

Introduction

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.

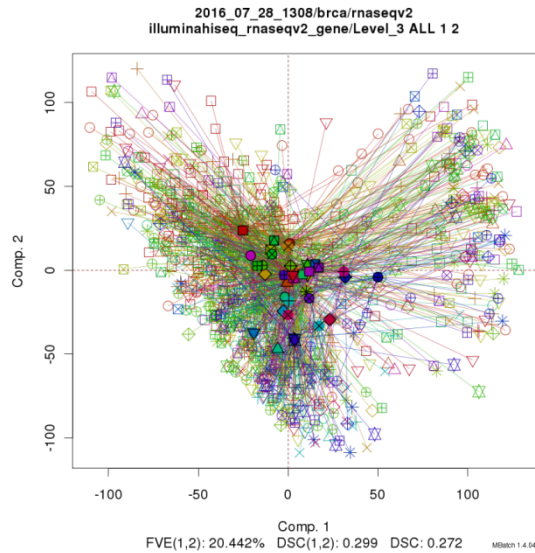
Algorithm

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.

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>.



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)`

Arguments

`##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.

##theTitle A string title to use in PNG files.

##theOutputPath String giving directory in which to place output PNG files.

##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.

##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.

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

##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")))} }

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

##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.

##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.

##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)

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

##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.

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

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

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

##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.

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)

  inputDir <- getTestInputDir()
  outputDir <- getTestOutputDir()
  compareDir <- getTestCompareDir()

  # set the paths
  theGeneFile=file.path(inputDir, "matrix_data-Tumor.tsv")
  theBatchFile=file.path(inputDir, "batches-Tumor.tsv")
  theOutputDir=file.path(outputDir, "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))
}

## All sorting in this package requires using a Sys.setlocale("LC_COLLATE","C"). MBatch Vers
## 2020 11 18 16:18:02.301 DEBUG ab7c64738d52 Changing LC_COLLATE to C for duration of run
## 2020 11 18 16:18:02.301 INFO ab7c64738d52 \ / \ / \ / \ / \ / \ / \ / \ / \ /
## 2020 11 18 16:18:02.301 INFO ab7c64738d52 Starting mbatchLoadFiles
## 2020 11 18 16:18:02.302 INFO ab7c64738d52 MBatch Version: BEA_VERSION_TIMESTAMP
## 2020 11 18 16:18:02.302 INFO ab7c64738d52 read batch file= /builds/BatchEffects_clean/Bat
## 2020 11 18 16:18:02.305 INFO ab7c64738d52 read gene file= /builds/BatchEffects_clean/Bat
## 2020 11 18 16:18:05.810 INFO ab7c64738d52 filter samples in batches using gene samples
## 2020 11 18 16:18:05.811 INFO ab7c64738d52 sort batches by gene file samples
## 2020 11 18 16:18:05.886 INFO ab7c64738d52 Finishing mbatchLoadFiles
## 2020 11 18 16:18:05.886 INFO ab7c64738d52 ~~~~~
## 2020 11 18 16:18:05.887 DEBUG ab7c64738d52 Changing LC_COLLATE to C for duration of run
## 2020 11 18 16:18:05.887 INFO ab7c64738d52 \ / \ / \ / \ / \ / \ / \ / \ / \ /
## 2020 11 18 16:18:05.887 INFO ab7c64738d52 mbatchTrimData Starting
## 2020 11 18 16:18:05.887 INFO ab7c64738d52 MBatch Version: BEA_VERSION_TIMESTAMP
## 2020 11 18 16:18:13.444 INFO ab7c64738d52 mbatchTrimData theMaxSize= 1e+05
## 2020 11 18 16:18:13.444 INFO ab7c64738d52 mbatchTrimData ncol(theMatrix)= 80
## 2020 11 18 16:18:13.444 INFO ab7c64738d52 mbatchTrimData nrow(theMatrix)= 1250
## 2020 11 18 16:18:13.445 INFO ab7c64738d52 mbatchTrimData Finishing
## 2020 11 18 16:18:13.445 INFO ab7c64738d52 ~~~~~
## 2020 11 18 16:18:13.447 DEBUG ab7c64738d52 Changing LC_COLLATE to C for duration of run
## 2020 11 18 16:18:13.447 INFO ab7c64738d52 \ / \ / \ / \ / \ / \ / \ / \ / \ /
## 2020 11 18 16:18:13.447 INFO ab7c64738d52 mbatchFilterData Starting
## 2020 11 18 16:18:13.447 INFO ab7c64738d52 MBatch Version: BEA_VERSION_TIMESTAMP
## 2020 11 18 16:18:13.448 DEBUG ab7c64738d52 rows pre filter 1250
## 2020 11 18 16:18:13.647 DEBUG ab7c64738d52 rows post filter 1250
## 2020 11 18 16:18:13.647 DEBUG ab7c64738d52 mbatchFilterData Prefilter, gene data had 125
## 2020 11 18 16:18:13.648 DEBUG ab7c64738d52 mbatchFilterData Prefilter, batch data had 80
## 2020 11 18 16:18:13.648 INFO ab7c64738d52 mbatchFilterData Finishing

```

