

**فاز دوم پروژه درس بیوانفورماتیک**

تهیه شده توسط:

سیاوش رحیمی۹۹۱۰۹۲۸۵

امیرمحمد ایمانی

**دانشگاه صنعتی شریف**

**پائیز ۱۴۰۱**

در این فاز با بررسی و استفاده از داده‌های موجود از فاز قبل قصد داریم تا ژن‌هایی را که به شکل معنادار میزان بیان ژن‌هایشان با دو گروه بیمار و CD34HSPC متفاوت باشند را پیدا کنیم. توجه کنید که با توجه به داده‌های فاز قبل به طور میانگین بیشترین همبستگی میان نمونه‌های گروه *CD34HSPC* بود که مربوط به سلولهای بنیادی خون میشود.

تعریف دقیق‌تر از :CD34HSPC

human hematopoietic stem/progenitor cell

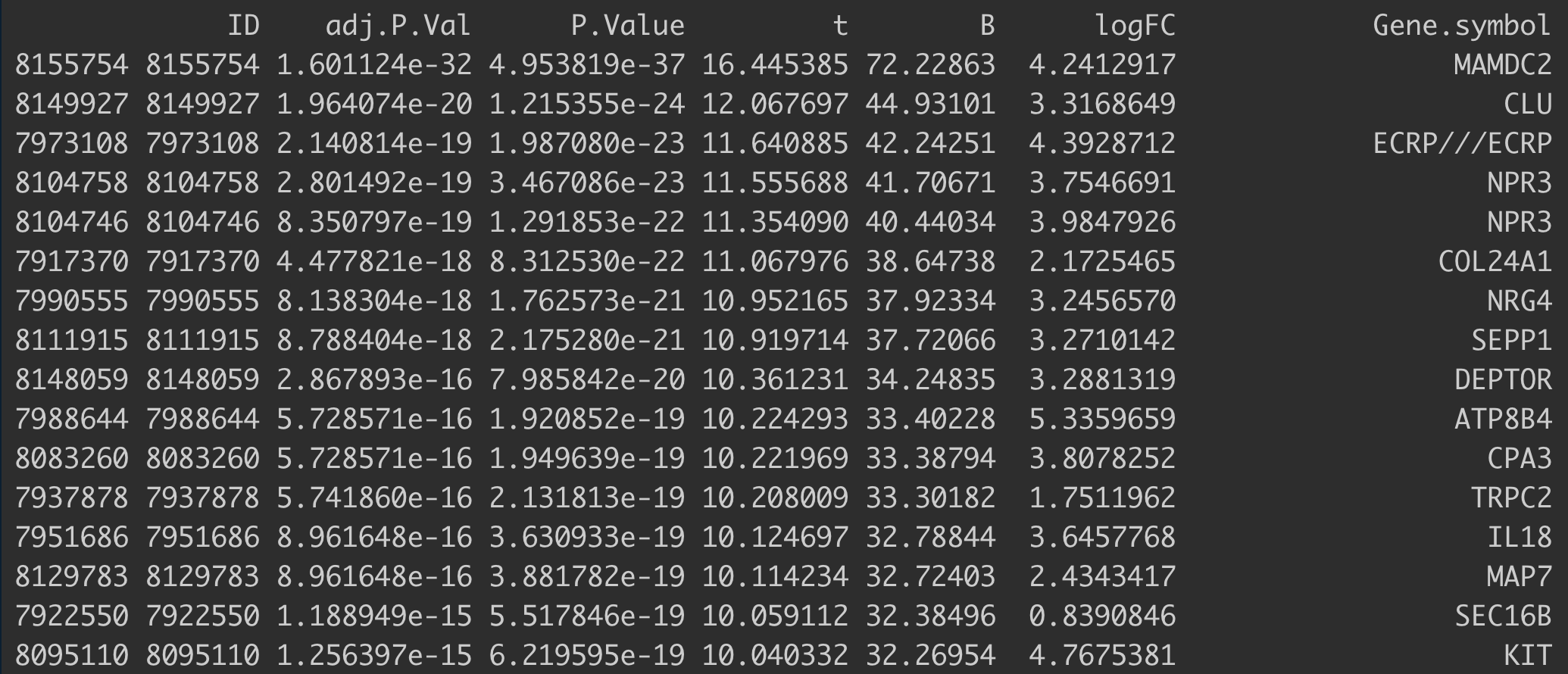
**بخش اول:**

در ابتدا با توجه به Desing Matrix و Contrast Matrix ای که در فاز قبل بین دو گروه AML patient و CD34HSPC محاسبه کردیم جدولی متناسب با داده‌های خواسته شده ایجاد می‌کنیم.

