About pcxn

Sokratis Kariotis, Yered Pita-Juarez, Winston Hide, Wenbin Wei

October 27, 2020

Contents

1 Description

Discover the correlated pathways/gene sets of a single pathway/gene set or discover correlation relationships among multiple pathways/gene sets. Draw a heatmap or create a network of your query and extract members of each pathway/gene set found in the available collections (MSigDB H hallmark, MSigDB C2 Canonical pathways, MSigDB C5 GO BP and Pathprint).

2 Using pcxn wrapper functions to query and visualize the data

The data from the pcxnData package can be utilized using the following functions

- Select a single pathway/gene set from one of the four collections (MSigDB H hallmark gene sets, MSigDB C2 Canonical pathways, MSigDB C5 GO BP gene sets, and Pathprint) and discover its correlated pathway/gene sets within the same collection.

- Discover correlation relationships among multiple pathways/gene sets identified by GSEA (gene set enrichment analysis). All the input pathways/gene sets should come from the same collection. MSigDB H hallmark gene sets, MSigDB C2 Canonical pathways, MSigDB C5 GO BP gene sets, and Pathprint are treated as four separate collections.

- Acquire the gene members of one of the available pathways.

- Draw a heatmap of a pcxn object.

- Create a network of a pcxn object.

2.1   Explore using a pathway or gene set

Each pathway collection (Pathprint, MSigDB H, MSigDB C2 CP and MSigDB C5 GO BP) can be explored using a single pathway of interest as a query. Alternatively, the relationships between groups of pathways, shown to be enriched by gene set enrichment, can be analysed. Both queries can be refined by choosing the number of top correlated gene sets, the minimum absolute correlation and maximum p-value allowed.

```
> # Select a query gene set from a specific collection while requesting
> # the 10 most correlated neighbors, whether the correlation coefficients are
> # adjusted for gene overlap and specific cut-offs for minimum absolute
> # correlation and maximum p-value
>
> library(pcxn)
> library(pcxnData)
> data(list = c("pathprint.Hs.gs",
+                "pathCor_pathprint_v1.2.3_dframe",
+                "pathCor_pathprint_v1.2.3_unadjusted_dframe",
+                "pathCor_CPv5.1_dframe", "gobp_gs_v5.1"))
> pcxn.obj <- pcxn_explore(collection = "pathprint",
+                          query_geneset = "Alzheimer's disease (KEGG)",
+                          adj_overlap = TRUE,
+                          top = 10,
+                          min_abs_corr = 0.05,
+                          max_pval = 0.05)
```

```
[1] "Successful exploring: Based on Alzheimer's disease (KEGG) and 10 top correlated genesets, 27
```

```
> # Select one or two pathway groups from a specific collection while requesting
> # the 10 most correlated neighbors, whether the correlation coefficients are
> # adjusted for gene overlap and specific cut-offs for minimum absolute
> # correlation and maximum p-value
>
> pcxn.obj <- pcxn_analyze("pathprint",c("ABC transporters (KEGG)",
+                                        "ACE Inhibitor Pathway (Wikipathways)",
+                                        "AR down reg. targets (Netpath)"),
+                                        c("DNA Repair (Reactome)"),
+                                        adj_overlap = FALSE, 10, 0.05, 0.05)
```

```
[1] "Successful analyzing: Based on phenotype 0 [ABC transporters (KEGG), ACE Inhibitor Pathway (W
```

2.2 Acquire gene members

Acquire the list of gene members that belong to any of the pathways or gene sets found in the pcxnData package

```
> # Use the pcxn package to select pathway from any collection and get it's
> # gene members as a list
>
> gene_member_list <- pcxn_gene_members("Alzheimer's disease (KEGG)")
```

2.3 Draw a heatmap of a pcxn object

Draw a heatmap of a pcxn object. A number of clustering methods are available. Color represents correlation coefficients.

```
> # Creare the pcxn object needed as an input
> pcxn.obj <- pcxn_analyze("pathprint",c("ABC transporters (KEGG)",
+                                        "ACE Inhibitor Pathway (Wikipathways)",
+                                        "AR down reg. targets (Netpath)"),
+                              c("DNA Repair (Reactome)"), 10, 0.05, 0.05)
[1] "Successful analyzing: Based on phenotype 0 [ABC transporters (KEGG), ACE Inhibitor Pathway (W

> # Draw the heatmap
> heatmap <- pcxn_heatmap(pcxn.obj, "complete")
```

2.4 Create a network of a pcxn object

Create a network visualisation of a pcxn object.

```
> # Create the pcxn object needed as an input
> pcxn.obj <- pcxn_analyze("pathprint",c("ABC transporters (KEGG)",
+                                        "ACE Inhibitor Pathway (Wikipathways)",
+                                        "AR down reg. targets (Netpath)"),
+                              c("DNA Repair (Reactome)"), 10, 0.05, 0.05)
[1] "Successful analyzing: Based on phenotype 0 [ABC transporters (KEGG), ACE Inhibitor Pathway (W

>
> # Create the network
> # network <- pcxn_network(pcxn.obj)
```