

Using MBatch Assessments: PCA_DualBatch_Structures

Tod Casasent

2018-06-21

1 Introduction

These instructions are aimed at people familiar with R and familiar with TCGA/GDC platforms and data types. They are intended to introduce the reader to producing the given assessment. These instructions will only rarely, if ever, touch on the appropriateness of the assessment algorithm or interpretation of output. See MBatch_01_InstallLinux for instructions on downloading test data.

2 Algorithm

PCA_Regular_Structures is a function used to perform batch effects assessments using the PCA-Plus algorithm, which adds centroids to the class PCA diagram. PCA-Plus is performed on each batch type available by default.

3 Output

The primary output method for MBatch is to view results in the Batch Effects Website, described elsewhere. The PNG files are rough versions of the website output.

Graphical output is a PCA-Plus plot, which is a PCA diagram with centroids marked and a DSC value. For information, see <http://bioinformatics.mdanderson.org/main/TCGABatchEffects:Overview>.

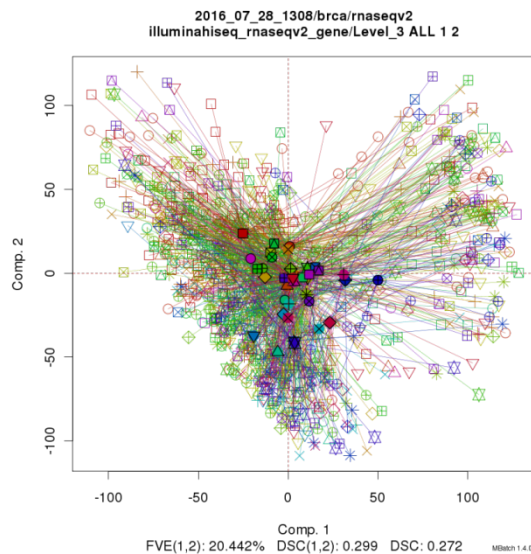


Figure 1: PCAPlus Example

4 Usage

`PCA_DualBatch_Structures(theData, theTitle, theOutputPath, theBatchTypeAndValuePairsToRemove, theBatchTypeAndValuePairsToKeep, theListForDoCentroidDualBatchType, theIsPcaTrendFunction=function(...) {FALSE}, theDoDSCFlag=TRUE, theDoDscPermsFileFlag=FALSE, theDoSampleLocatorFlag=TRUE, theListOfComponentsToPlot=c(1, 2, 1, 3, 1, 4, 2, 3, 2, 4, 3, 4), theDSCPermutations=2000, theDSCThreads=1, theMinBatchSize=2, theJavaParameters="-Xms1200m", theSeed=NULL, theMaxGeneCount=20000)`

5 Arguments

5.1 theData

An instance of `BEA_DATA`.

`BEA_DATA` objects can be created by calls of the form `new("BEA_DATA", theData, theBatches, theCovariates)`. If you have no covariate data, use an empty data.frame created with `data.frame()`

`mData`: Object of class "matrix" A matrix where the colnames are sample ids and the rownames are gene equivalents. All names should be strings, not factors.

`mBatches`: Object of class "data.frame" A data.frame where the column "names" are batch types. The first batch "type" is "Sample". All names and values should be strings, not factors or numeric.

`mCovariates`: Object of class "data.frame" A data.frame where the column "names" are covariate types. The first covariate "type" is "Sample". All names and values should be strings, not factors or numeric.

5.2 theTitle

A string title to use in PNG files.

5.3 theOutputPath

String giving directory in which to place output PNG files.

5.4 theBatchTypeAndValuePairsToRemove

A list of vectors containing the batch type (or * for all types) and the value to remove. `list()` indicates none while `NULL` will cause an error.

5.5 theBatchTypeAndValuePairsToKeep

A list of vectors containing the batch type (or * for all types) and a vector of the the value(s) to keep. `list()` indicates none while `NULL` will cause an error.

5.6 theListForDoCentroidDualBatchType

A vector of strings. Each pair of strings should be batch types and are used to generate dual-batch plots.

5.7 theIsPcaTrendFunction

This function is used by the PCA Centroids algorithm to determine if trending lines should be added. If the function returns true, a line is drawn between the centroids, using the string sorted order. This means, for date or time fields, the batch ids should be formatted such that when sorted as strings, the batch ids are in chronological order from oldest to newest. The function receives arguments theBatchTypeName (a string identifying the batch type) and theListOfBatchIds (a list of strings, which are the batch ids). Default to returning false.

In this sample function, it checks if the batch type is ShipDate. `isTrendBatch<-function(theBatchTypeName, theListOfBatchIds) { return(is.element(theBatchTypeName, c("ShipDate"))) }`

5.8 theDoDSCFlag

A flag indicating whether or not to perform the DSC calculation for PCA-Plus plots. TRUE means perform DSC calculation.

5.9 theDoDscPermsFileFlag

A flag indicating whether or not to create the ANY_DSCPerms.txt and ANY_CompX_CompY_DSCPerms.txt files, which contains contain DSC, Db, Dw, a list of DSC, list of Dw, and list of Db for each permutation. File format consists of six lines per permutation, each line of the format "...". The labels provided are DSC, Db, Dw, DSCList, DwList and DbList.

5.10 theDoSampleLocatorFlag

A flag indicating whether or not to create the ANY_Scores.txt file. File format tab delimited with a header line of "Component Number" followed by sample ids. The first column of subsequent rows has PCA N with N replaced with a numeric. Other cells contain the PCA scores.

5.11 theListOfComponentsToPlot

A vector of integers. Each pair of integers is used to create a PCA plot of those two components. Defaults to `c(1, 2, 1, 3, 1, 4, 2, 3, 2, 4, 3, 4)`

5.12 theDSCPermutations

The number of permutations to perform for DSC calculations. Defaults to 2000

5.13 theDSCThreads

The number of threads (or CPUs or cores) to use to perform DSC calculations. Defaults to 1. Generally limit this to at most the number available minus 1, particularly on HPC systems.

5.14 theMinBatchSize

Minimum batch to allow in PCA calculations. Defaults to 2.

5.15 theJavaParameters

Command line arguments to Java for DSC calculations—for example if memory needs to be more than “-Xms1200m”.

5.16 theSeed

If non-zero, a seed to use when calling sample to select genes used in PCA plots. Default 0.

5.17 theMaxGeneCount

If non-zero, maximum number of genes allowed for PCA plots. If count is greater than this, sample is used to select a subset of genes. Default 20,000.

6 Example Call

If calling this function more than once, either change the output directory or call the clearDSCOverviewFiles function, to clear temporary RData files.

The following code is adapted from the tests/PCA_DualBatch_Structures.R file. Data used is from the testing data as per the MBatch_01_InstallLinux document. In the future, we plan to make the output from MBatch more user friendly, but currently, this produces the following output at the command line.

This output can generally be skipped as very long and generally obscure. After the output is an explanation of files and directories created.

```
{
  library(MBatch)

  # set the paths
  theGeneFile="/bea_testing/MATRIX_DATA/matrix_data-Tumor.tsv"
  theBatchFile="/bea_testing/MATRIX_DATA/batches-Tumor.tsv"
  theOutputDir="/bea_testing/output/PCA_DualBatch_Structures"
  theRandomSeed=314

  # make sure the output dir exists and is empty
  unlink(theOutputDir, recursive=TRUE)
  dir.create(theOutputDir, showWarnings=FALSE, recursive=TRUE)
  # load data
  myData <- mbatchLoadFiles(theGeneFile, theBatchFile)
  myData@mData <- mbatchTrimData(myData@mData, 100000)
  # trend function to handle time data
  isTrendBatch<-function(theBatchTypeName, theListOfBatchIds)
  {
    return(is.element(theBatchTypeName, c("ShipDate")))
  }
  ###
  # do two plots PlateId versus TSS and BatchId versus TSS
  PCA_DualBatch_Structures(myData,
    "Example Title for Test Data", theOutputDir, list(), list(),
    theIsPcaTrendFunction=isTrendBatch,
    theListForDoCentroidDualBatchType=c("PlateId", "TSS", "BatchId", "TSS"),
```

```

    theDoDSCFlag=TRUE, theDoDscPermsFileFlag=TRUE, theDoSampleLocatorFlag=TRUE,
    theListOfComponentsToPlot=c(1, 2, 1, 3), theDSCPermutations=100,
    theDSCThreads=5, theMinBatchSize=0,
    theJavaParameters="-Xms1000m", theSeed=0, theMaxGeneCount=0)
print(dir(theOutputDir, recursive=TRUE))
}