کد متناظر:



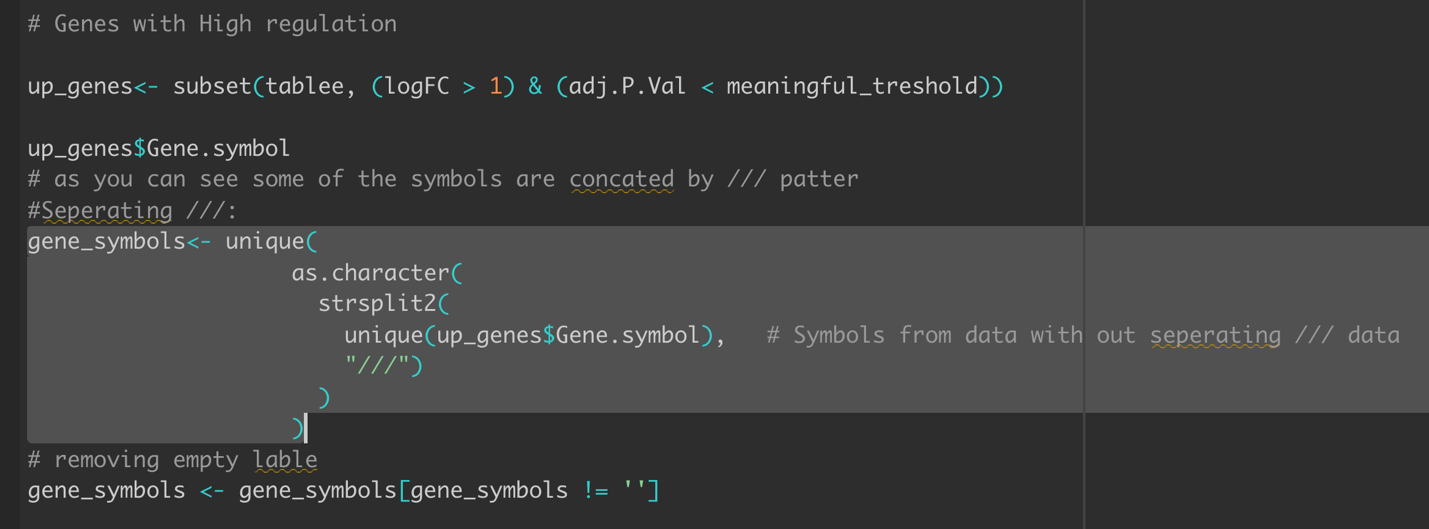
بخشی از جدول ساخته‌شده به صورت:



همانطور که میتوان مشاهده کرد بعضی از داده‌ها چند Gene.symbol دارند که با /// از یکدیگر جدا شده‌اند که ما باید این داده‌هارا در ادامه درست بکنیم.

می‌دانیم که ژن هایی که logFC آن ها از ۱ بیشتر می‌باشد به معنای بیان بالای آن‌ها هست و آن ژن‌هایی که logFC آن‌ها از -۱ کمتر هستند بیان کمی دارند. از طرفی فرض صفر ما عدم تفاوت میان میزان بیان ژنهاست، ژنه'ایی ک'ه مق'دار *adj p-value* آن‌ها از ۰.۰۵ کمتر است دارای تفاوت معنادار در میزان بیان ژن هستند. حال با توجه به عبارات فوق، ژن‌هایی میزان بیانشان تفاوت معنادار دارند را به دست می‌آوریم.

