StarBioTrek:Application Examples

Claudia Cava, Isabella Castiglioni

2018-11-13

Introduction

In this vignette, we demonstrate the application of `StarBioTrek` as tool for pathways analysis integrating different data types. For basic use of the `StarBioTrek` package, please refer to the vignette `Working with StarBioTrek package`.

`StarBioTrek` is used as tool to measure pathway activity and pathway cross-talk integrating TCGA data.

Case Study n 1 Relationship between metabolism and cell growth and death in cancer

The aim is the study of ratio among metabolism and cellular processes as cell growth and death in the cancer.

According to KEGG pathway there are different pathways involved in the metabolism that can be grouped in six big sets as: Carbohydrate metabolism, Energy metabolism, Lipid metabolism, Aminoacid metabolism, Glycan biosynthesis and metabolism and Metabolism of cofactors and vitamins.

In this case we want to see if there is a correlation between lipid metabolism and cellular processes in cancer.

First of all we download the set of pathways for the analyses.

For lipid metabolism:

```
path_lip<-getKEGGdata(KEGG_path="Lip_met")
```

The set of pathways for lipid metabolism includes:Fatty acid biosynthesis, Fatty acid elongation, Fatty acid degradation, Synthesis and degradation of ketone bodies, Cutin, suberine and wax biosynthesis, Steroid biosynthesis, Primary bile acid biosynthesis, Secondary bile acid biosynthesis, Steroid hormone biosynthesis, Glycerolipid metabolism, Glycerophospholipid metabolism, Ether lipid metabolism, Sphingolipid metabolism, Arachidonic acid metabolism, Linoleic acid metabolism, alpha-Linolenic acid metabolism and Biosynthesis of unsaturated fatty acids.

For cellular processes:

```
pathcell_grow_d<-getKEGGdata(KEGG_path="cell_grow_d")
```

The set of pathways for cellular processes includes:Cell cycle, Apoptosis and p53 signaling pathway.

Then, we use the function `dev_std_crtlk` to create a measure of pathway cross-talk (pairwise pathway measure) using TCGA data (e.g. Data_CANCER_normUQ_filt).

```
score_euc_dist_Lip_met<-dev_std_crtlk(dataFilt=Data_CANCER_normUQ_filt,path_lip)
```

The function `svm_classification` is used to obtain the best pairwise of pathway able to classify normal vs breast cancer. The training dataset was 60/100 of the data while the testing 40/100. In this analysis we considered the two classes from TCGA: normal and tumour. The output will be a list of AUC value for each pairwise meeasure of pathway.

```
tumo<-SelectedSample(Dataset=Data_CANCER_normUQ_filt,typesample="tumor")[,1:100]
norm<-SelectedSample(Dataset=Data_CANCER_normUQ_filt,typesample="normal")[,1:100]
nf <- 60
res_class<-svm_classification(TCGA_matrix=score_euc_dist_Lip_met,nfs=nf,
                              normal=colnames(norm),tumour=colnames(tumo))
```

We considered the pairwise of pathways that obtained a performance of AUC major 0.80.

```
better_perf<-select_class(auc.df=res_class,cutoff=0.80)
```

The function `process_matrix` creates a TCGA matrix with the measure of cross-talk previously used, only for the pairwise pathway obtained by `select_class`.

```
matrix_best_perf<-process_matrix(measure=score_euc_dist_Lip_met,list_perf=better_perf)
tumo_bestlipd<-SelectedSample(Dataset=matrix_best_perf,typesample="tumor")[,1:100]
score_bestlipd<-colMeans(tumo_bestlipd)
```

Now we want to create a pathawy cross-talk also for the pathways of cellular processes.

First of all we select the tumour samples and then create a matrix of distance using `dev_std_crtlk`

```
tumo_cell_grow_d<-SelectedSample(Dataset=Data_CANCER_normUQ_filt,typesample="tumor")[,1:100]
score_euc_dist_cell_grow_d<-dev_std_crtlk(dataFilt=tumo_cell_grow_d,pathcell_grow_d)
```

We process the matrix in order to harmonize the structure with `matrix_best_perf`.

```
score__cell_grow_d<-process_matrix_cell_process(score_euc_dist_cell_grow_d)
 score__cell_grow_d_mean<-colMeans(score__cell_grow_d)
```

Now we want to see if there is a correlation among cellular processes and the lipid metabolism in breast cancer.

```
correlazione<-cor(score__cell_grow_d_mean,score_bestlipd)
plot_matrix<-cbind(score__cell_grow_d_mean,score_bestlipd)
```

References

Cava C, Colaprico A, Bertoli G, Bontempi G, Mauri G, Castiglioni I. How interacting pathways are regulated by miRNAs in breast cancer subtypes. BMC Bioinformatics. 2016. In Press.

Colaprico A, Cava C, Bertoli G, Bontempi G, Castiglioni I. Integrative Analysis with Monte Carlo Cross-Validation Reveals miRNAs Regulating Pathways Cross-Talk in Aggressive Breast Cancer. Biomed Res Int. 2015;2015:831314.

Cava, C., Bertoli, G., & Castiglioni, I. (2014, August). Pathway-based expression profile for breast cancer diagnoses. In 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 1151-1154). IEEE.