Identification of Neurodegenerative Factors Using Translatome-Interactome Analysis : Documentation

William Shin

December 26, 2014

1 Introduction

This document is designed to outline the set-up and running of MARINa on the saline- and MPTP-specific translatomes. When compiled, the document will run the computational analysis described in the manuscript and output the results as a table and plot (corresponding to Figure 3C in the main text).

1.1 Installing the MARINa package

The version of MARINa that was used in this analysis can be found as part of the ssMarina package, which can be downloaded through figshare `http://dx.doi.org/10.6084/m9.figshare.785718`, as well as DESeq, which is used for the normalization of the RNASEQ dataset. An additional package called xtable is also required for the correct display of the results tables. In a new R-session, the libraries can be opened with the following commands:

```
> library(ssmarina)
> library(DESeq)
> library(xtable) # needed for printing out the MARINA results later
```

2 Run MARINa and Output Results

The following code is use to to load the expression file and interactome, run MARINa, output the results as a table and plot. The code itself has the following structure :

1. The function loads the regulon and expression file, and stores the annotation information.
2. The function performs the normalization of the expression file using DETransform.
3. The function then generates the null distrubution by shuffling the sample labels 1000 times.
4. The function runs marina with the regulon, signature, and null distribution.
5. The MARINa results are also annotated, which allows the gene names to be displayed in the pdf file.
6. The table of the full MARINa results is generated and printed.
7. The pdf plot file of the MARINa results is also generated and printed.
8. The function outputs the identity and differential expression of the ARACNe-predicted targets for each MR as a table.

```
> # reading in the regulon and expression file
> load("mus_brain_regulon.rda")
> dset = read.table("Parkinsons_CountFile_20141110_annot.txt", header=T, sep="\t",as.is=T)
> #storing annotation information
> entrez2symbol<-dset[,2]
> names(entrez2symbol)<- as.character(dset[,1])
> symbol2entrez <- as.character(dset[,1])
> names(symbol2entrez) <- dset[,2]
> # convert expression file to matrix
> dset.matrix = as.matrix(dset[,3:ncol(dset)])
> dset = dset.matrix
> rownames(dset) = names(entrez2symbol)
> # DE transform to perform normalization
> dset.DESeq = newCountDataSet( countData = dset, conditions =  rep(1, ncol(dset)))
> dset.DEtransformed = DEtransform(dset.DESeq)
> dset = dset.DEtransformed
> # samples for day-4 after MPTP treatment
> Treatment_4 = grep( "M+._4$", colnames(dset))
> Control_4 = grep( "S+._4$", colnames(dset))
> # generate signature, null distribution, and run marina
> signature <- rowTtest( dset[, Treatment_4], dset[, Control_4], alternative="two.sided")$statistic
> dnull = ttestNull( dset[ , Treatment_4] , dset[ ,  Control_4] , per = 1000, verbose=F, seed=539 )
> marina_results <- marina(signature, regul, dnull, minsize=20,verbose=F)
> marina_results_annot = marinaAnnot(marina_results, entrez2symbol) # annotate results
> # table to print out the full MARINa Results.
> toPrint = data.frame( names(marina_results_annot$es$nes),
+                       marina_results_annot$es$nes,
+                       marina_results_annot$es$p.value )
> colnames( toPrint ) = c("Entrez", "NES", "pvalue")
> sortedind = sort(toPrint$NES, index.return = T, decreasing=T)
> toPrint = toPrint[sortedind$ix, ]
> toPrint.sig = toPrint[which(toPrint$pvalue < 0.01),]
> # writing table of full results
> write.table( toPrint.sig, file="MarinaResults.txt", sep="\t", quote=F, row.names=F)
> # find significant results
> # generating the pdf of the significant (pvalue < 0.01) MR candidates
> num = length(toPrint$pvalue[which(toPrint$pvalue < 0.01)])
> # setting the size of the pdf file
> height_num  = ceiling( num / 5 )
> if(height_num < 5){
+   height_num = 5
+ }
> # plot the significant MR results as a pdf file.
> pdf("MarinaResults_Final.pdf", height = height_num)
> plot(marina_results_annot, num)
> dev.off()
```

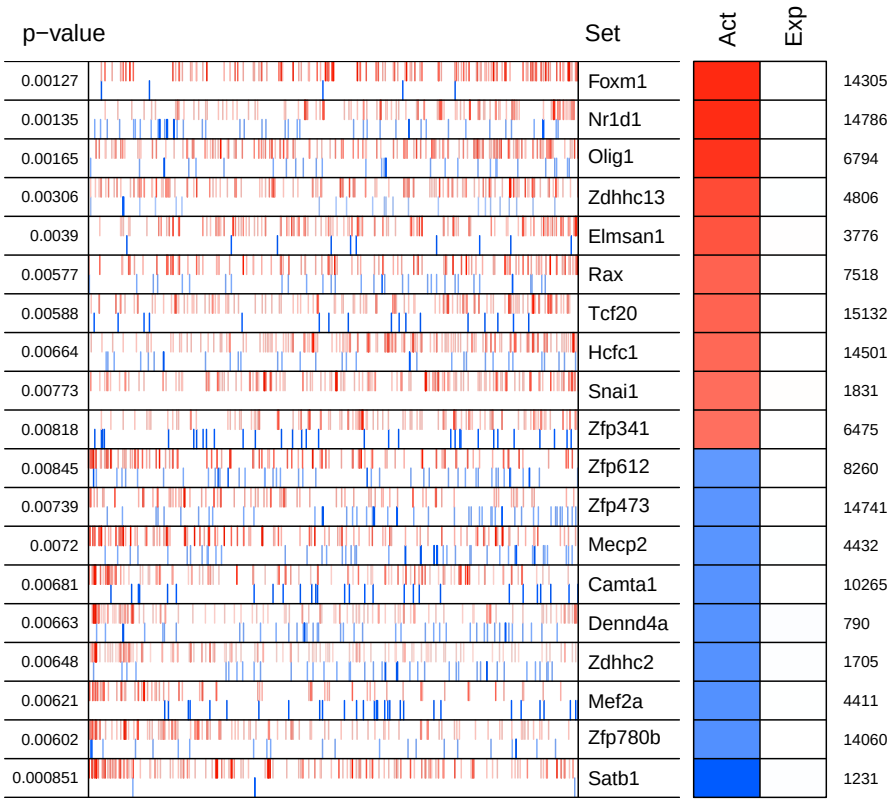

Figure 1: MR Candidates for saline- and MPTP-specific translatomes

Entrez	NES	pvalue
Foxm1	3.223	0.001
Nr1d1	3.205	0.001
Olig1	3.146	0.002
Zdhhc13	2.962	0.003
Elmsan1	2.886	0.004
Rax	2.761	0.006
Tcf20	2.754	0.006
Hcfc1	2.714	0.007
Snai1	2.664	0.008
Zfp341	2.645	0.008
Zfp612	-2.634	0.008
Zfp473	-2.679	0.007
Mecp2	-2.687	0.007
Camta1	-2.706	0.007
Dennd4a	-2.715	0.007
Zdhhc2	-2.722	0.006
Mef2a	-2.737	0.006
Zfp780b	-2.747	0.006
Satb1	-3.336	0.001

Table 1: MR candidates for saline- and MPTP-specific translatomes