کد متناظر بخش محاسبه ژن‌های با بیان بالا:



کد متناظر بخش محاسبه ژن‌های با بیان پایین:

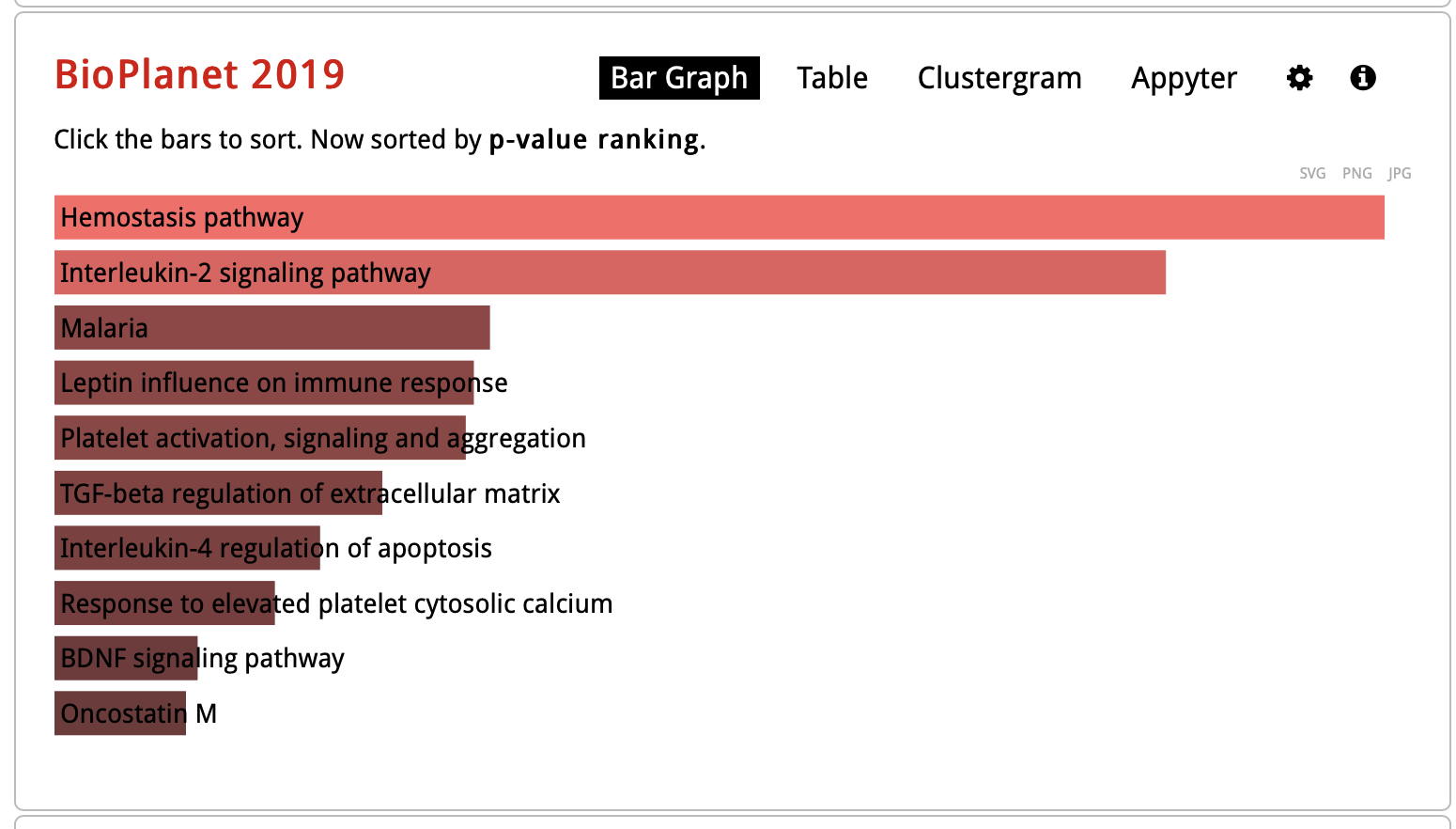


حال سمبل ژن‌های با تفاوت معنادار را در فایل‌های مربوط ذخیره کردیم، و آن‌هارا در سایت Enrchr آپلود می‌کنیم.

