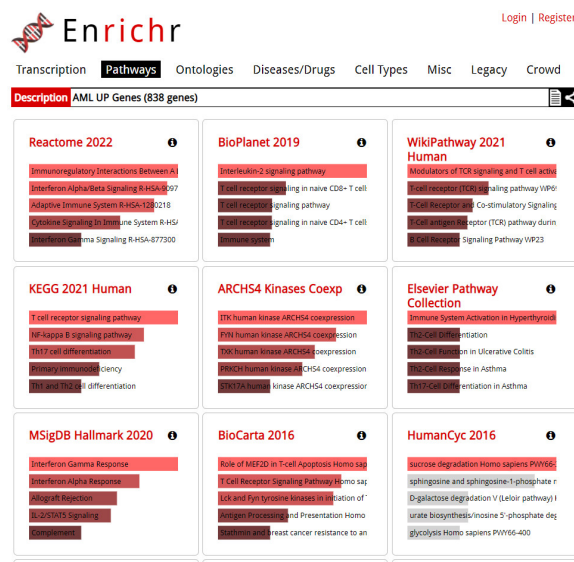
شکل ۶: biological processes در بخش ontologies



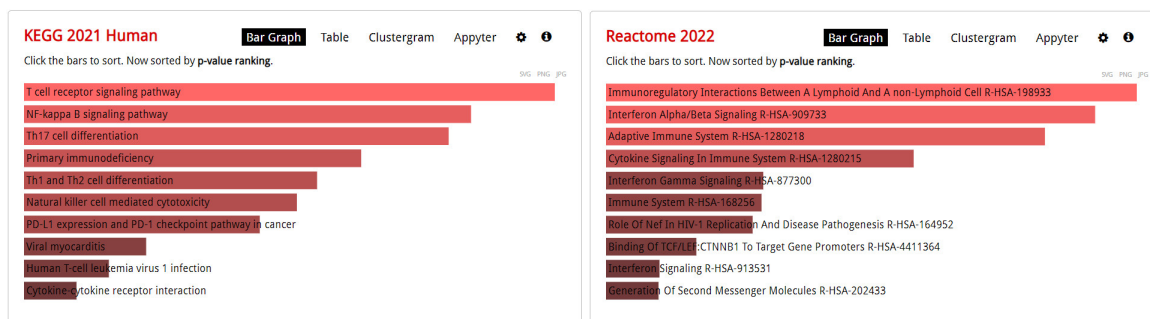
شکل ۷: cellular components در بخش ontologies



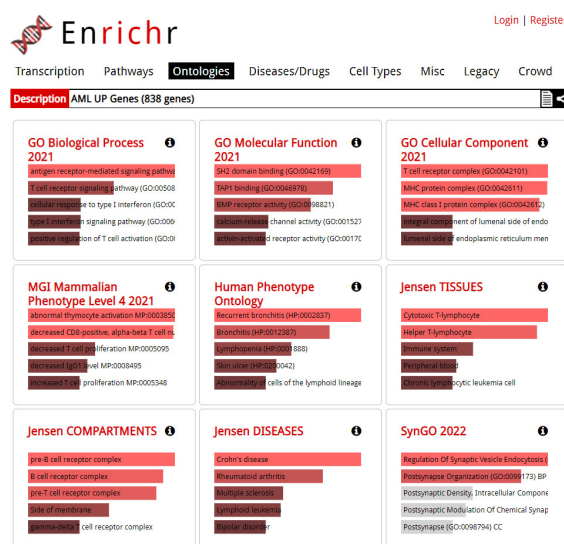
شکل ۸: molecular functions در بخش ontologies



شکل ۹: pathways در بخش enrichr



شکل ۱۰: pathway های پایگاه داده KEGG و Reactome



شکل ۱۱: بخش ontologies در enrichr

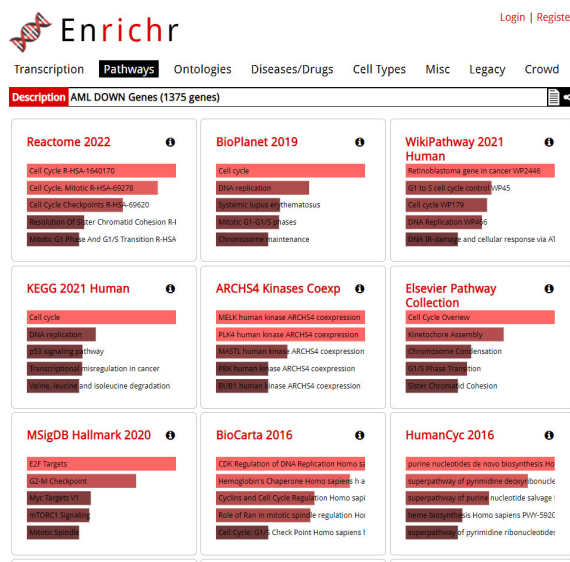
## ۱.۲.۲ انتخاب تمام نمونه‌های سالم

در بخش pathways (شکل ۱۲) می‌توان پایگاه داده‌های pathway مختلف را بررسی کرد. در بخش ontologies می‌توان biological processes و cellular components و molecular functions را بررسی کرد (شکل ۱۳).

