STAT1 SCEPTRE vs Seurat With Score QC Individual

2023-04-04

Introduction

This is a followup analysis of Gene's `STAT1-analysis-v2` writeup, with two main differences: **The ChIP-seq data is from CD14+ monocytes rather than K562 and the Seurat QC is matched to that of SCEPTRE.** Here, we consider two ways of determining transcription factor target genes: those that are in the hTFtarget database and those that have at least one ChIP-seq peak within 5kb of their TSS. These two align reasonably well; see the table below.

Table 1: Comparing target genes identified based on database and based directly on ChIP-seq data (at least one peak within 5kb).

Database \ ChIP-seq	FALSE	TRUE
FALSE	4164	725
TRUE	5789	3827

Compare SCEPTRE and Seurat results with ChIP-seq scores

Let's first look at how the SCEPTRE and Seurat discoveries align with each other.

Table 2: Comparing SCEPTRE versus Seurat discoveries.

SCEPTRE \ Seurat	FALSE	TRUE
FALSE	6424	1407
TRUE	596	4280

This table suggests that SCEPTRE and Seurat results have decent, but imperfect, agreement. Also note that the total numbers of discoveries made by the two methods are nearly the same. Next, let's look at how the SCEPTRE and Seurat discoveries align with the ChIP-seq target genes (as identified by either the hTFtarget database or directly from the ChIP-seq data).

Table 3: Comparing SCEPTRE to database.

SCEPTRE Database	FALSE	TRUE	Prop
FALSE	2640	1259	0.323
TRUE	5191	3617	0.411
Prop	0.663	0.742	

Table 4: Comparing Seurat to database.

Seurat Database	FALSE	TRUE	Prop
FALSE	2282	1617	0.415
TRUE	4738	4070	0.462
Prop	0.675	0.716	

Table 5: Comparing SCEPTRE to ChIP-seq binary scores.

SCEPTRE ChIP-seq	FALSE	TRUE	Prop
FALSE	5417	3024	0.358
TRUE	2414	1852	0.434
Prop	0.308	0.38	

Table 6: Comparing Seurat to ChIP-seq binary scores.

Seurat ChIP-seq	FALSE	TRUE	Prop
FALSE	4836	3605	0.427
TRUE	2184	2082	0.488
Prop	0.311	0.366	

Table 7: Enrichment odds ratios, comparing to database and our ChIP-seq target assignments with pvalue quantile threshold equal to 0.75

Ground truth ⟍ Method	SCEPTRE	Seurat
database	1.461	1.212
ChIP-seq	1.374	1.279