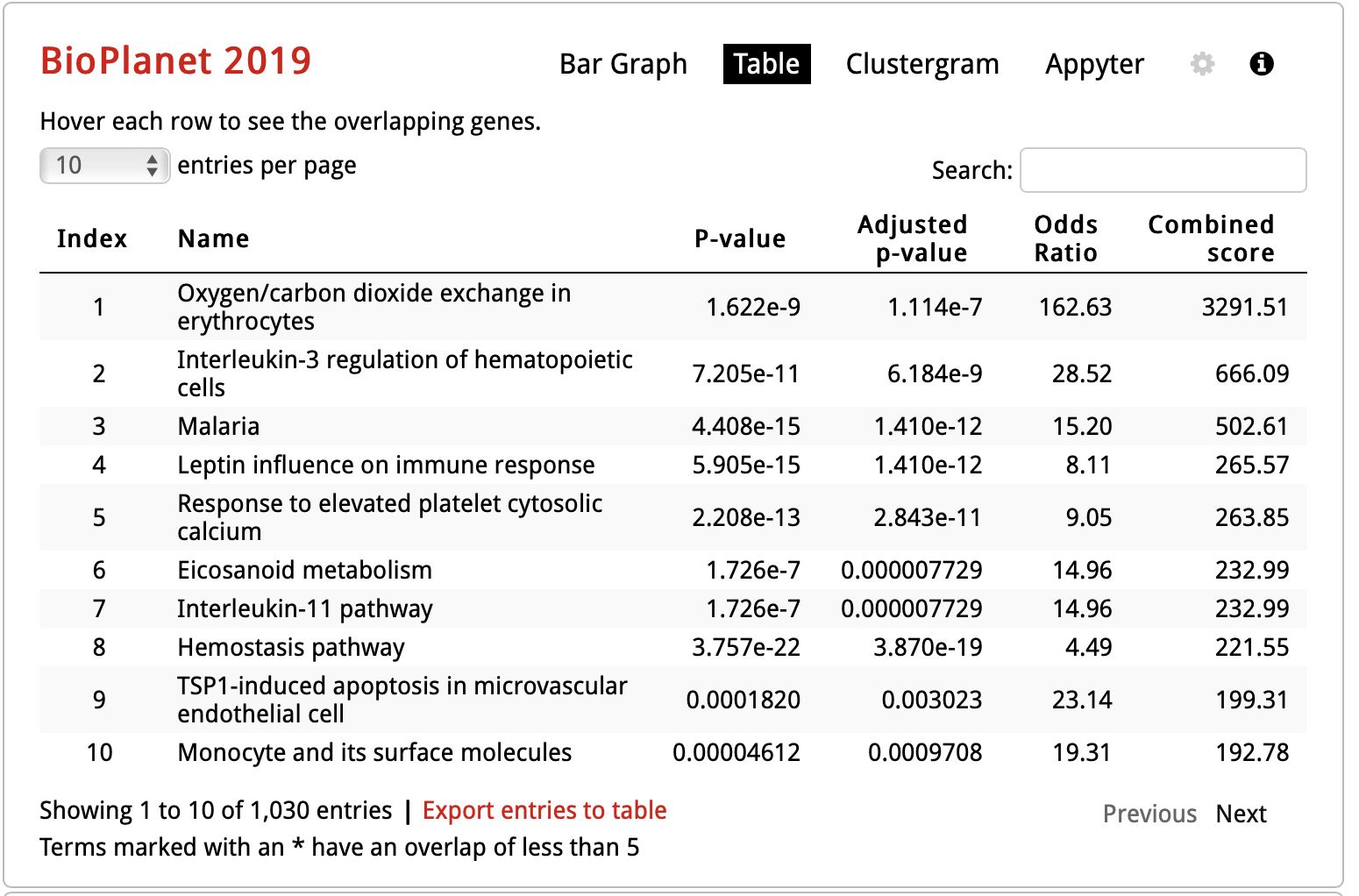
**بخش دوم:**

نتایج بدست آمده برای داده های با بیان‌ بالا:

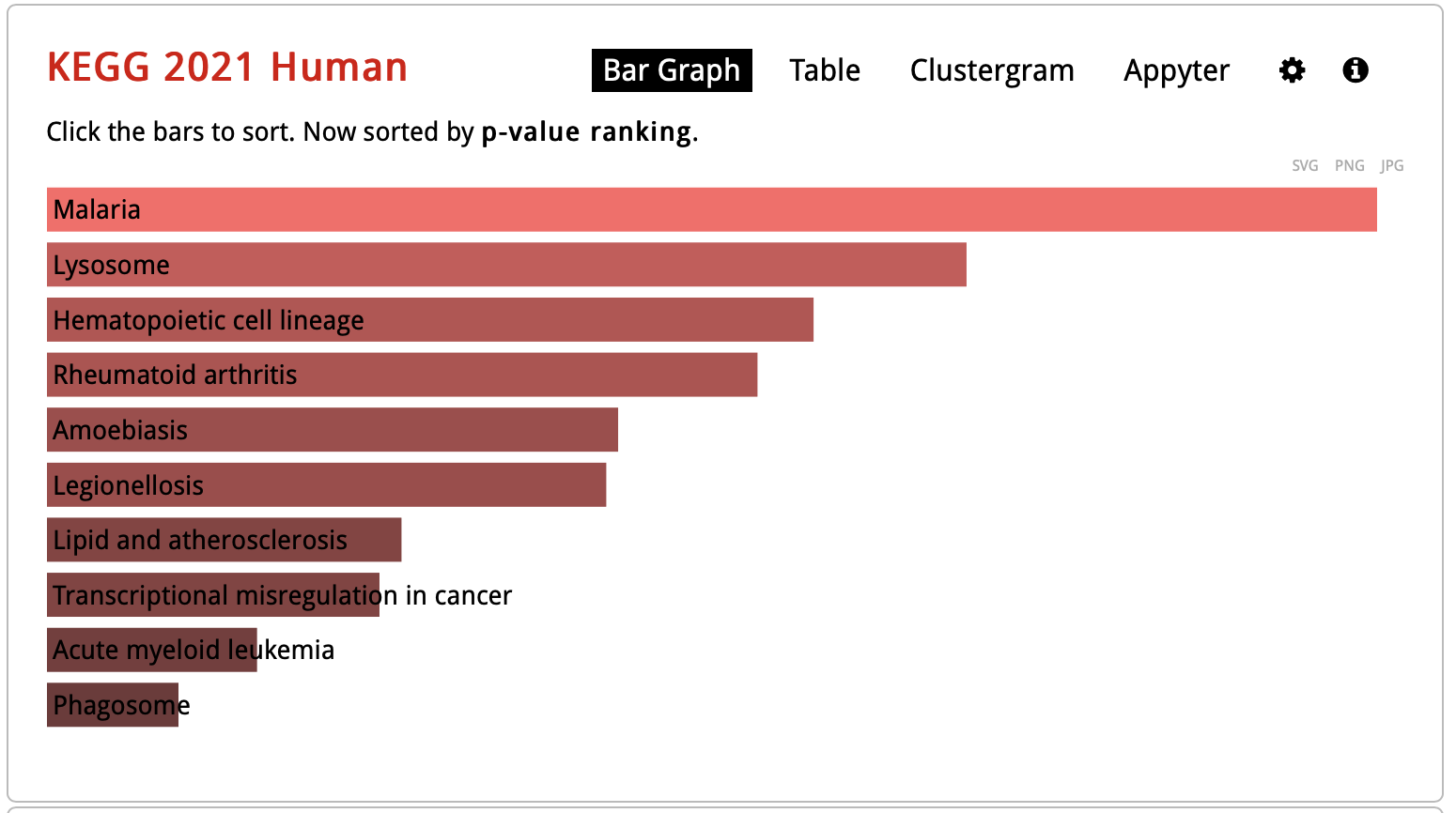
در ابتدا با دیتا ست BioPlanet 2019 اگر برسی کنیم میتوانیم ببینیم که چه pathway هایی که تحت تاثیر ژن های با بیان بالا بودند کدام pathway ها هستند که در این مورد Hemostasis pathway کمترین p-value ممکن را دارد.