```

## 2020 11 18 16:18:13.649 INFO ab7c64738d52 ~~~~~
## 2020 11 18 16:18:13.650 DEBUG ab7c64738d52 Calling .jinit for PCA
## 2020 11 18 16:18:13.650 DEBUG ab7c64738d52 classpath=, /tmp/RtmpeZh3Bb/Rinst18036e4eb2d/M
## 2020 11 18 16:18:13.790 DEBUG ab7c64738d52 Changing LC_COLLATE to C for duration of run
## 2020 11 18 16:18:13.791 DEBUG ab7c64738d52 createBatchEffectsOutput_pca_dualBatch - start
## 2020 11 18 16:18:13.792 DEBUG ab7c64738d52 doSamplePcaCall - pre-filter matrix size= 1250
## 2020 11 18 16:18:13.793 DEBUG ab7c64738d52 filterBasedOnGeneLimit theMatrixGeneData nrow=
## 2020 11 18 16:18:13.793 DEBUG ab7c64738d52 filterBasedOnGeneLimit theMatrixGeneData ncol=
## 2020 11 18 16:18:13.793 DEBUG ab7c64738d52 doSamplePcaCall - pre-filter matrix size= 1250
## 2020 11 18 16:18:13.794 DEBUG ab7c64738d52 doSamplePcaCall - no-NA matrix length= 1247
## 2020 11 18 16:18:13.821 DEBUG ab7c64738d52 doSamplePcaCall - max(theCompPairList)= 3
## 2020 11 18 16:18:13.821 DEBUG ab7c64738d52 doSamplePcaCall - pca scores size= 80,80
## 2020 11 18 16:18:13.822 DEBUG ab7c64738d52 checkCreateDir: /builds/BatchEffects_clean/Ba
## 2020 11 18 16:18:13.823 DEBUG ab7c64738d52 checkCreateDir: /builds/BatchEffects_clean/Ba
## 2020 11 18 16:18:13.823 DEBUG ab7c64738d52 pvalueDSC start
## 2020 11 18 16:18:13.823 DEBUG ab7c64738d52 pvalueDSC component A < 1
## 2020 11 18 16:18:13.824 DEBUG ab7c64738d52 pvalueDSC call pvalueDSCwithExcerpt
## 2020 11 18 16:18:13.824 DEBUG ab7c64738d52 pvalueDSCwithExcerpt start
## 2020 11 18 16:18:13.825 DEBUG ab7c64738d52 nrow(thePcaDataExcerpt)= 80
## 2020 11 18 16:18:13.825 DEBUG ab7c64738d52 ncol(thePcaDataExcerpt)= 80
## 2020 11 18 16:18:13.825 DEBUG ab7c64738d52 length(theBatchIdsForSamples)= 80
## 2020 11 18 16:18:13.829 DEBUG ab7c64738d52 getDSCwithExcerpt before java
## 2020 11 18 16:18:13.880 DEBUG ab7c64738d52 getDSCwithExcerpt after java
## 2020 11 18 16:18:13.881 DEBUG ab7c64738d52 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2020 11 18 16:18:13.882 DEBUG ab7c64738d52 pvalueDSCwithExcerpt length(unique(theBatchIds
## 2020 11 18 16:18:13.882 DEBUG ab7c64738d52 pvalueDSCwithExcerpt 2222 length(permResults)=
## 2020 11 18 16:18:13.882 DEBUG ab7c64738d52 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2020 11 18 16:18:13.889 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- start
## 2020 11 18 16:18:13.889 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- redo name
## 2020 11 18 16:18:13.889 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- check file
## 2020 11 18 16:18:13.890 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- no file, write
## 2020 11 18 16:18:13.890 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- nothing to write
## 2020 11 18 16:18:13.890 DEBUG ab7c64738d52 write writePcaDataFilesForDataset
## 2020 11 18 16:18:13.890 DEBUG ab7c64738d52 checkCreateDir: /builds/BatchEffects_clean/Ba
## 2020 11 18 16:18:13.891 DEBUG ab7c64738d52 writePCAValuesTSV
## 2020 11 18 16:18:13.891 DEBUG ab7c64738d52 writePCAValuesTSV outputFile
## 2020 11 18 16:18:13.891 DEBUG ab7c64738d52 writePCAValuesTSV components
## 2020 11 18 16:18:13.891 DEBUG ab7c64738d52 writePCAValuesTSV myDataFrame
## 2020 11 18 16:18:13.892 DEBUG ab7c64738d52 writePCAValuesTSV - values
## 2020 11 18 16:18:13.892 DEBUG ab7c64738d52 writePCAValuesTSV - matrix
## 2020 11 18 16:18:13.893 DEBUG ab7c64738d52 writePCAValuesTSV - write
## 2020 11 18 16:18:13.895 DEBUG ab7c64738d52 writePCAValuesTSV finished /builds/BatchEffec
## 2020 11 18 16:18:13.896 DEBUG ab7c64738d52 writePCAAnnotations
## 2020 11 18 16:18:13.896 DEBUG ab7c64738d52 writePCAAnnotations - header and all data
## 2020 11 18 16:18:14.159 DEBUG ab7c64738d52 doInternalPca before filter - Number of rownan
## 2020 11 18 16:18:14.159 DEBUG ab7c64738d52 doInternalPca before filter - length of batch

```