```

```

## 2018 06 21 10:30:26.751 DEBUG megazone23 Changing LC_COLLATE to C for duration of run
## 2018 06 21 10:30:26.752 INFO megazone23 \ / \ / \ / \ / \ / \ / \ / \ /
## 2018 06 21 10:30:26.752 INFO megazone23 Starting mbatchLoadFiles
## 2018 06 21 10:30:26.752 INFO megazone23 MBatch Version: 2017-09-19-1530
## 2018 06 21 10:30:26.753 INFO megazone23 read batch file= /bea_testing/MATRIX_DATA/batches-Tumor.tsv
## 2018 06 21 10:30:26.754 INFO megazone23 read gene file= /bea_testing/MATRIX_DATA/matrix_data-Tumor.t
## 2018 06 21 10:30:36.795 INFO megazone23 filter samples in batches using gene samples
## 2018 06 21 10:30:36.797 INFO megazone23 sort batches by gene file samples
## 2018 06 21 10:30:36.885 INFO megazone23 Finishing mbatchLoadFiles
## 2018 06 21 10:30:36.886 INFO megazone23 ~~~~~
## 2018 06 21 10:30:36.887 DEBUG megazone23 Changing LC_COLLATE to C for duration of run
## 2018 06 21 10:30:36.887 INFO megazone23 \ / \ / \ / \ / \ / \ / \ / \ /
## 2018 06 21 10:30:36.888 INFO megazone23 mbatchTrimData Starting
## 2018 06 21 10:30:36.888 INFO megazone23 MBatch Version: 2017-09-19-1530
## 2018 06 21 10:30:44.775 INFO megazone23 mbatchTrimData Finishing
## 2018 06 21 10:30:44.776 INFO megazone23 ~~~~~
## 2018 06 21 10:30:44.777 DEBUG megazone23 Changing LC_COLLATE to C for duration of run
## 2018 06 21 10:30:44.777 INFO megazone23 \ / \ / \ / \ / \ / \ / \ / \ /
## 2018 06 21 10:30:44.777 INFO megazone23 mbatchFilterData Starting
## 2018 06 21 10:30:44.777 INFO megazone23 MBatch Version: 2017-09-19-1530
## 2018 06 21 10:30:44.778 DEBUG megazone23 rows pre filter 1250
## 2018 06 21 10:30:45.044 DEBUG megazone23 rows post filter 1250
## 2018 06 21 10:30:45.044 DEBUG megazone23 mbatchFilterData Prefilter, gene data had 1250 while post
## 2018 06 21 10:30:45.045 DEBUG megazone23 mbatchFilterData Prefilter, batch data had 80 while post
## 2018 06 21 10:30:45.046 INFO megazone23 mbatchFilterData Finishing
## 2018 06 21 10:30:45.046 INFO megazone23 ~~~~~
## 2018 06 21 10:30:45.048 DEBUG megazone23 Calling .jinit for PCA
## 2018 06 21 10:30:45.048 DEBUG megazone23 classpath=, /home/linux/R/x86_64-pc-linux-gnu-library/3.4/M
## 2018 06 21 10:30:45.327 DEBUG megazone23 Changing LC_COLLATE to C for duration of run
## 2018 06 21 10:30:45.328 DEBUG megazone23 createBatchEffectsOutput_pca_dualBatch - start loop
## 2018 06 21 10:30:45.329 DEBUG megazone23 doSamplePcaCall - pre-filter matrix size= 1250,80
## 2018 06 21 10:30:45.329 DEBUG megazone23 filterBasedOnGeneLimit theMatrixGeneData nrow= 1250
## 2018 06 21 10:30:45.330 DEBUG megazone23 filterBasedOnGeneLimit theMatrixGeneData ncol= 80
## 2018 06 21 10:30:45.330 DEBUG megazone23 doSamplePcaCall - pre-filter matrix size= 1250,80
## 2018 06 21 10:30:45.330 DEBUG megazone23 doSamplePcaCall - no-NA matrix length= 1247
## 2018 06 21 10:30:45.377 DEBUG megazone23 doSamplePcaCall - max(theCompPairList)= 3
## 2018 06 21 10:30:45.377 DEBUG megazone23 doSamplePcaCall - pca scores size= 80,80
## 2018 06 21 10:30:45.378 DEBUG megazone23 checkCreateDir: /bea_testing/output/PCA_DualBatch_Structure
## 2018 06 21 10:30:45.378 DEBUG megazone23 checkCreateDir: /bea_testing/output/PCA_DualBatch_Structure
## 2018 06 21 10:30:45.379 DEBUG megazone23 pvalueDSC start
## 2018 06 21 10:30:45.379 DEBUG megazone23 pvalueDSC component A < 1
## 2018 06 21 10:30:45.379 DEBUG megazone23 pvalueDSC call pvalueDSCwithExcerpt
## 2018 06 21 10:30:45.380 DEBUG megazone23 pvalueDSCwithExcerpt start
## 2018 06 21 10:30:45.380 DEBUG megazone23 nrow(thePcaDataExcerpt)= 80
## 2018 06 21 10:30:45.380 DEBUG megazone23 ncol(thePcaDataExcerpt)= 80
## 2018 06 21 10:30:45.380 DEBUG megazone23 length(theBatchIdsForSamples)= 80
## 2018 06 21 10:30:45.381 DEBUG megazone23 getDSCwithExcerpt before java