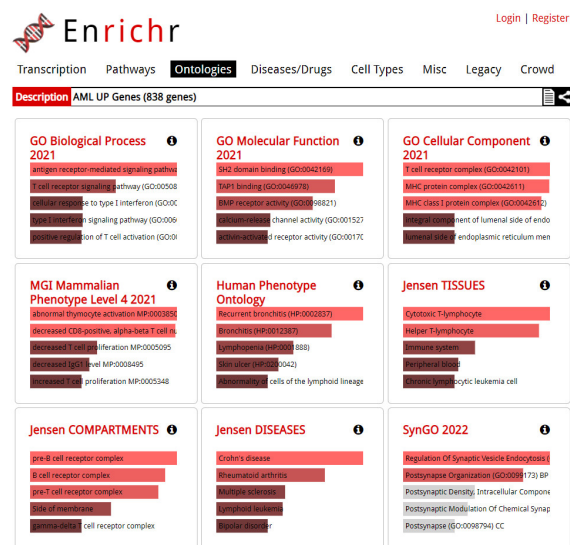
## ۲.۲.۲ انتخاب نمونه‌های سالم با بیشترین هم‌بستگی

دوباره با رفتن به بخش pathways (شکل ۱۴) می‌توان پایگاه داده‌های pathway مختلف را این بار برای نمونه‌های جدید بررسی کرد. در بخش ontologies می‌توان biological processes و cellular components و molecular functions را بررسی کرد (شکل ۱۵).





شکل ۱۲: بخش pathways در enrichr

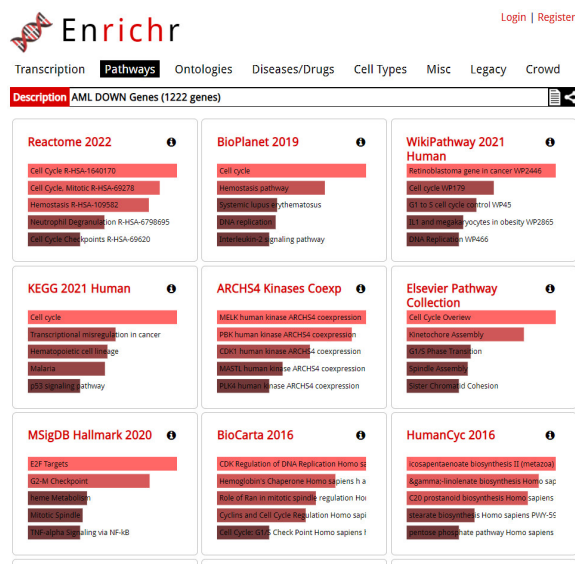


شکل ۱۳: بخش ontologies در enrichr

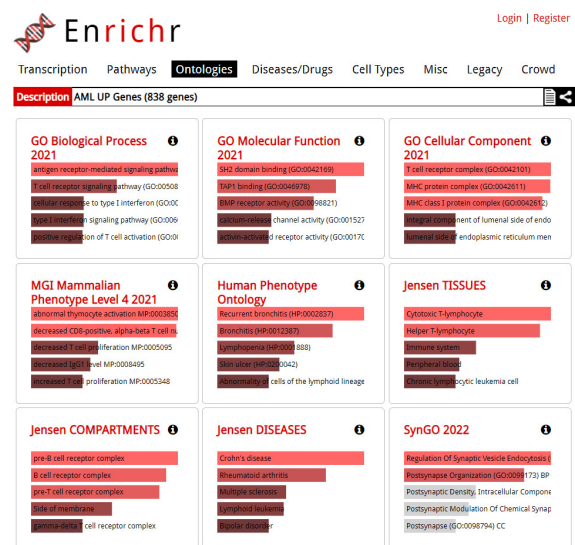
## ۳ جست‌وجو در مقالات

(الف)

پس



شکل ۱۴: بخش pathways در enrichr



شکل ۱۵: بخش ontologies در enrichr

(ب)  
پس



## مراجع