Chapter 1 - Quickstart Guide

Gabriel Odom

2018-04-18

Table of Contents

[Overview 1](#_Toc511835002)

[Import Data 2](#_Toc511835003)

[Import .gmt Files 2](#_Toc511835004)

[Import and Tidy Assay Data 2](#_Toc511835005)

[Import Patient Info 2](#_Toc511835006)

[Create an Omics Data Object 3](#_Toc511835007)

[Create an Object 3](#_Toc511835008)

[Inspect the Object Contents 3](#_Toc511835009)

[Modify Object Contents 3](#_Toc511835010)

[Test Pathways for Significance 3](#_Toc511835011)

[AES-PCA 3](#_Toc511835012)

[Supervised PCA 3](#_Toc511835013)

[Inspect Results 3](#_Toc511835014)

[Graph of Top Pathways 4](#_Toc511835015)

[Extract Genes from the Top Pathways 4](#_Toc511835016)

# Overview

This guide will serve as a brief introduction to pathway significance testing with the pathwayPCA package. We will discuss four steps. For detailed discussion of these four steps, see the appropriate vignettes. This vignette is the introduction and overview chapter in the “Pathway Significance Testing with pathwayPCA” workflow.

1. Download Packages
2. Import Data ([*vignette*](https://gabrielodom.github.io/pathwayPCA/articles/C2-Importing_Data.html))
3. Create Omics Data Object ([*vignette*](https://gabrielodom.github.io/pathwayPCA/articles/C3-Create_Omics_Objects.html))
4. Test Pathway Significance ([*vignette*](https://gabrielodom.github.io/pathwayPCA/articles/C4-Methods_Walkthrough.html))
5. Inspect Results

Before we get started, you need the pathwayPCA package to run your analysis. You can install it from GitHub. If you want your analysis to be performed with parallel computing, you will need a package to help you. We recommend the parallel package. We also recommend the tidyverse package to help you run some of the examples in these vignettes.

devtools::install\_github("gabrielodom/pathwayPCA")

library(pathwayPCA)  
library(tidyverse)  
#> -- Attaching packages ------------------------------------------------------------------------------------------------------------------- tidyverse 1.2.1 --  
#> v ggplot2 2.2.1 v purrr 0.2.4  
#> v tibble 1.4.2 v dplyr 0.7.4  
#> v tidyr 0.8.0 v stringr 1.3.0  
#> v readr 1.1.1 v forcats 0.3.0  
#> -- Conflicts ---------------------------------------------------------------------------------------------------------------------- tidyverse\_conflicts() --  
#> x dplyr::filter() masks stats::filter()  
#> x dplyr::lag() masks stats::lag()  
library(parallel)

# Import Data

This section is a quick overview of the material covered in the [Import and Tidy Data](https://gabrielodom.github.io/pathwayPCA/articles/C2-Importing_Data.html) vignette. We will cover three data import steps.

### Import .gmt Files

Use the read\_gmt function. Create your own with the create\_pathwaySet function.

### Import and Tidy Assay Data

Use the read\_csv function from the readr package. Tidy the assay with the transpose\_assay function.

### Import Patient Info

Use the read\_csv function to import the patient data, then the inner\_join function from the dplyr package to match the assay measurements to patient information by subject identifier.

# Create an Omics Data Object

This section is a quick overview of the material covered in the [Create an Omics\* Data Container](https://gabrielodom.github.io/pathwayPCA/articles/C3-Create_Omics_Objects.html) vignette.

### Create an Object

Create a data container specific to survival, regression, or categorical responses.

### Inspect the Object Contents

Print the object to the screen to see a summary of the data contained therein.

### Modify Object Contents

Use accessor functions to extract, edit, or replace data contained in the object.

# Test Pathways for Significance

This section is a quick overview of the material covered in the “AES-PCA” and “Supervised PCA” sections of the [Test Pathway Significance](https://gabrielodom.github.io/pathwayPCA/articles/C4-Methods_Walkthrough.html) vignette.

### AES-PCA

Perform AES-PCA on the object with the AESPCA\_pVals function.

### Supervised PCA

Perform Supervised PCA on the object with the superPCA\_pVals function.

# Inspect Results

This section is a quick overview of the material covered in the “Analyze the Results” section of the [Test Pathway Significance](https://gabrielodom.github.io/pathwayPCA/articles/C4-Methods_Walkthrough.html) vignette.

### Graph of Top Pathways

Use ggplot2 to create summary graphics of the analysis results.

### Extract Genes from the Top Pathways

Use the topGenes function to “score” the genes contained in the top significant pathways.