```

```

## 2018 06 21 10:30:45.384 DEBUG megazone23 getDSCwithExcerpt after java
## 2018 06 21 10:30:45.385 DEBUG megazone23 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2018 06 21 10:30:45.386 DEBUG megazone23 pvalueDSCwithExcerpt length(unique(theBatchIdsForSamples))=
## 2018 06 21 10:30:45.387 DEBUG megazone23 pvalueDSCwithExcerpt 2222 length(permResults)= 1
## 2018 06 21 10:30:45.387 DEBUG megazone23 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2018 06 21 10:30:45.396 DEBUG megazone23 openAndWriteDscPermsFile -- start
## 2018 06 21 10:30:45.397 DEBUG megazone23 openAndWriteDscPermsFile -- redo name
## 2018 06 21 10:30:45.397 DEBUG megazone23 openAndWriteDscPermsFile -- check file
## 2018 06 21 10:30:45.397 DEBUG megazone23 openAndWriteDscPermsFile -- no file, write
## 2018 06 21 10:30:45.398 DEBUG megazone23 openAndWriteDscPermsFile -- nothing to write
## 2018 06 21 10:30:45.398 DEBUG megazone23 write writePcaDataFilesForDataset
## 2018 06 21 10:30:45.399 DEBUG megazone23 checkCreateDir: /bea_testing/output/PCA_DualBatch_Structur
## 2018 06 21 10:30:45.399 DEBUG megazone23 writePCAValuesTSV
## 2018 06 21 10:30:45.399 DEBUG megazone23 writePCAValuesTSV outputFile
## 2018 06 21 10:30:45.400 DEBUG megazone23 writePCAValuesTSV components
## 2018 06 21 10:30:45.400 DEBUG megazone23 writePCAValuesTSV myDataFrame
## 2018 06 21 10:30:45.400 DEBUG megazone23 writePCAValuesTSV - values
## 2018 06 21 10:30:45.401 DEBUG megazone23 writePCAValuesTSV - matrix
## 2018 06 21 10:30:45.402 DEBUG megazone23 writePCAValuesTSV - write
## 2018 06 21 10:30:45.404 DEBUG megazone23 writePCAValuesTSV finished /bea_testing/output/PCA_DualBat
## 2018 06 21 10:30:45.404 DEBUG megazone23 writePCAAnnotations
## 2018 06 21 10:30:45.405 DEBUG megazone23 writePCAAnnotations - header and all data
## 2018 06 21 10:30:45.684 DEBUG megazone23 doInternalPca before filter - Number of rownames for pca sco
## 2018 06 21 10:30:45.684 DEBUG megazone23 doInternalPca before filter - length of batch ids for sampl
## 2018 06 21 10:30:45.684 DEBUG megazone23 doInternalPca after filter - Number of rownames for pca sco
## 2018 06 21 10:30:45.684 DEBUG megazone23 doInternalPca after filter - length of batch ids for sampl
## 2018 06 21 10:30:45.685 DEBUG megazone23 doInternalPca - pvalueDSC
## 2018 06 21 10:30:45.685 DEBUG megazone23 pvalueDSC start
## 2018 06 21 10:30:45.686 DEBUG megazone23 pvalueDSC transpose
## 2018 06 21 10:30:45.686 DEBUG megazone23 pvalueDSC call pvalueDSCwithExcerpt
## 2018 06 21 10:30:45.687 DEBUG megazone23 pvalueDSCwithExcerpt start
## 2018 06 21 10:30:45.687 DEBUG megazone23 nrow(thePcaDataExcerpt)= 2
## 2018 06 21 10:30:45.687 DEBUG megazone23 ncol(thePcaDataExcerpt)= 80
## 2018 06 21 10:30:45.688 DEBUG megazone23 length(theBatchIdsForSamples)= 80
## 2018 06 21 10:30:45.689 DEBUG megazone23 getDSCwithExcerpt before java
## 2018 06 21 10:30:45.691 DEBUG megazone23 getDSCwithExcerpt after java
## 2018 06 21 10:30:45.692 DEBUG megazone23 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2018 06 21 10:30:45.692 DEBUG megazone23 pvalueDSCwithExcerpt length(unique(theBatchIdsForSamples))=
## 2018 06 21 10:30:45.692 DEBUG megazone23 pvalueDSCwithExcerpt 2222 length(permResults)= 1
## 2018 06 21 10:30:45.693 DEBUG megazone23 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2018 06 21 10:30:45.693 DEBUG megazone23 openAndWriteDscPermsFile -- start
## 2018 06 21 10:30:45.693 DEBUG megazone23 openAndWriteDscPermsFile -- redo name
## 2018 06 21 10:30:45.694 DEBUG megazone23 openAndWriteDscPermsFile -- check file
## 2018 06 21 10:30:45.694 DEBUG megazone23 openAndWriteDscPermsFile -- no file, write
## 2018 06 21 10:30:45.694 DEBUG megazone23 openAndWriteDscPermsFile -- nothing to write
## 2018 06 21 10:30:45.695 DEBUG megazone23 writePCAAnnotations
## 2018 06 21 10:30:45.696 DEBUG megazone23 writePCAAnnotations - data PC1,PC2
## 2018 06 21 10:30:45.699 DEBUG megazone23 writePcaAnalysis_Image - Number of rownames for pca score (
## 2018 06 21 10:30:45.699 DEBUG megazone23 writePcaAnalysis_Image - length of batch ids for samples 8
## 2018 06 21 10:30:45.783 DEBUG megazone23 mbatchStandardLegend - Calling .jinit /home/linux/R/x86_64
## 2018 06 21 10:30:46.060 DEBUG megazone23 mbatchStandardLegend - .jinit complete
## 2018 06 21 10:30:46.061 DEBUG megazone23 mbatchStandardLegend - theTitle Rays: PlateId (80)
## 2018 06 21 10:30:46.061 DEBUG megazone23 mbatchStandardLegend - theVersion MBatch 1.4.16

```

```

## 2018 06 21 10:30:46.062 DEBUG megazone23 mbatchStandardLegend - theFilenamePath /bea_testing/output,
## 2018 06 21 10:30:46.062 DEBUG megazone23 mbatchStandardLegend - theLegendNames A29J (80)
## 2018 06 21 10:30:46.062 DEBUG megazone23 mbatchStandardLegend - theLegendNames 1
## 2018 06 21 10:30:46.063 DEBUG megazone23 mbatchStandardLegend - theLegendColors 1
## 2018 06 21 10:30:46.063 DEBUG megazone23 mbatchStandardLegend - theLegendSymbols 0
## 2018 06 21 10:30:46.063 DEBUG megazone23 mbatchStandardLegend - myColors #b30000
## 2018 06 21 10:30:46.064 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:46.082 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:46.085 DEBUG megazone23 mbatchStandardLegend - Calling .jinit /home/linux/R/x86_64
## 2018 06 21 10:30:46.361 DEBUG megazone23 mbatchStandardLegend - .jinit complete
## 2018 06 21 10:30:46.362 DEBUG megazone23 mbatchStandardLegend - theTitle Points: TSS (80)
## 2018 06 21 10:30:46.362 DEBUG megazone23 mbatchStandardLegend - theVersion MBatch 1.4.16
## 2018 06 21 10:30:46.363 DEBUG megazone23 mbatchStandardLegend - theFilenamePath /bea_testing/output,
## 2018 06 21 10:30:46.363 DEBUG megazone23 mbatchStandardLegend - theLegendNames OR - University of M
## (72), OU - Roswell Park (1), P6 - Translational Genomics
## Research Institute (2), PA - University of Minnesota
## (1), PK - University Health
## Network (4)
## 2018 06 21 10:30:46.363 DEBUG megazone23 mbatchStandardLegend - theLegendNames 5
## 2018 06 21 10:30:46.364 DEBUG megazone23 mbatchStandardLegend - theLegendColors 5
## 2018 06 21 10:30:46.364 DEBUG megazone23 mbatchStandardLegend - theLegendSymbols 5
## 2018 06 21 10:30:46.364 DEBUG megazone23 mbatchStandardLegend - myColors #b30000,#8fb300,#00b347,#0
## 2018 06 21 10:30:46.365 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:46.432 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:46.435 DEBUG megazone23 mbatchStandardLegend - Calling .jinit /home/linux/R/x86_64
## 2018 06 21 10:30:46.709 DEBUG megazone23 mbatchStandardLegend - .jinit complete
## 2018 06 21 10:30:46.710 DEBUG megazone23 mbatchStandardLegend - theTitle Dispersion Metrics
## 2018 06 21 10:30:46.710 DEBUG megazone23 mbatchStandardLegend - theVersion MBatch 1.4.16
## 2018 06 21 10:30:46.711 DEBUG megazone23 mbatchStandardLegend - theFilenamePath /bea_testing/output,
## 2018 06 21 10:30:46.711 DEBUG megazone23 mbatchStandardLegend - theLegendNames First PCA Component:
## 2018 06 21 10:30:46.712 DEBUG megazone23 mbatchStandardLegend - theLegendNames 14
## 2018 06 21 10:30:46.712 DEBUG megazone23 mbatchStandardLegend - theLegendColors 0
## 2018 06 21 10:30:46.712 DEBUG megazone23 mbatchStandardLegend - theLegendSymbols 0
## 2018 06 21 10:30:46.713 DEBUG megazone23 mbatchStandardLegend - myColors
## 2018 06 21 10:30:46.713 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:46.838 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:46.841 DEBUG megazone23 mbatchStandardCombineLegends - Calling .jinit /home/linux/
## 2018 06 21 10:30:47.109 DEBUG megazone23 mbatchStandardCombineLegends - .jinit complete
## 2018 06 21 10:30:47.109 DEBUG megazone23 mbatchStandardCombineLegends - theTitle Example Title for
## 2018 06 21 10:30:47.109 DEBUG megazone23 mbatchStandardCombineLegends - theFilenamePath /bea_testing
## 2018 06 21 10:30:47.110 DEBUG megazone23 mbatchStandardCombineLegends - theListOfFiles /bea_testing
## 2018 06 21 10:30:47.110 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:47.363 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:47.368 DEBUG megazone23 doInternalPca before filter - Number of rownames for pca sco
## 2018 06 21 10:30:47.369 DEBUG megazone23 doInternalPca before filter - length of batch ids for sample
## 2018 06 21 10:30:47.370 DEBUG megazone23 doInternalPca after filter - Number of rownames for pca sco
## 2018 06 21 10:30:47.370 DEBUG megazone23 doInternalPca after filter - length of batch ids for sample
## 2018 06 21 10:30:47.371 DEBUG megazone23 doInternalPca - pvalueDSC
## 2018 06 21 10:30:47.371 DEBUG megazone23 pvalueDSC start
## 2018 06 21 10:30:47.372 DEBUG megazone23 pvalueDSC transpose
## 2018 06 21 10:30:47.373 DEBUG megazone23 pvalueDSC call pvalueDSCwithExcerpt
## 2018 06 21 10:30:47.373 DEBUG megazone23 pvalueDSCwithExcerpt start
## 2018 06 21 10:30:47.374 DEBUG megazone23 nrow(thePcaDataExcerpt)= 2
## 2018 06 21 10:30:47.374 DEBUG megazone23 ncol(thePcaDataExcerpt)= 80

