

Quantify deregulation of pathways in cancer

Yotam Drier

April 26, 2022

Contents

1	Introduction	1
2	Analyzing data set using pathifier library	1
2.1	Load the library	1
2.2	Load the data set	2
2.3	Load pathway database	2
2.4	Quantify deregulation of pathways	2
2.5	Brief analysis	2

1 Introduction

The *pathifier* is an algorithm that infers pathway deregulation scores for each tumor sample on the basis of expression data. This score is determined, in a context-specific manner, for every particular dataset and type of cancer that is being investigated. The algorithm transforms gene-level information into pathway-level information, generating a compact and biologically relevant representation of each sample [1]. As the algorithm learns the pathway normal flow from normal samples, expression of relevant normal samples must be supplied, however, the structure of the pathway need not be given, just the identity of genes taking part in each pathway.

2 Analyzing data set using pathifier library

2.1 Load the library

Load the pathifier library:

```
> library(pathifier)
```

2.2 Load the data set

Load the data set ‘Sheffer’ (built in Pathifier package) [2].

```
> data(Sheffer)
```

2.3 Load pathway database

Load the two pathways (MISMATCH REPAIR and REGULATION OF AUTOPHAGY) by KEGG [3] (supplied with the package)

```
> data(KEGG)
```

2.4 Quantify deregulation of pathways

Calculate the deregulation score by running pathifier

```
> PDS<-quantify_pathways_deregulation(sheffer$data, sheffer$allgenes,  
+ kegg$gs, kegg$pathwaynames, sheffer$normals, attempts = 100,  
+ min_exp=sheffer$minexp, min_std=sheffer$minstd)
```

```
robust_score_bydist. min_exp= 4 , min_std= 0.2254005  
pathway 1 > sig: 0.07988982  
pathway 2 > sig: 0.1022222  
2 pathways processed with start= by ranks
```

The deregulation scores are now in `PDS$scores`, ready for further analysis.

2.5 Brief analysis

Show scores for normals samples are generally lower

```
> x<-NULL  
> x$normals<-PDS$scores$MISMATCH_REPAIR[sheffer$normals]  
> x$tumors<-PDS$scores$MISMATCH_REPAIR[!sheffer$normals]  
> boxplot(x)  
> boxplot(x,ylab="score")
```



List samples whose regulation of autophagy is highly deregulated

```
> as.character(sheffer$samples[PDS$scores$REGULATION_OF_AUTOPHAGY>0.8])
```

```
[1] "00485_U_13/02/2004_63_F_WN_2_1_0_4_DOD_80_Y_19_S_B_W_M_TT_M_0_8_L_4.8_XX_2_2_8
[2] "00888_B_2/7/2003_69_M_WN_3_0_0_0_DUN_46_X_46_S_?_W_W_TT_M_0_9_L_4_XX_X_0_21
[3] "00947_U_2/7/2003_62_M_WN_3_1_1_4_DOD_39_Y_5_0_B_W_M_TT_0_0_7_L_X_XX_2_1_17
[4] "03752_A_3/7/2003_54_M_WN_2_0_0_1_NED_136_X_136_S_G_W_M_TT_M_Y_8_L_6_S3_2_0_11
[5] "03820_V_22/07/2003_72_F_WN_1_0_0_4_NED_200_Y_83_0_?_W_W_0_M_0_7_L_X_XX_2_0_X
```

References

- [1] Drier Y, Sheffer M and Domany E, Pathway-based personalized analysis of cancer, *PNAS*, 2013, vol. 110(16) pp:6388-6393.
- [2] Sheffer M, Bacolod MD, Zuk O, Giardina SF, Pincas H, et al. Association of survival and disease progression with chromosomal instability: A genomic exploration of colorectal cancer., *PNAS*, 2009, Vol 106(17) pp:7131–7136.

- [3] Kanehisa M, Goto S, Sato Y, Furumichi M and Tanabe M. KEGG for integration and interpretation of large-scale molecular datasets, *Nucleic Acids Res*, 2012, Vol 40(Database issue):D109–D114.