```

## 2020 11 18 16:18:14.159 DEBUG ab7c64738d52 doInternalPca after filter - Number of rownames
## 2020 11 18 16:18:14.160 DEBUG ab7c64738d52 doInternalPca after filter - length of batch 1
## 2020 11 18 16:18:14.160 DEBUG ab7c64738d52 doInternalPca - pvalueDSC
## 2020 11 18 16:18:14.160 DEBUG ab7c64738d52 pvalueDSC start
## 2020 11 18 16:18:14.160 DEBUG ab7c64738d52 pvalueDSC transpose
## 2020 11 18 16:18:14.160 DEBUG ab7c64738d52 pvalueDSC call pvalueDSCwithExcerpt
## 2020 11 18 16:18:14.161 DEBUG ab7c64738d52 pvalueDSCwithExcerpt start
## 2020 11 18 16:18:14.161 DEBUG ab7c64738d52 nrow(thePcaDataExcerpt)= 2
## 2020 11 18 16:18:14.161 DEBUG ab7c64738d52 ncol(thePcaDataExcerpt)= 80
## 2020 11 18 16:18:14.161 DEBUG ab7c64738d52 length(theBatchIdsForSamples)= 80
## 2020 11 18 16:18:14.162 DEBUG ab7c64738d52 getDSCwithExcerpt before java
## 2020 11 18 16:18:14.163 DEBUG ab7c64738d52 getDSCwithExcerpt after java
## 2020 11 18 16:18:14.163 DEBUG ab7c64738d52 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2020 11 18 16:18:14.164 DEBUG ab7c64738d52 pvalueDSCwithExcerpt length(unique(theBatchIds
## 2020 11 18 16:18:14.164 DEBUG ab7c64738d52 pvalueDSCwithExcerpt 2222 length(permResults)=
## 2020 11 18 16:18:14.164 DEBUG ab7c64738d52 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2020 11 18 16:18:14.164 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- start
## 2020 11 18 16:18:14.165 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- redo name
## 2020 11 18 16:18:14.165 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- check file
## 2020 11 18 16:18:14.165 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- no file, write
## 2020 11 18 16:18:14.165 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- nothing to write
## 2020 11 18 16:18:14.166 DEBUG ab7c64738d52 writePCAAnnotations
## 2020 11 18 16:18:14.167 DEBUG ab7c64738d52 writePCAAnnotations - data PC1,PC2
## 2020 11 18 16:18:14.179 DEBUG ab7c64738d52 writePcaAnalysis_Image - Number of rownames fo
## 2020 11 18 16:18:14.179 DEBUG ab7c64738d52 writePcaAnalysis_Image - length of batch ids 1

## 2020 11 18 16:18:14.268 DEBUG ab7c64738d52 mbatchStandardLegend - Calling .jinit /tmp/Rt
## 2020 11 18 16:18:14.270 DEBUG ab7c64738d52 mbatchStandardLegend - .jinit complete
## 2020 11 18 16:18:14.270 DEBUG ab7c64738d52 mbatchStandardLegend - theTitle Rays: PlateId
## 2020 11 18 16:18:14.270 DEBUG ab7c64738d52 mbatchStandardLegend - theVersion MBatch 1.7
## 2020 11 18 16:18:14.271 DEBUG ab7c64738d52 mbatchStandardLegend - theFilenamePath /builc
## 2020 11 18 16:18:14.271 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames A29J (8
## 2020 11 18 16:18:14.271 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames 1
## 2020 11 18 16:18:14.272 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendColors 1
## 2020 11 18 16:18:14.272 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendSymbols 0
## 2020 11 18 16:18:14.272 DEBUG ab7c64738d52 mbatchStandardLegend - myColors #b30000
## 2020 11 18 16:18:14.272 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:14.448 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:14.450 DEBUG ab7c64738d52 mbatchStandardLegend - Calling .jinit /tmp/Rt
## 2020 11 18 16:18:14.451 DEBUG ab7c64738d52 mbatchStandardLegend - .jinit complete
## 2020 11 18 16:18:14.452 DEBUG ab7c64738d52 mbatchStandardLegend - theTitle Points: TSS (
## 2020 11 18 16:18:14.452 DEBUG ab7c64738d52 mbatchStandardLegend - theVersion MBatch 1.7
## 2020 11 18 16:18:14.452 DEBUG ab7c64738d52 mbatchStandardLegend - theFilenamePath /builc
## 2020 11 18 16:18:14.453 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames OR - Un
## (72), OU - Roswell Park (1), P6 - Translational Genomics
## Research Institute (2), PA - University of Minnesota

```

```

## (1), PK - University Health
## Network (4)
## 2020 11 18 16:18:14.453 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames 5
## 2020 11 18 16:18:14.453 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendColors 5
## 2020 11 18 16:18:14.453 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendSymbols 5
## 2020 11 18 16:18:14.454 DEBUG ab7c64738d52 mbatchStandardLegend - myColors #b30000,#8fb3
## 2020 11 18 16:18:14.454 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:14.531 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:14.532 DEBUG ab7c64738d52 mbatchStandardLegend - Calling .jinit /tmp/Rt
## 2020 11 18 16:18:14.534 DEBUG ab7c64738d52 mbatchStandardLegend - .jinit complete
## 2020 11 18 16:18:14.534 DEBUG ab7c64738d52 mbatchStandardLegend - theTitle Dispersion Me
## 2020 11 18 16:18:14.535 DEBUG ab7c64738d52 mbatchStandardLegend - theVersion MBatch 1.7
## 2020 11 18 16:18:14.535 DEBUG ab7c64738d52 mbatchStandardLegend - theFilenamePath /build
## 2020 11 18 16:18:14.535 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames First P
## 2020 11 18 16:18:14.536 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames 14
## 2020 11 18 16:18:14.536 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendColors 0
## 2020 11 18 16:18:14.536 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendSymbols 0
## 2020 11 18 16:18:14.536 DEBUG ab7c64738d52 mbatchStandardLegend - myColors
## 2020 11 18 16:18:14.536 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:14.669 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:14.670 DEBUG ab7c64738d52 mbatchStandardCombineLegends - Calling .jinit
## 2020 11 18 16:18:14.672 DEBUG ab7c64738d52 mbatchStandardCombineLegends - .jinit complete
## 2020 11 18 16:18:14.672 DEBUG ab7c64738d52 mbatchStandardCombineLegends - theTitle Examp
## 2020 11 18 16:18:14.673 DEBUG ab7c64738d52 mbatchStandardCombineLegends - theFilenamePat
## 2020 11 18 16:18:14.673 DEBUG ab7c64738d52 mbatchStandardCombineLegends - theListOfFiles
## 2020 11 18 16:18:14.673 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:15.016 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:15.020 DEBUG ab7c64738d52 doInternalPca before filter - Number of rowname
## 2020 11 18 16:18:15.020 DEBUG ab7c64738d52 doInternalPca before filter - length of batch
## 2020 11 18 16:18:15.020 DEBUG ab7c64738d52 doInternalPca after filter - Number of rowname
## 2020 11 18 16:18:15.020 DEBUG ab7c64738d52 doInternalPca after filter - length of batch
## 2020 11 18 16:18:15.020 DEBUG ab7c64738d52 doInternalPca - pvalueDSC
## 2020 11 18 16:18:15.021 DEBUG ab7c64738d52 pvalueDSC start
## 2020 11 18 16:18:15.021 DEBUG ab7c64738d52 pvalueDSC transpose
## 2020 11 18 16:18:15.021 DEBUG ab7c64738d52 pvalueDSC call pvalueDSCwithExcerpt
## 2020 11 18 16:18:15.021 DEBUG ab7c64738d52 pvalueDSCwithExcerpt start
## 2020 11 18 16:18:15.022 DEBUG ab7c64738d52 nrow(thePcaDataExcerpt)= 2
## 2020 11 18 16:18:15.022 DEBUG ab7c64738d52 ncol(thePcaDataExcerpt)= 80
## 2020 11 18 16:18:15.022 DEBUG ab7c64738d52 length(theBatchIdsForSamples)= 80
## 2020 11 18 16:18:15.022 DEBUG ab7c64738d52 getDSCwithExcerpt before java
## 2020 11 18 16:18:15.023 DEBUG ab7c64738d52 getDSCwithExcerpt after java
## 2020 11 18 16:18:15.024 DEBUG ab7c64738d52 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2020 11 18 16:18:15.024 DEBUG ab7c64738d52 pvalueDSCwithExcerpt length(unique(theBatchIds
## 2020 11 18 16:18:15.024 DEBUG ab7c64738d52 pvalueDSCwithExcerpt 2222 length(permResults)=
## 2020 11 18 16:18:15.024 DEBUG ab7c64738d52 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2020 11 18 16:18:15.025 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- start

```



```

## 2020 11 18 16:18:15.025 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- redo name
## 2020 11 18 16:18:15.025 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- check file
## 2020 11 18 16:18:15.025 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- no file, write
## 2020 11 18 16:18:15.026 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- nothing to write
## 2020 11 18 16:18:15.026 DEBUG ab7c64738d52 writePCAAnnotations
## 2020 11 18 16:18:15.027 DEBUG ab7c64738d52 writePCAAnnotations - data PC1,PC3
## 2020 11 18 16:18:15.029 DEBUG ab7c64738d52 writePcaAnalysis_Image - Number of rownames fo
## 2020 11 18 16:18:15.029 DEBUG ab7c64738d52 writePcaAnalysis_Image - length of batch ids 1

## 2020 11 18 16:18:15.110 DEBUG ab7c64738d52 mbatchStandardLegend - Calling .jinit /tmp/Rt
## 2020 11 18 16:18:15.112 DEBUG ab7c64738d52 mbatchStandardLegend - .jinit complete
## 2020 11 18 16:18:15.112 DEBUG ab7c64738d52 mbatchStandardLegend - theTitle Rays: PlateId
## 2020 11 18 16:18:15.112 DEBUG ab7c64738d52 mbatchStandardLegend - theVersion MBatch 1.7
## 2020 11 18 16:18:15.113 DEBUG ab7c64738d52 mbatchStandardLegend - theFilenamePath /build
## 2020 11 18 16:18:15.113 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames A29J (8
## 2020 11 18 16:18:15.113 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames 1
## 2020 11 18 16:18:15.113 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendColors 1
## 2020 11 18 16:18:15.114 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendSymbols 0
## 2020 11 18 16:18:15.114 DEBUG ab7c64738d52 mbatchStandardLegend - myColors #b30000
## 2020 11 18 16:18:15.114 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:15.132 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:15.133 DEBUG ab7c64738d52 mbatchStandardLegend - Calling .jinit /tmp/Rt
## 2020 11 18 16:18:15.135 DEBUG ab7c64738d52 mbatchStandardLegend - .jinit complete
## 2020 11 18 16:18:15.135 DEBUG ab7c64738d52 mbatchStandardLegend - theTitle Points: TSS (
## 2020 11 18 16:18:15.135 DEBUG ab7c64738d52 mbatchStandardLegend - theVersion MBatch 1.7
## 2020 11 18 16:18:15.136 DEBUG ab7c64738d52 mbatchStandardLegend - theFilenamePath /build
## 2020 11 18 16:18:15.136 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames OR - Un
## (72), OU - Roswell Park (1), P6 - Translational Genomics
## Research Institute (2), PA - University of Minnesota
## (1), PK - University Health
## Network (4)
## 2020 11 18 16:18:15.136 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames 5
## 2020 11 18 16:18:15.136 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendColors 5
## 2020 11 18 16:18:15.137 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendSymbols 5
## 2020 11 18 16:18:15.137 DEBUG ab7c64738d52 mbatchStandardLegend - myColors #b30000,#8fb3
## 2020 11 18 16:18:15.137 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:15.203 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:15.204 DEBUG ab7c64738d52 mbatchStandardLegend - Calling .jinit /tmp/Rt
## 2020 11 18 16:18:15.206 DEBUG ab7c64738d52 mbatchStandardLegend - .jinit complete
## 2020 11 18 16:18:15.206 DEBUG ab7c64738d52 mbatchStandardLegend - theTitle Dispersion Me
## 2020 11 18 16:18:15.206 DEBUG ab7c64738d52 mbatchStandardLegend - theVersion MBatch 1.7
## 2020 11 18 16:18:15.207 DEBUG ab7c64738d52 mbatchStandardLegend - theFilenamePath /build
## 2020 11 18 16:18:15.207 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames First P
## 2020 11 18 16:18:15.207 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames 14
## 2020 11 18 16:18:15.208 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendColors 0
## 2020 11 18 16:18:15.208 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendSymbols 0

```

```

## 2020 11 18 16:18:15.208 DEBUG ab7c64738d52 mbatchStandardLegend - myColors
## 2020 11 18 16:18:15.208 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:15.337 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:15.338 DEBUG ab7c64738d52 mbatchStandardCombineLegends - Calling .jinit
## 2020 11 18 16:18:15.340 DEBUG ab7c64738d52 mbatchStandardCombineLegends - .jinit complete
## 2020 11 18 16:18:15.340 DEBUG ab7c64738d52 mbatchStandardCombineLegends - theTitle Examp
## 2020 11 18 16:18:15.340 DEBUG ab7c64738d52 mbatchStandardCombineLegends - theFilenamePath
## 2020 11 18 16:18:15.340 DEBUG ab7c64738d52 mbatchStandardCombineLegends - theListOfFiles
## 2020 11 18 16:18:15.341 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:15.600 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:15.601 DEBUG ab7c64738d52 doSamplePcaCall - pre-filter matrix size= 1250
## 2020 11 18 16:18:15.601 DEBUG ab7c64738d52 filterBasedOnGeneLimit theMatrixGeneData nrow=
## 2020 11 18 16:18:15.601 DEBUG ab7c64738d52 filterBasedOnGeneLimit theMatrixGeneData ncol=
## 2020 11 18 16:18:15.602 DEBUG ab7c64738d52 doSamplePcaCall - pre-filter matrix size= 1250
## 2020 11 18 16:18:15.602 DEBUG ab7c64738d52 doSamplePcaCall - no-NA matrix length= 1247
## 2020 11 18 16:18:15.621 DEBUG ab7c64738d52 doSamplePcaCall - max(theCompPairList)= 3
## 2020 11 18 16:18:15.621 DEBUG ab7c64738d52 doSamplePcaCall - pca scores size= 80,80
## 2020 11 18 16:18:15.622 DEBUG ab7c64738d52 checkCreateDir: /builds/BatchEffects_clean/Ba
## 2020 11 18 16:18:15.622 DEBUG ab7c64738d52 checkCreateDir: /builds/BatchEffects_clean/Ba
## 2020 11 18 16:18:15.623 DEBUG ab7c64738d52 pvalueDSC start
## 2020 11 18 16:18:15.623 DEBUG ab7c64738d52 pvalueDSC component A < 1
## 2020 11 18 16:18:15.623 DEBUG ab7c64738d52 pvalueDSC call pvalueDSCwithExcerpt
## 2020 11 18 16:18:15.623 DEBUG ab7c64738d52 pvalueDSCwithExcerpt start
## 2020 11 18 16:18:15.624 DEBUG ab7c64738d52 nrow(thePcaDataExcerpt)= 80
## 2020 11 18 16:18:15.624 DEBUG ab7c64738d52 ncol(thePcaDataExcerpt)= 80
## 2020 11 18 16:18:15.624 DEBUG ab7c64738d52 length(theBatchIdsForSamples)= 80
## 2020 11 18 16:18:15.624 DEBUG ab7c64738d52 getDSCwithExcerpt before java
## 2020 11 18 16:18:15.628 DEBUG ab7c64738d52 getDSCwithExcerpt after java
## 2020 11 18 16:18:15.628 DEBUG ab7c64738d52 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2020 11 18 16:18:15.629 DEBUG ab7c64738d52 pvalueDSCwithExcerpt length(unique(theBatchIds
## 2020 11 18 16:18:15.629 DEBUG ab7c64738d52 pvalueDSCwithExcerpt 2222 length(permResults)=
## 2020 11 18 16:18:15.629 DEBUG ab7c64738d52 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2020 11 18 16:18:15.634 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- start
## 2020 11 18 16:18:15.635 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- redo name
## 2020 11 18 16:18:15.635 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- check file
## 2020 11 18 16:18:15.635 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- no file, write
## 2020 11 18 16:18:15.635 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- nothing to write
## 2020 11 18 16:18:15.636 DEBUG ab7c64738d52 write writePcaDataFilesForDataset
## 2020 11 18 16:18:15.636 DEBUG ab7c64738d52 checkCreateDir: /builds/BatchEffects_clean/Ba
## 2020 11 18 16:18:15.636 DEBUG ab7c64738d52 writePCAValuesTSV
## 2020 11 18 16:18:15.636 DEBUG ab7c64738d52 writePCAValuesTSV outputFile
## 2020 11 18 16:18:15.637 DEBUG ab7c64738d52 writePCAValuesTSV components
## 2020 11 18 16:18:15.637 DEBUG ab7c64738d52 writePCAValuesTSV myDataFrame
## 2020 11 18 16:18:15.637 DEBUG ab7c64738d52 writePCAValuesTSV - values
## 2020 11 18 16:18:15.637 DEBUG ab7c64738d52 writePCAValuesTSV - matrix
## 2020 11 18 16:18:15.638 DEBUG ab7c64738d52 writePCAValuesTSV - write

```

```

## 2020 11 18 16:18:15.639 DEBUG ab7c64738d52 writePCAValuesTSV finished /builds/BatchEffect
## 2020 11 18 16:18:15.640 DEBUG ab7c64738d52 writePCAAnnotations
## 2020 11 18 16:18:15.640 DEBUG ab7c64738d52 writePCAAnnotations - header and all data
## 2020 11 18 16:18:15.814 DEBUG ab7c64738d52 doInternalPca before filter - Number of rownames
## 2020 11 18 16:18:15.815 DEBUG ab7c64738d52 doInternalPca before filter - length of batch
## 2020 11 18 16:18:15.815 DEBUG ab7c64738d52 doInternalPca after filter - Number of rownames
## 2020 11 18 16:18:15.815 DEBUG ab7c64738d52 doInternalPca after filter - length of batch
## 2020 11 18 16:18:15.815 DEBUG ab7c64738d52 doInternalPca - pvalueDSC
## 2020 11 18 16:18:15.815 DEBUG ab7c64738d52 pvalueDSC start
## 2020 11 18 16:18:15.816 DEBUG ab7c64738d52 pvalueDSC transpose
## 2020 11 18 16:18:15.816 DEBUG ab7c64738d52 pvalueDSC call pvalueDSCwithExcerpt
## 2020 11 18 16:18:15.816 DEBUG ab7c64738d52 pvalueDSCwithExcerpt start
## 2020 11 18 16:18:15.816 DEBUG ab7c64738d52 nrow(thePcaDataExcerpt)= 2
## 2020 11 18 16:18:15.817 DEBUG ab7c64738d52 ncol(thePcaDataExcerpt)= 80
## 2020 11 18 16:18:15.817 DEBUG ab7c64738d52 length(theBatchIdsForSamples)= 80
## 2020 11 18 16:18:15.817 DEBUG ab7c64738d52 getDSCwithExcerpt before java
## 2020 11 18 16:18:15.818 DEBUG ab7c64738d52 getDSCwithExcerpt after java
## 2020 11 18 16:18:15.819 DEBUG ab7c64738d52 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2020 11 18 16:18:15.819 DEBUG ab7c64738d52 pvalueDSCwithExcerpt length(unique(theBatchIds
## 2020 11 18 16:18:15.819 DEBUG ab7c64738d52 pvalueDSCwithExcerpt 2222 length(permResults)=
## 2020 11 18 16:18:15.819 DEBUG ab7c64738d52 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2020 11 18 16:18:15.820 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- start
## 2020 11 18 16:18:15.820 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- redo name
## 2020 11 18 16:18:15.820 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- check file
## 2020 11 18 16:18:15.820 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- no file, write
## 2020 11 18 16:18:15.821 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- nothing to write
## 2020 11 18 16:18:15.821 DEBUG ab7c64738d52 writePCAAnnotations
## 2020 11 18 16:18:15.822 DEBUG ab7c64738d52 writePCAAnnotations - data PC1,PC2
## 2020 11 18 16:18:15.823 DEBUG ab7c64738d52 writePcaAnalysis_Image - Number of rownames fo
## 2020 11 18 16:18:15.824 DEBUG ab7c64738d52 writePcaAnalysis_Image - length of batch ids f

## 2020 11 18 16:18:15.908 DEBUG ab7c64738d52 mbatchStandardLegend - Calling .jinit /tmp/Rt
## 2020 11 18 16:18:15.910 DEBUG ab7c64738d52 mbatchStandardLegend - .jinit complete
## 2020 11 18 16:18:15.910 DEBUG ab7c64738d52 mbatchStandardLegend - theTitle Rays: BatchId
## 2020 11 18 16:18:15.910 DEBUG ab7c64738d52 mbatchStandardLegend - theVersion MBatch 1.7
## 2020 11 18 16:18:15.911 DEBUG ab7c64738d52 mbatchStandardLegend - theFilenamePath /build
## 2020 11 18 16:18:15.911 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames 00304
## 2020 11 18 16:18:15.911 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames 1
## 2020 11 18 16:18:15.911 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendColors 1
## 2020 11 18 16:18:15.912 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendSymbols 0
## 2020 11 18 16:18:15.912 DEBUG ab7c64738d52 mbatchStandardLegend - myColors #b30000
## 2020 11 18 16:18:15.912 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:15.931 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:15.932 DEBUG ab7c64738d52 mbatchStandardLegend - Calling .jinit /tmp/Rt
## 2020 11 18 16:18:15.934 DEBUG ab7c64738d52 mbatchStandardLegend - .jinit complete
## 2020 11 18 16:18:15.934 DEBUG ab7c64738d52 mbatchStandardLegend - theTitle Points: TSS

```

```

## 2020 11 18 16:18:15.934 DEBUG ab7c64738d52 mbatchStandardLegend - theVersion MBatch 1.7
## 2020 11 18 16:18:15.935 DEBUG ab7c64738d52 mbatchStandardLegend - theFilenamePath /build
## 2020 11 18 16:18:15.935 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames OR - Un
## (72), OU - Roswell Park (1), P6 - Translational Genomics
## Research Institute (2), PA - University of Minnesota
## (1), PK - University Health
## Network (4)
## 2020 11 18 16:18:15.935 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames 5
## 2020 11 18 16:18:15.936 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendColors 5
## 2020 11 18 16:18:15.936 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendSymbols 5
## 2020 11 18 16:18:15.936 DEBUG ab7c64738d52 mbatchStandardLegend - myColors #b30000,#8fb3
## 2020 11 18 16:18:15.936 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:16.003 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:16.004 DEBUG ab7c64738d52 mbatchStandardLegend - Calling .jinit /tmp/Rt
## 2020 11 18 16:18:16.006 DEBUG ab7c64738d52 mbatchStandardLegend - .jinit complete
## 2020 11 18 16:18:16.006 DEBUG ab7c64738d52 mbatchStandardLegend - theTitle Dispersion Me
## 2020 11 18 16:18:16.006 DEBUG ab7c64738d52 mbatchStandardLegend - theVersion MBatch 1.7
## 2020 11 18 16:18:16.007 DEBUG ab7c64738d52 mbatchStandardLegend - theFilenamePath /build
## 2020 11 18 16:18:16.007 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames First P
## 2020 11 18 16:18:16.007 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames 14
## 2020 11 18 16:18:16.008 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendColors 0
## 2020 11 18 16:18:16.008 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendSymbols 0
## 2020 11 18 16:18:16.008 DEBUG ab7c64738d52 mbatchStandardLegend - myColors
## 2020 11 18 16:18:16.008 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:16.139 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:16.140 DEBUG ab7c64738d52 mbatchStandardCombineLegends - Calling .jinit
## 2020 11 18 16:18:16.142 DEBUG ab7c64738d52 mbatchStandardCombineLegends - .jinit complete
## 2020 11 18 16:18:16.142 DEBUG ab7c64738d52 mbatchStandardCombineLegends - theTitle Examp
## 2020 11 18 16:18:16.142 DEBUG ab7c64738d52 mbatchStandardCombineLegends - theFilenamePath
## 2020 11 18 16:18:16.142 DEBUG ab7c64738d52 mbatchStandardCombineLegends - theListOfFiles
## 2020 11 18 16:18:16.143 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:16.408 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:16.409 DEBUG ab7c64738d52 doInternalPca before filter - Number of rowname
## 2020 11 18 16:18:16.410 DEBUG ab7c64738d52 doInternalPca before filter - length of batch
## 2020 11 18 16:18:16.410 DEBUG ab7c64738d52 doInternalPca after filter - Number of rowname
## 2020 11 18 16:18:16.410 DEBUG ab7c64738d52 doInternalPca after filter - length of batch i
## 2020 11 18 16:18:16.410 DEBUG ab7c64738d52 doInternalPca - pvalueDSC
## 2020 11 18 16:18:16.411 DEBUG ab7c64738d52 pvalueDSC start
## 2020 11 18 16:18:16.411 DEBUG ab7c64738d52 pvalueDSC transpose
## 2020 11 18 16:18:16.411 DEBUG ab7c64738d52 pvalueDSC call pvalueDSCwithExcerpt
## 2020 11 18 16:18:16.411 DEBUG ab7c64738d52 pvalueDSCwithExcerpt start
## 2020 11 18 16:18:16.411 DEBUG ab7c64738d52 nrow(thePcaDataExcerpt)= 2
## 2020 11 18 16:18:16.412 DEBUG ab7c64738d52 ncol(thePcaDataExcerpt)= 80
## 2020 11 18 16:18:16.412 DEBUG ab7c64738d52 length(theBatchIdsForSamples)= 80
## 2020 11 18 16:18:16.412 DEBUG ab7c64738d52 getDSCwithExcerpt before java
## 2020 11 18 16:18:16.413 DEBUG ab7c64738d52 getDSCwithExcerpt after java

```

```

## 2020 11 18 16:18:16.414 DEBUG ab7c64738d52 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2020 11 18 16:18:16.414 DEBUG ab7c64738d52 pvalueDSCwithExcerpt length(unique(theBatchIds
## 2020 11 18 16:18:16.414 DEBUG ab7c64738d52 pvalueDSCwithExcerpt 2222 length(permResults)=
## 2020 11 18 16:18:16.414 DEBUG ab7c64738d52 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2020 11 18 16:18:16.415 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- start
## 2020 11 18 16:18:16.415 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- redo name
## 2020 11 18 16:18:16.415 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- check file
## 2020 11 18 16:18:16.415 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- no file, write
## 2020 11 18 16:18:16.416 DEBUG ab7c64738d52 openAndWriteDscPermsFile -- nothing to write
## 2020 11 18 16:18:16.416 DEBUG ab7c64738d52 writePCAAnnotations
## 2020 11 18 16:18:16.416 DEBUG ab7c64738d52 writePCAAnnotations - data PC1,PC3
## 2020 11 18 16:18:16.419 DEBUG ab7c64738d52 writePcaAnalysis_Image - Number of rownames fo
## 2020 11 18 16:18:16.419 DEBUG ab7c64738d52 writePcaAnalysis_Image - length of batch ids f

## 2020 11 18 16:18:16.500 DEBUG ab7c64738d52 mbatchStandardLegend - Calling .jinit /tmp/Rt
## 2020 11 18 16:18:16.502 DEBUG ab7c64738d52 mbatchStandardLegend - .jinit complete
## 2020 11 18 16:18:16.502 DEBUG ab7c64738d52 mbatchStandardLegend - theTitle Rays: BatchId
## 2020 11 18 16:18:16.502 DEBUG ab7c64738d52 mbatchStandardLegend - theVersion MBatch 1.7
## 2020 11 18 16:18:16.503 DEBUG ab7c64738d52 mbatchStandardLegend - theFilenamePath /builc
## 2020 11 18 16:18:16.503 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames 00304
## 2020 11 18 16:18:16.503 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames 1
## 2020 11 18 16:18:16.504 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendColors 1
## 2020 11 18 16:18:16.504 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendSymbols 0
## 2020 11 18 16:18:16.504 DEBUG ab7c64738d52 mbatchStandardLegend - myColors #b30000
## 2020 11 18 16:18:16.504 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:16.522 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:16.524 DEBUG ab7c64738d52 mbatchStandardLegend - Calling .jinit /tmp/Rt
## 2020 11 18 16:18:16.525 DEBUG ab7c64738d52 mbatchStandardLegend - .jinit complete
## 2020 11 18 16:18:16.526 DEBUG ab7c64738d52 mbatchStandardLegend - theTitle Points: TSS
## 2020 11 18 16:18:16.526 DEBUG ab7c64738d52 mbatchStandardLegend - theVersion MBatch 1.7
## 2020 11 18 16:18:16.526 DEBUG ab7c64738d52 mbatchStandardLegend - theFilenamePath /builc
## 2020 11 18 16:18:16.527 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames OR - Un
## (72), OU - Roswell Park (1), P6 - Translational Genomics
## Research Institute (2), PA - University of Minnesota
## (1), PK - University Health
## Network (4)
## 2020 11 18 16:18:16.527 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames 5
## 2020 11 18 16:18:16.527 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendColors 5
## 2020 11 18 16:18:16.527 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendSymbols 5
## 2020 11 18 16:18:16.528 DEBUG ab7c64738d52 mbatchStandardLegend - myColors #b30000,#8fb3
## 2020 11 18 16:18:16.528 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:16.594 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:16.595 DEBUG ab7c64738d52 mbatchStandardLegend - Calling .jinit /tmp/Rt
## 2020 11 18 16:18:16.596 DEBUG ab7c64738d52 mbatchStandardLegend - .jinit complete
## 2020 11 18 16:18:16.597 DEBUG ab7c64738d52 mbatchStandardLegend - theTitle Dispersion Me
## 2020 11 18 16:18:16.597 DEBUG ab7c64738d52 mbatchStandardLegend - theVersion MBatch 1.7

```

```

## 2020 11 18 16:18:16.598 DEBUG ab7c64738d52 mbatchStandardLegend - theFilenamePath /build
## 2020 11 18 16:18:16.598 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames First P
## 2020 11 18 16:18:16.598 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendNames 14
## 2020 11 18 16:18:16.598 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendColors 0
## 2020 11 18 16:18:16.598 DEBUG ab7c64738d52 mbatchStandardLegend - theLegendSymbols 0
## 2020 11 18 16:18:16.599 DEBUG ab7c64738d52 mbatchStandardLegend - myColors
## 2020 11 18 16:18:16.599 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:16.727 DEBUG ab7c64738d52 mbatchStandardLegend after java
## 2020 11 18 16:18:16.728 DEBUG ab7c64738d52 mbatchStandardCombineLegends - Calling .jinit
## 2020 11 18 16:18:16.730 DEBUG ab7c64738d52 mbatchStandardCombineLegends - .jinit complete
## 2020 11 18 16:18:16.730 DEBUG ab7c64738d52 mbatchStandardCombineLegends - theTitle Examp
## 2020 11 18 16:18:16.731 DEBUG ab7c64738d52 mbatchStandardCombineLegends - theFilenamePat
## 2020 11 18 16:18:16.731 DEBUG ab7c64738d52 mbatchStandardCombineLegends - theListOfFiles
## 2020 11 18 16:18:16.731 DEBUG ab7c64738d52 mbatchStandardLegend before java
## 2020 11 18 16:18:16.992 DEBUG ab7c64738d52 mbatchStandardLegend after java
## [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/PCAValues.tsv"
## [23] "PlateIdwithTSS/DualBatch/ANY_Comp1_Comp2_DSC.txt"
## [24] "PlateIdwithTSS/DualBatch/ANY_Comp1_Comp2_DSC.txt__CompDSC.RData"
## [25] "PlateIdwithTSS/DualBatch/ANY_Comp1_Comp3_DSC.txt"
## [26] "PlateIdwithTSS/DualBatch/ANY_Comp1_Comp3_DSC.txt__CompDSC.RData"
## [27] "PlateIdwithTSS/DualBatch/ANY_DSC.txt"
## [28] "PlateIdwithTSS/DualBatch/ANY_DSC.txt__OverallDSC.RData"
## [29] "PlateIdwithTSS/DualBatch/ANY_FVE.txt"
## [30] "PlateIdwithTSS/DualBatch/ANY_Scores.txt"
## [31] "PlateIdwithTSS/DualBatch/ANY_Weights.txt"

```

```
## [32] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp2_Diagram.png"
## [33] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp2_Legend-ALL.png"
## [34] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp2_Legend-DSC.png"
## [35] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp2_Legend-Points.png"
## [36] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp2_Legend-Rays.png"
## [37] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp3_Diagram.png"
## [38] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp3_Legend-ALL.png"
## [39] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp3_Legend-DSC.png"
## [40] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp3_Legend-Points.png"
## [41] "PlateIdwithTSS/DualBatch/PCA-Plus/PlateIdwithTSS_Comp1_Comp3_Legend-Rays.png"
## [42] "PlateIdwithTSS/DualBatch/PCAAnnotations.tsv"
## [43] "PlateIdwithTSS/DualBatch/PCAValues.tsv"
```

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.

```
/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 subdirectory.

DSC txt files are described above in theDoDscPermsFileFlag.

ANY_Scores.txt is described above in theDoSampleLocatorFlag.

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.

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

```
/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.

```

/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