```

```

## 2018 06 21 10:30:47.375 DEBUG megazone23 length(theBatchIdsForSamples)= 80
## 2018 06 21 10:30:47.375 DEBUG megazone23 getDSCwithExcerpt before java
## 2018 06 21 10:30:47.376 DEBUG megazone23 getDSCwithExcerpt after java
## 2018 06 21 10:30:47.377 DEBUG megazone23 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2018 06 21 10:30:47.378 DEBUG megazone23 pvalueDSCwithExcerpt length(unique(theBatchIdsForSamples))=
## 2018 06 21 10:30:47.378 DEBUG megazone23 pvalueDSCwithExcerpt 2222 length(permResults)= 1
## 2018 06 21 10:30:47.379 DEBUG megazone23 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2018 06 21 10:30:47.379 DEBUG megazone23 openAndWriteDscPermsFile -- start
## 2018 06 21 10:30:47.379 DEBUG megazone23 openAndWriteDscPermsFile -- redo name
## 2018 06 21 10:30:47.380 DEBUG megazone23 openAndWriteDscPermsFile -- check file
## 2018 06 21 10:30:47.380 DEBUG megazone23 openAndWriteDscPermsFile -- no file, write
## 2018 06 21 10:30:47.381 DEBUG megazone23 openAndWriteDscPermsFile -- nothing to write
## 2018 06 21 10:30:47.381 DEBUG megazone23 writePCAAnnotations
## 2018 06 21 10:30:47.382 DEBUG megazone23 writePCAAnnotations - data PC1,PC3
## 2018 06 21 10:30:47.387 DEBUG megazone23 writePcaAnalysis_Image - Number of rownames for pca score (
## 2018 06 21 10:30:47.387 DEBUG megazone23 writePcaAnalysis_Image - length of batch ids for samples 8

## 2018 06 21 10:30:47.473 DEBUG megazone23 mbatchStandardLegend - Calling .jinit /home/linux/R/x86_64
## 2018 06 21 10:30:47.736 DEBUG megazone23 mbatchStandardLegend - .jinit complete
## 2018 06 21 10:30:47.737 DEBUG megazone23 mbatchStandardLegend - theTitle Rays: PlateId (80)
## 2018 06 21 10:30:47.737 DEBUG megazone23 mbatchStandardLegend - theVersion MBatch 1.4.16
## 2018 06 21 10:30:47.738 DEBUG megazone23 mbatchStandardLegend - theFilenamePath /bea_testing/output,
## 2018 06 21 10:30:47.739 DEBUG megazone23 mbatchStandardLegend - theLegendNames A29J (80)
## 2018 06 21 10:30:47.740 DEBUG megazone23 mbatchStandardLegend - theLegendNames 1
## 2018 06 21 10:30:47.740 DEBUG megazone23 mbatchStandardLegend - theLegendColors 1
## 2018 06 21 10:30:47.741 DEBUG megazone23 mbatchStandardLegend - theLegendSymbols 0
## 2018 06 21 10:30:47.741 DEBUG megazone23 mbatchStandardLegend - myColors #b30000
## 2018 06 21 10:30:47.741 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:47.761 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:47.763 DEBUG megazone23 mbatchStandardLegend - Calling .jinit /home/linux/R/x86_64
## 2018 06 21 10:30:48.033 DEBUG megazone23 mbatchStandardLegend - .jinit complete
## 2018 06 21 10:30:48.034 DEBUG megazone23 mbatchStandardLegend - theTitle Points: TSS (80)
## 2018 06 21 10:30:48.034 DEBUG megazone23 mbatchStandardLegend - theVersion MBatch 1.4.16
## 2018 06 21 10:30:48.035 DEBUG megazone23 mbatchStandardLegend - theFilenamePath /bea_testing/output,
## 2018 06 21 10:30:48.036 DEBUG megazone23 mbatchStandardLegend - theLegendNames OR - University of M
## (72), OU - Roswell Park (1), P6 - Translational Genomics
## Research Institute (2), PA - University of Minnesota
## (1), PK - University Health
## Network (4)
## 2018 06 21 10:30:48.036 DEBUG megazone23 mbatchStandardLegend - theLegendNames 5
## 2018 06 21 10:30:48.037 DEBUG megazone23 mbatchStandardLegend - theLegendColors 5
## 2018 06 21 10:30:48.037 DEBUG megazone23 mbatchStandardLegend - theLegendSymbols 5
## 2018 06 21 10:30:48.037 DEBUG megazone23 mbatchStandardLegend - myColors #b30000,#8fb300,#00b347,#0
## 2018 06 21 10:30:48.038 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:48.104 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:48.106 DEBUG megazone23 mbatchStandardLegend - Calling .jinit /home/linux/R/x86_64
## 2018 06 21 10:30:48.394 DEBUG megazone23 mbatchStandardLegend - .jinit complete
## 2018 06 21 10:30:48.395 DEBUG megazone23 mbatchStandardLegend - theTitle Dispersion Metrics
## 2018 06 21 10:30:48.395 DEBUG megazone23 mbatchStandardLegend - theVersion MBatch 1.4.16
## 2018 06 21 10:30:48.396 DEBUG megazone23 mbatchStandardLegend - theFilenamePath /bea_testing/output,
## 2018 06 21 10:30:48.396 DEBUG megazone23 mbatchStandardLegend - theLegendNames First PCA Component:
## 2018 06 21 10:30:48.396 DEBUG megazone23 mbatchStandardLegend - theLegendNames 14
## 2018 06 21 10:30:48.397 DEBUG megazone23 mbatchStandardLegend - theLegendColors 0
## 2018 06 21 10:30:48.397 DEBUG megazone23 mbatchStandardLegend - theLegendSymbols 0

```



```

## 2018 06 21 10:30:48.397 DEBUG megazone23 mbatchStandardLegend - myColors
## 2018 06 21 10:30:48.398 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:48.529 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:48.531 DEBUG megazone23 mbatchStandardCombineLegends - Calling .jinit /home/linux/
## 2018 06 21 10:30:48.815 DEBUG megazone23 mbatchStandardCombineLegends - .jinit complete
## 2018 06 21 10:30:48.815 DEBUG megazone23 mbatchStandardCombineLegends - theTitle Example Title for
## 2018 06 21 10:30:48.815 DEBUG megazone23 mbatchStandardCombineLegends - theFilenamePath /bea_testing
## 2018 06 21 10:30:48.816 DEBUG megazone23 mbatchStandardCombineLegends - theListOfFiles /bea_testing
## 2018 06 21 10:30:48.816 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:49.060 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:49.061 DEBUG megazone23 doSamplePcaCall - pre-filter matrix size= 1250,80
## 2018 06 21 10:30:49.062 DEBUG megazone23 filterBasedOnGeneLimit theMatrixGeneData nrow= 1250
## 2018 06 21 10:30:49.062 DEBUG megazone23 filterBasedOnGeneLimit theMatrixGeneData ncol= 80
## 2018 06 21 10:30:49.063 DEBUG megazone23 doSamplePcaCall - pre-filter matrix size= 1250,80
## 2018 06 21 10:30:49.063 DEBUG megazone23 doSamplePcaCall - no-NA matrix length= 1247
## 2018 06 21 10:30:49.110 DEBUG megazone23 doSamplePcaCall - max(theCompPairList)= 3
## 2018 06 21 10:30:49.110 DEBUG megazone23 doSamplePcaCall - pca scores size= 80,80
## 2018 06 21 10:30:49.111 DEBUG megazone23 checkCreateDir: /bea_testing/output/PCA_DualBatch_Structur
## 2018 06 21 10:30:49.112 DEBUG megazone23 checkCreateDir: /bea_testing/output/PCA_DualBatch_Structur
## 2018 06 21 10:30:49.112 DEBUG megazone23 pvalueDSC start
## 2018 06 21 10:30:49.112 DEBUG megazone23 pvalueDSC component A < 1
## 2018 06 21 10:30:49.113 DEBUG megazone23 pvalueDSC call pvalueDSCwithExcerpt
## 2018 06 21 10:30:49.113 DEBUG megazone23 pvalueDSCwithExcerpt start
## 2018 06 21 10:30:49.113 DEBUG megazone23 nrow(thePcaDataExcerpt)= 80
## 2018 06 21 10:30:49.114 DEBUG megazone23 ncol(thePcaDataExcerpt)= 80
## 2018 06 21 10:30:49.114 DEBUG megazone23 length(theBatchIdsForSamples)= 80
## 2018 06 21 10:30:49.115 DEBUG megazone23 getDSCwithExcerpt before java
## 2018 06 21 10:30:49.116 DEBUG megazone23 getDSCwithExcerpt after java
## 2018 06 21 10:30:49.117 DEBUG megazone23 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2018 06 21 10:30:49.118 DEBUG megazone23 pvalueDSCwithExcerpt length(unique(theBatchIdsForSamples))=
## 2018 06 21 10:30:49.118 DEBUG megazone23 pvalueDSCwithExcerpt 2222 length(permResults)= 1
## 2018 06 21 10:30:49.119 DEBUG megazone23 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2018 06 21 10:30:49.128 DEBUG megazone23 openAndWriteDscPermsFile -- start
## 2018 06 21 10:30:49.128 DEBUG megazone23 openAndWriteDscPermsFile -- redo name
## 2018 06 21 10:30:49.129 DEBUG megazone23 openAndWriteDscPermsFile -- check file
## 2018 06 21 10:30:49.129 DEBUG megazone23 openAndWriteDscPermsFile -- no file, write
## 2018 06 21 10:30:49.129 DEBUG megazone23 openAndWriteDscPermsFile -- nothing to write
## 2018 06 21 10:30:49.130 DEBUG megazone23 write writePcaDataFilesForDataset
## 2018 06 21 10:30:49.130 DEBUG megazone23 checkCreateDir: /bea_testing/output/PCA_DualBatch_Structur
## 2018 06 21 10:30:49.130 DEBUG megazone23 writePCAValuesTSV
## 2018 06 21 10:30:49.131 DEBUG megazone23 writePCAValuesTSV outputFile
## 2018 06 21 10:30:49.131 DEBUG megazone23 writePCAValuesTSV components
## 2018 06 21 10:30:49.131 DEBUG megazone23 writePCAValuesTSV myDataFrame
## 2018 06 21 10:30:49.132 DEBUG megazone23 writePCAValuesTSV - values
## 2018 06 21 10:30:49.132 DEBUG megazone23 writePCAValuesTSV - matrix
## 2018 06 21 10:30:49.133 DEBUG megazone23 writePCAValuesTSV - write
## 2018 06 21 10:30:49.135 DEBUG megazone23 writePCAValuesTSV finished /bea_testing/output/PCA_DualBat
## 2018 06 21 10:30:49.135 DEBUG megazone23 writePCAAnnotations
## 2018 06 21 10:30:49.136 DEBUG megazone23 writePCAAnnotations - header and all data
## 2018 06 21 10:30:49.348 DEBUG megazone23 doInternalPca before filter - Number of rownames for pca sco
## 2018 06 21 10:30:49.348 DEBUG megazone23 doInternalPca before filter - length of batch ids for sampl
## 2018 06 21 10:30:49.349 DEBUG megazone23 doInternalPca after filter - Number of rownames for pca sco
## 2018 06 21 10:30:49.349 DEBUG megazone23 doInternalPca after filter - length of batch ids for sampl
## 2018 06 21 10:30:49.349 DEBUG megazone23 doInternalPca - pvalueDSC

```

```

## 2018 06 21 10:30:49.350 DEBUG megazone23 pvalueDSC start
## 2018 06 21 10:30:49.350 DEBUG megazone23 pvalueDSC transpose
## 2018 06 21 10:30:49.351 DEBUG megazone23 pvalueDSC call pvalueDSCwithExcerpt
## 2018 06 21 10:30:49.351 DEBUG megazone23 pvalueDSCwithExcerpt start
## 2018 06 21 10:30:49.351 DEBUG megazone23 nrow(thePcaDataExcerpt)= 2
## 2018 06 21 10:30:49.351 DEBUG megazone23 ncol(thePcaDataExcerpt)= 80
## 2018 06 21 10:30:49.352 DEBUG megazone23 length(theBatchIdsForSamples)= 80
## 2018 06 21 10:30:49.353 DEBUG megazone23 getDSCwithExcerpt before java
## 2018 06 21 10:30:49.355 DEBUG megazone23 getDSCwithExcerpt after java
## 2018 06 21 10:30:49.356 DEBUG megazone23 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2018 06 21 10:30:49.357 DEBUG megazone23 pvalueDSCwithExcerpt length(unique(theBatchIdsForSamples))=
## 2018 06 21 10:30:49.357 DEBUG megazone23 pvalueDSCwithExcerpt 2222 length(permResults)= 1
## 2018 06 21 10:30:49.358 DEBUG megazone23 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2018 06 21 10:30:49.359 DEBUG megazone23 openAndWriteDscPermsFile -- start
## 2018 06 21 10:30:49.359 DEBUG megazone23 openAndWriteDscPermsFile -- redo name
## 2018 06 21 10:30:49.359 DEBUG megazone23 openAndWriteDscPermsFile -- check file
## 2018 06 21 10:30:49.360 DEBUG megazone23 openAndWriteDscPermsFile -- no file, write
## 2018 06 21 10:30:49.360 DEBUG megazone23 openAndWriteDscPermsFile -- nothing to write
## 2018 06 21 10:30:49.361 DEBUG megazone23 writePCAAnnotations
## 2018 06 21 10:30:49.361 DEBUG megazone23 writePCAAnnotations - data PC1,PC2
## 2018 06 21 10:30:49.365 DEBUG megazone23 writePcaAnalysis_Image - Number of rownames for pca score (
## 2018 06 21 10:30:49.365 DEBUG megazone23 writePcaAnalysis_Image - length of batch ids for samples 8

## 2018 06 21 10:30:49.449 DEBUG megazone23 mbatchStandardLegend - Calling .jinit /home/linux/R/x86_64
## 2018 06 21 10:30:49.709 DEBUG megazone23 mbatchStandardLegend - .jinit complete
## 2018 06 21 10:30:49.709 DEBUG megazone23 mbatchStandardLegend - theTitle Rays: BatchId (80)
## 2018 06 21 10:30:49.709 DEBUG megazone23 mbatchStandardLegend - theVersion MBatch 1.4.16
## 2018 06 21 10:30:49.710 DEBUG megazone23 mbatchStandardLegend - theFilenamePath /bea_testing/output,
## 2018 06 21 10:30:49.711 DEBUG megazone23 mbatchStandardLegend - theLegendNames 00304 (80)
## 2018 06 21 10:30:49.711 DEBUG megazone23 mbatchStandardLegend - theLegendNames 1
## 2018 06 21 10:30:49.711 DEBUG megazone23 mbatchStandardLegend - theLegendColors 1
## 2018 06 21 10:30:49.711 DEBUG megazone23 mbatchStandardLegend - theLegendSymbols 0
## 2018 06 21 10:30:49.712 DEBUG megazone23 mbatchStandardLegend - myColors #b30000
## 2018 06 21 10:30:49.712 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:49.731 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:49.736 DEBUG megazone23 mbatchStandardLegend - Calling .jinit /home/linux/R/x86_64
## 2018 06 21 10:30:50.001 DEBUG megazone23 mbatchStandardLegend - .jinit complete
## 2018 06 21 10:30:50.001 DEBUG megazone23 mbatchStandardLegend - theTitle Points: TSS (80)
## 2018 06 21 10:30:50.002 DEBUG megazone23 mbatchStandardLegend - theVersion MBatch 1.4.16
## 2018 06 21 10:30:50.003 DEBUG megazone23 mbatchStandardLegend - theFilenamePath /bea_testing/output,
## 2018 06 21 10:30:50.004 DEBUG megazone23 mbatchStandardLegend - theLegendNames OR - University of M
## (72), OU - Roswell Park (1), P6 - Translational Genomics
## Research Institute (2), PA - University of Minnesota
## (1), PK - University Health
## Network (4)
## 2018 06 21 10:30:50.004 DEBUG megazone23 mbatchStandardLegend - theLegendNames 5
## 2018 06 21 10:30:50.005 DEBUG megazone23 mbatchStandardLegend - theLegendColors 5
## 2018 06 21 10:30:50.005 DEBUG megazone23 mbatchStandardLegend - theLegendSymbols 5
## 2018 06 21 10:30:50.006 DEBUG megazone23 mbatchStandardLegend - myColors #b30000,#8fb300,#00b347,#0
## 2018 06 21 10:30:50.007 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:50.072 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:50.075 DEBUG megazone23 mbatchStandardLegend - Calling .jinit /home/linux/R/x86_64
## 2018 06 21 10:30:50.362 DEBUG megazone23 mbatchStandardLegend - .jinit complete
## 2018 06 21 10:30:50.363 DEBUG megazone23 mbatchStandardLegend - theTitle Dispersion Metrics

```

```

## 2018 06 21 10:30:50.363 DEBUG megazone23 mbatchStandardLegend - theVersion MBatch 1.4.16
## 2018 06 21 10:30:50.364 DEBUG megazone23 mbatchStandardLegend - theFilenamePath /bea_testing/output
## 2018 06 21 10:30:50.364 DEBUG megazone23 mbatchStandardLegend - theLegendNames First PCA Component:
## 2018 06 21 10:30:50.365 DEBUG megazone23 mbatchStandardLegend - theLegendNames 14
## 2018 06 21 10:30:50.365 DEBUG megazone23 mbatchStandardLegend - theLegendColors 0
## 2018 06 21 10:30:50.366 DEBUG megazone23 mbatchStandardLegend - theLegendSymbols 0
## 2018 06 21 10:30:50.366 DEBUG megazone23 mbatchStandardLegend - myColors
## 2018 06 21 10:30:50.366 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:50.499 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:50.501 DEBUG megazone23 mbatchStandardCombineLegends - Calling .jinit /home/linux/
## 2018 06 21 10:30:50.775 DEBUG megazone23 mbatchStandardCombineLegends - .jinit complete
## 2018 06 21 10:30:50.776 DEBUG megazone23 mbatchStandardCombineLegends - theTitle Example Title for '
## 2018 06 21 10:30:50.776 DEBUG megazone23 mbatchStandardCombineLegends - theFilenamePath /bea_testing
## 2018 06 21 10:30:50.776 DEBUG megazone23 mbatchStandardCombineLegends - theListOfFiles /bea_testing
## 2018 06 21 10:30:50.777 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:51.022 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:51.087 DEBUG megazone23 doInternalPca before filter - Number of rownames for pca sco
## 2018 06 21 10:30:51.091 DEBUG megazone23 doInternalPca before filter - length of batch ids for sampl
## 2018 06 21 10:30:51.092 DEBUG megazone23 doInternalPca after filter - Number of rownames for pca sco
## 2018 06 21 10:30:51.093 DEBUG megazone23 doInternalPca after filter - length of batch ids for sampl
## 2018 06 21 10:30:51.094 DEBUG megazone23 doInternalPca - pvalueDSC
## 2018 06 21 10:30:51.096 DEBUG megazone23 pvalueDSC start
## 2018 06 21 10:30:51.097 DEBUG megazone23 pvalueDSC transpose
## 2018 06 21 10:30:51.097 DEBUG megazone23 pvalueDSC call pvalueDSCwithExcerpt
## 2018 06 21 10:30:51.098 DEBUG megazone23 pvalueDSCwithExcerpt start
## 2018 06 21 10:30:51.099 DEBUG megazone23 nrow(thePcaDataExcerpt)= 2
## 2018 06 21 10:30:51.100 DEBUG megazone23 ncol(thePcaDataExcerpt)= 80
## 2018 06 21 10:30:51.100 DEBUG megazone23 length(theBatchIdsForSamples)= 80
## 2018 06 21 10:30:51.102 DEBUG megazone23 getDSCwithExcerpt before java
## 2018 06 21 10:30:51.105 DEBUG megazone23 getDSCwithExcerpt after java
## 2018 06 21 10:30:51.107 DEBUG megazone23 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2018 06 21 10:30:51.108 DEBUG megazone23 pvalueDSCwithExcerpt length(unique(theBatchIdsForSamples))=
## 2018 06 21 10:30:51.108 DEBUG megazone23 pvalueDSCwithExcerpt 2222 length(permutationResults)= 1
## 2018 06 21 10:30:51.109 DEBUG megazone23 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2018 06 21 10:30:51.109 DEBUG megazone23 openAndWriteDscPermsFile -- start
## 2018 06 21 10:30:51.110 DEBUG megazone23 openAndWriteDscPermsFile -- redo name
## 2018 06 21 10:30:51.110 DEBUG megazone23 openAndWriteDscPermsFile -- check file
## 2018 06 21 10:30:51.110 DEBUG megazone23 openAndWriteDscPermsFile -- no file, write
## 2018 06 21 10:30:51.111 DEBUG megazone23 openAndWriteDscPermsFile -- nothing to write
## 2018 06 21 10:30:51.111 DEBUG megazone23 writePCAAnnotations
## 2018 06 21 10:30:51.112 DEBUG megazone23 writePCAAnnotations - data PC1,PC3
## 2018 06 21 10:30:51.116 DEBUG megazone23 writePcaAnalysis_Image - Number of rownames for pca score (
## 2018 06 21 10:30:51.116 DEBUG megazone23 writePcaAnalysis_Image - length of batch ids for samples 80

## 2018 06 21 10:30:51.200 DEBUG megazone23 mbatchStandardLegend - Calling .jinit /home/linux/R/x86_64
## 2018 06 21 10:30:51.471 DEBUG megazone23 mbatchStandardLegend - .jinit complete
## 2018 06 21 10:30:51.472 DEBUG megazone23 mbatchStandardLegend - theTitle Rays: BatchId (80)
## 2018 06 21 10:30:51.472 DEBUG megazone23 mbatchStandardLegend - theVersion MBatch 1.4.16
## 2018 06 21 10:30:51.473 DEBUG megazone23 mbatchStandardLegend - theFilenamePath /bea_testing/output
## 2018 06 21 10:30:51.473 DEBUG megazone23 mbatchStandardLegend - theLegendNames 00304 (80)
## 2018 06 21 10:30:51.474 DEBUG megazone23 mbatchStandardLegend - theLegendNames 1
## 2018 06 21 10:30:51.474 DEBUG megazone23 mbatchStandardLegend - theLegendColors 1
## 2018 06 21 10:30:51.474 DEBUG megazone23 mbatchStandardLegend - theLegendSymbols 0
## 2018 06 21 10:30:51.475 DEBUG megazone23 mbatchStandardLegend - myColors #b30000

```

```

## 2018 06 21 10:30:51.475 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:51.495 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:51.498 DEBUG megazone23 mbatchStandardLegend - Calling .jinit /home/linux/R/x86_64
## 2018 06 21 10:30:51.787 DEBUG megazone23 mbatchStandardLegend - .jinit complete
## 2018 06 21 10:30:51.788 DEBUG megazone23 mbatchStandardLegend - theTitle Points: TSS (80)
## 2018 06 21 10:30:51.789 DEBUG megazone23 mbatchStandardLegend - theVersion MBatch 1.4.16
## 2018 06 21 10:30:51.790 DEBUG megazone23 mbatchStandardLegend - theFilenamePath /bea_testing/output
## 2018 06 21 10:30:51.790 DEBUG megazone23 mbatchStandardLegend - theLegendNames OR - University of M
## (72), OU - Roswell Park (1), P6 - Translational Genomics
## Research Institute (2), PA - University of Minnesota
## (1), PK - University Health
## Network (4)
## 2018 06 21 10:30:51.791 DEBUG megazone23 mbatchStandardLegend - theLegendNames 5
## 2018 06 21 10:30:51.791 DEBUG megazone23 mbatchStandardLegend - theLegendColors 5
## 2018 06 21 10:30:51.791 DEBUG megazone23 mbatchStandardLegend - theLegendSymbols 5
## 2018 06 21 10:30:51.792 DEBUG megazone23 mbatchStandardLegend - myColors #b30000,#8fb300,#00b347,#0
## 2018 06 21 10:30:51.792 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:51.860 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:51.862 DEBUG megazone23 mbatchStandardLegend - Calling .jinit /home/linux/R/x86_64
## 2018 06 21 10:30:52.131 DEBUG megazone23 mbatchStandardLegend - .jinit complete
## 2018 06 21 10:30:52.132 DEBUG megazone23 mbatchStandardLegend - theTitle Dispersion Metrics
## 2018 06 21 10:30:52.132 DEBUG megazone23 mbatchStandardLegend - theVersion MBatch 1.4.16
## 2018 06 21 10:30:52.133 DEBUG megazone23 mbatchStandardLegend - theFilenamePath /bea_testing/output
## 2018 06 21 10:30:52.133 DEBUG megazone23 mbatchStandardLegend - theLegendNames First PCA Component:
## 2018 06 21 10:30:52.133 DEBUG megazone23 mbatchStandardLegend - theLegendNames 14
## 2018 06 21 10:30:52.134 DEBUG megazone23 mbatchStandardLegend - theLegendColors 0
## 2018 06 21 10:30:52.134 DEBUG megazone23 mbatchStandardLegend - theLegendSymbols 0
## 2018 06 21 10:30:52.134 DEBUG megazone23 mbatchStandardLegend - myColors
## 2018 06 21 10:30:52.135 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:52.265 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:52.267 DEBUG megazone23 mbatchStandardCombineLegends - Calling .jinit /home/linux/
## 2018 06 21 10:30:52.530 DEBUG megazone23 mbatchStandardCombineLegends - .jinit complete
## 2018 06 21 10:30:52.531 DEBUG megazone23 mbatchStandardCombineLegends - theTitle Example Title for '
## 2018 06 21 10:30:52.531 DEBUG megazone23 mbatchStandardCombineLegends - theFilenamePath /bea_testing
## 2018 06 21 10:30:52.531 DEBUG megazone23 mbatchStandardCombineLegends - theListOfFiles /bea_testing
## 2018 06 21 10:30:52.532 DEBUG megazone23 mbatchStandardLegend before java
## 2018 06 21 10:30:52.774 DEBUG megazone23 mbatchStandardLegend after java
## 2018 06 21 10:30:52.775 DEBUG megazone23 myCompListDscFile
## [1] "ALL__CompListDSC.RData"
## [2] "BatchIdwithTSS/DualBatch/ANY_Comp1_Comp2_DSC.txt"
## [3] "BatchIdwithTSS/DualBatch/ANY_Comp1_Comp2_DSC.txt__CompDSC.RData"
## [4] "BatchIdwithTSS/DualBatch/ANY_Comp1_Comp3_DSC.txt"
## [5] "BatchIdwithTSS/DualBatch/ANY_Comp1_Comp3_DSC.txt__CompDSC.RData"
## [6] "BatchIdwithTSS/DualBatch/ANY_DSC.txt"
## [7] "BatchIdwithTSS/DualBatch/ANY_DSC.txt__OverallDSC.RData"
## [8] "BatchIdwithTSS/DualBatch/ANY_FVE.txt"
## [9] "BatchIdwithTSS/DualBatch/ANY_Scores.txt"
## [10] "BatchIdwithTSS/DualBatch/ANY_Weights.txt"
## [11] "BatchIdwithTSS/DualBatch/PCA-Plus/BatchIdwithTSS_Comp1_Comp2_Diagram.png"
## [12] "BatchIdwithTSS/DualBatch/PCA-Plus/BatchIdwithTSS_Comp1_Comp2_Legend-ALL.png"
## [13] "BatchIdwithTSS/DualBatch/PCA-Plus/BatchIdwithTSS_Comp1_Comp2_Legend-DSC.png"
## [14] "BatchIdwithTSS/DualBatch/PCA-Plus/BatchIdwithTSS_Comp1_Comp2_Legend-Points.png"
## [15] "BatchIdwithTSS/DualBatch/PCA-Plus/BatchIdwithTSS_Comp1_Comp2_Legend-Rays.png"
## [16] "BatchIdwithTSS/DualBatch/PCA-Plus/BatchIdwithTSS_Comp1_Comp3_Diagram.png"

```

```
## [17] "BatchIdwithTSS/DualBatch/PCA-Plus/BatchIdwithTSS_Comp1_Comp3_Legend-ALL.png"
## [18] "BatchIdwithTSS/DualBatch/PCA-Plus/BatchIdwithTSS_Comp1_Comp3_Legend-DSC.png"
## [19] "BatchIdwithTSS/DualBatch/PCA-Plus/BatchIdwithTSS_Comp1_Comp3_Legend-Points.png"
## [20] "BatchIdwithTSS/DualBatch/PCA-Plus/BatchIdwithTSS_Comp1_Comp3_Legend-Rays.png"
## [21] "BatchIdwithTSS/DualBatch/PCAAnnotations.tsv"
## [22] "BatchIdwithTSS/DualBatch/PCAVvalues.tsv"
## [23] "DSCOverview.tsv"
## [24] "PlateIdwithTSS/DualBatch/ANY_Comp1_Comp2_DSC.txt"
## [25] "PlateIdwithTSS/DualBatch/ANY_Comp1_Comp2_DSC.txt__CompDSC.RData"
## [26] "PlateIdwithTSS/DualBatch/ANY_Comp1_Comp3_DSC.txt"
## [27] "PlateIdwithTSS/DualBatch/ANY_Comp1_Comp3_DSC.txt__CompDSC.RData"
## [28] "PlateIdwithTSS/DualBatch/ANY_DSC.txt"
## [29] "PlateIdwithTSS/DualBatch/ANY_DSC.txt__OverallDSC.RData"
## [30] "PlateIdwithTSS/DualBatch/ANY_FVE.txt"
## [31] "PlateIdwithTSS/DualBatch/ANY_Scores.txt"
## [32] "PlateIdwithTSS/DualBatch/ANY_Weights.txt"
## [33] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp2_Diagram.png"
## [34] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp2_Legend-ALL.png"
## [35] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp2_Legend-DSC.png"
## [36] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp2_Legend-Points.png"
## [37] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp2_Legend-Rays.png"
## [38] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp3_Diagram.png"
## [39] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp3_Legend-ALL.png"
## [40] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp3_Legend-DSC.png"
## [41] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp3_Legend-Points.png"
## [42] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp3_Legend-Rays.png"
## [43] "PlateIdwithTSS/DualBatch/PCAAnnotations.tsv"
## [44] "PlateIdwithTSS/DualBatch/PCAVvalues.tsv"
```

7 Example File Output

The above code creates the following subdirectories and files. The RData and DSCOverview.tsv files are used internally. The subdirectories correspond to the Batch Types Pairs on which the PCA-Plus assessment was performed.

```
/bea_testing/output/PCA_DualBatch_Structures$ ls -l
total 16
-rw-r--r-- 1 linux linux 77 Jun 14 09:34 ALL__CompListDSC.RData
drwxr-xr-x 3 linux linux 4096 Jun 14 09:34 BatchIdwithTSS
-rw-r--r-- 1 linux linux 348 Jun 14 09:34 DSCOverview.tsv
drwxr-xr-x 3 linux linux 4096 Jun 14 09:34 PlateIdwithTSS
```

Looking at the “BatchIdwithTSS” subdirectory, it contains another directory called DualBatch (a different function provides “ManyToMany” comparisons). The DualBatch directory contains the following files and a PCA-Plus sub-directory.

DSC txt files are described above in the DoDscPermsFileFlag.

ANY_Scores.txt is described above in the DoSampleLocatorFlag.

The below files are tab delimited.

ANY_FVE.txt has three columns: “Component Number”, “FVE (%)”, “Cumulative FVE (%)”.

ANY_Weights.txt has PCA components as the headers, and features (such as, genes or probes) at the start of subsequent rows.

PCAVValues.tsv This has an initial header row, a row for FVE, and then rows giving sample ids and PCA values.

```
/bea_testing/output/PCA_DualBatch_Structures/BatchIdwithTSS/DualBatch$ ls -l
total 2120
-rw-r--r-- 1 linux linux      383 Jun 14 09:34 ANY_Comp1_Comp2_DSC.txt
-rw-r--r-- 1 linux linux      267 Jun 14 09:34 ANY_Comp1_Comp2_DSC.txt__CompDSC.RData
-rw-r--r-- 1 linux linux      383 Jun 14 09:34 ANY_Comp1_Comp3_DSC.txt
-rw-r--r-- 1 linux linux      267 Jun 14 09:34 ANY_Comp1_Comp3_DSC.txt__CompDSC.RData
-rw-r--r-- 1 linux linux     2266 Jun 14 09:34 ANY_DSC.txt
-rw-r--r-- 1 linux linux      923 Jun 14 09:34 ANY_DSC.txt__OverallDSC.RData
-rw-r--r-- 1 linux linux     1289 Jun 14 09:34 ANY_FVE.txt
-rw-r--r-- 1 linux linux    120163 Jun 14 09:34 ANY_Scores.txt
-rw-r--r-- 1 linux linux   2001324 Jun 14 09:34 ANY_Weights.txt
-rw-r--r-- 1 linux linux      639 Jun 14 09:34 PCAAnnotations.tsv
drwxr-xr-x 2 linux linux     4096 Jun 14 09:34 PCA-Plus
-rw-r--r-- 1 linux linux     6643 Jun 14 09:34 PCAValues.tsv
```

The PCA-Plus directory contains the different PCA-Plus PNG files and legends.

```
/bea_testing/output/PCA_DualBatch_Structures/BatchIdwithTSS/DualBatch/PCA-Plus$ ls -l
total 436
-rw-r--r-- 1 linux linux 120980 Jun 14 09:34 BatchIdwithTSS_Comp1_Comp2_Diagram.png
-rw-r--r-- 1 linux linux  51860 Jun 14 09:34 BatchIdwithTSS_Comp1_Comp2_Legend-ALL.png
-rw-r--r-- 1 linux linux  28370 Jun 14 09:34 BatchIdwithTSS_Comp1_Comp2_Legend-DSC.png
-rw-r--r-- 1 linux linux 14517 Jun 14 09:34 BatchIdwithTSS_Comp1_Comp2_Legend-Points.png
-rw-r--r-- 1 linux linux  3664 Jun 14 09:34 BatchIdwithTSS_Comp1_Comp2_Legend-Rays.png
-rw-r--r-- 1 linux linux 117616 Jun 14 09:34 BatchIdwithTSS_Comp1_Comp3_Diagram.png
-rw-r--r-- 1 linux linux  52066 Jun 14 09:34 BatchIdwithTSS_Comp1_Comp3_Legend-ALL.png
-rw-r--r-- 1 linux linux  28484 Jun 14 09:34 BatchIdwithTSS_Comp1_Comp3_Legend-DSC.png
-rw-r--r-- 1 linux linux 14517 Jun 14 09:34 BatchIdwithTSS_Comp1_Comp3_Legend-Points.png
-rw-r--r-- 1 linux linux  3664 Jun 14 09:34 BatchIdwithTSS_Comp1_Comp3_Legend-Rays.png
```

Here is a diagram generated from this code.

Example Title for Test Data BatchIdwithTSS 1 2

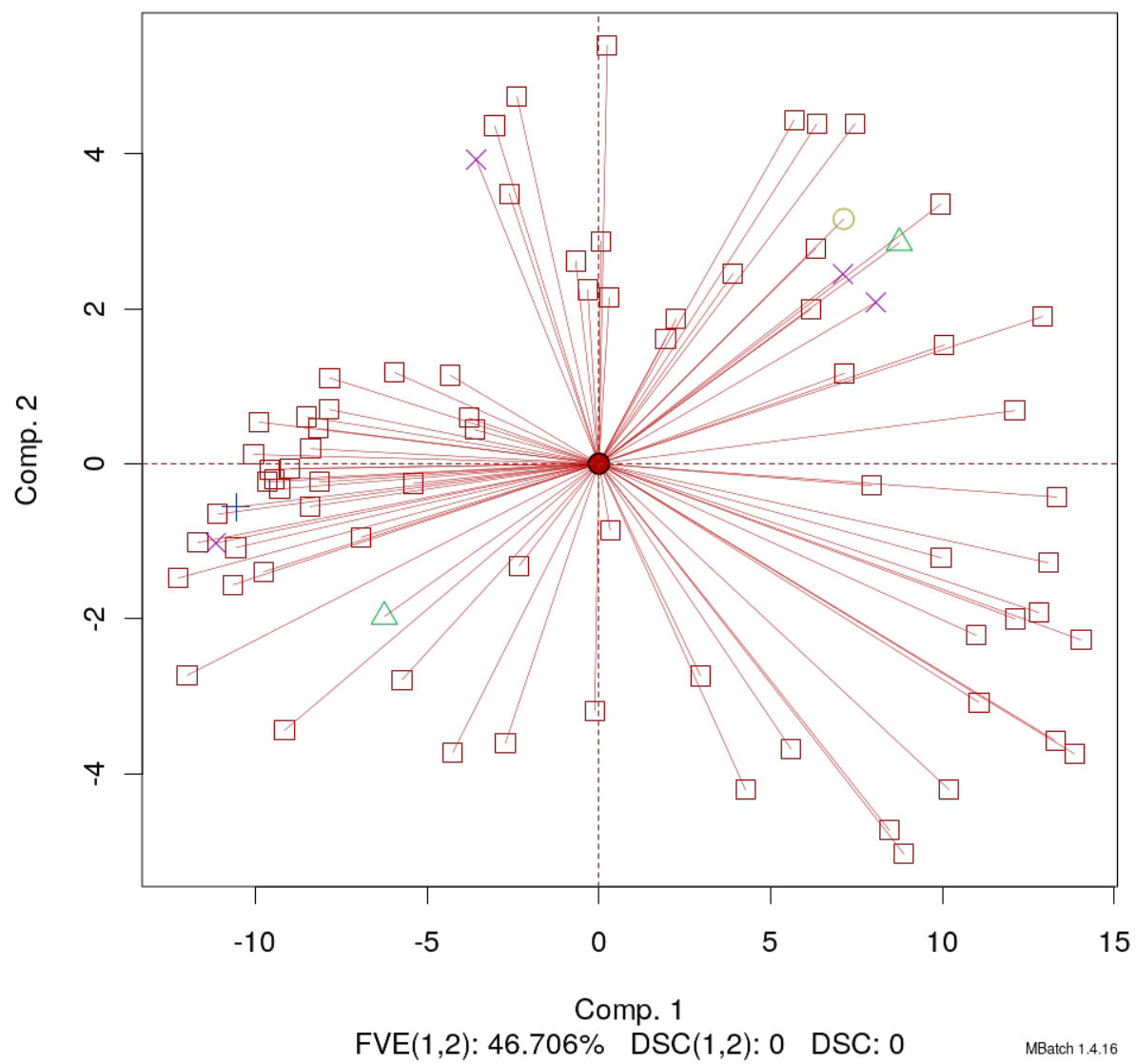


Figure 2: PCAPlus Output