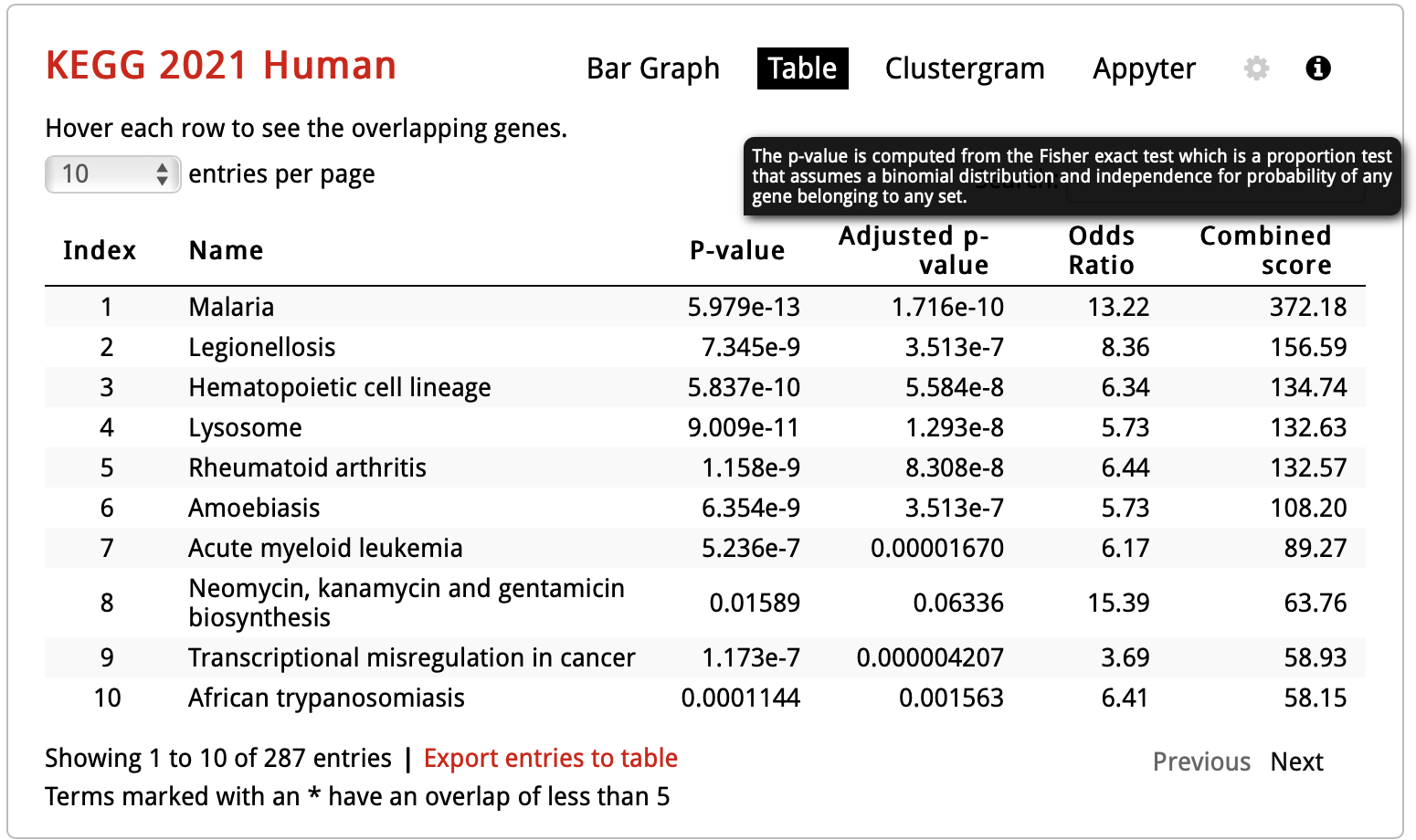


جدول داده متناظر بارپلات بالا:

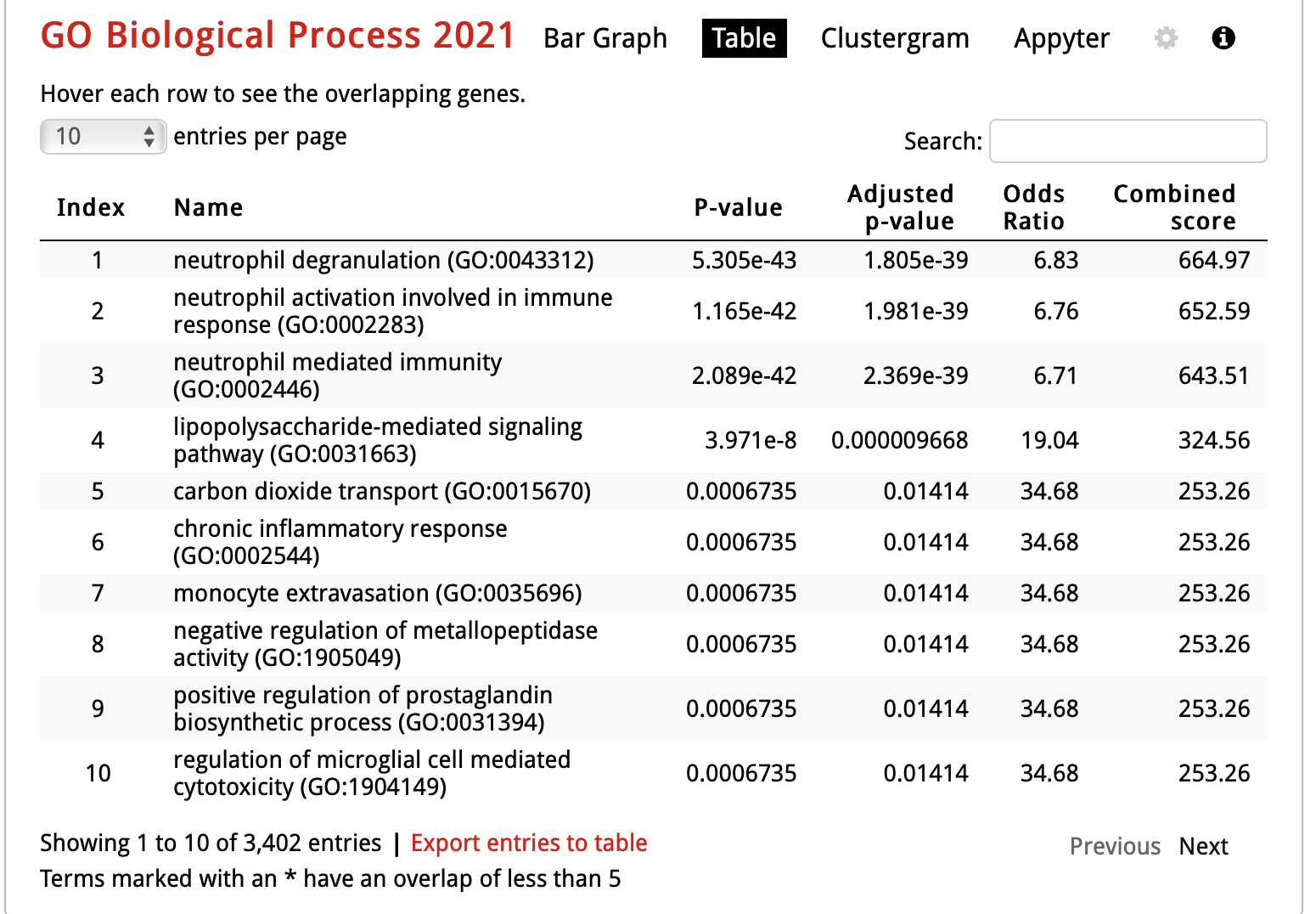


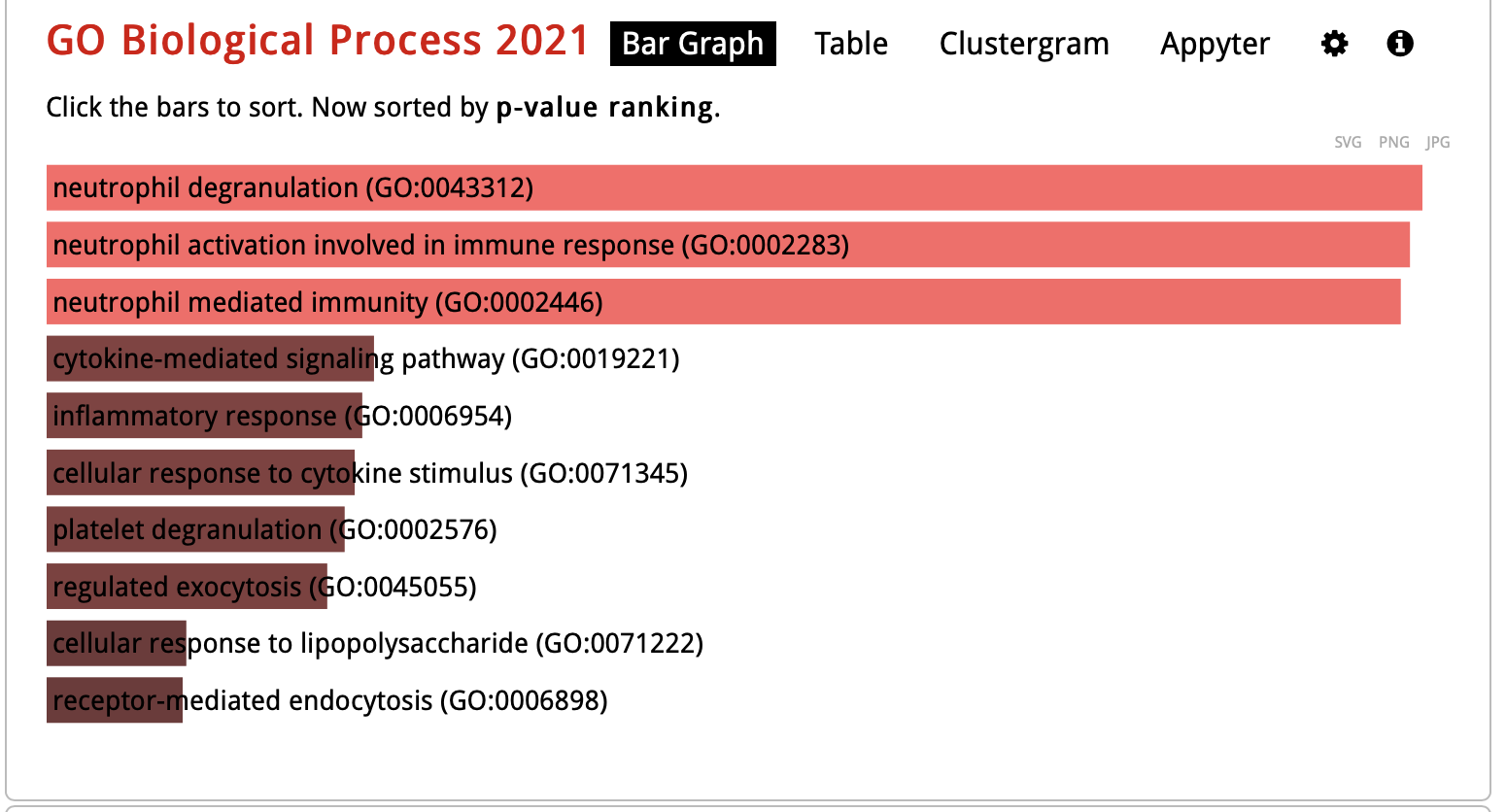
با استفاده از یک دیتا ست دیگر میتوان به داده های زیر هم رسید:





جدول‌های متناظر با gene antology، ژن‌های با بیان بالا:





حال جدول متناظر ژن‌های با بیان کم را هم در ادامه اضافه می‌کنیم:

