Using MBatch Assessments: PCA_DualBatch_Structures

Tod Casasent

2023-10-06

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.

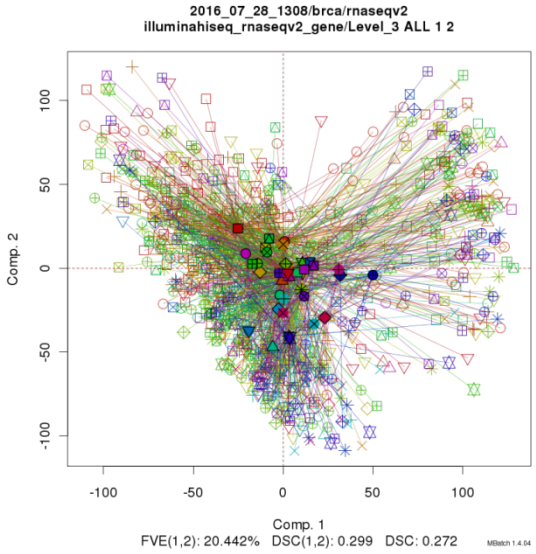

Figure 1: PCAPlus Example

4 Usage

PCA_DualBatch_Structures(theData, theTitle, theOutputPath, theBatchTypeAndValuePairsToRemove, theBatchTypeAndValuePairsToKeep, theListForDoCentroidDualBatchType, theDataVersion, theTestVersion, 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, theSeed=NULL, theMaxGeneCount=20000)

5 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.

##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.

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.

```
{
  require(MBatch)

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

  # set the paths
  theGeneFile=cleanFilePath(inputDir, "matrix_data-Tumor.tsv")
  theBatchFile=cleanFilePath(inputDir, "batches-Tumor.tsv")
  theOutputDir=cleanFilePath(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", list(), list(),
        theOutputDir=theOutputDir,
    theDataVersion="DATA_2022-09-09-1600",
    theTestVersion="TEST_2022-10-10-1300",
        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,
        theSeed=0, theMaxGeneCount=0)
    print(dir(theOutputDir, recursive=TRUE))
}
```

```
## Loading required package: MBatch

## Loading Conda Environment. MBatch Version: BEA_VERSION_TIMESTAMP

## All sorting in this package requires using a Sys.setlocale("LC_COLLATE","C"). MBatch Version: BEA_VE

## Packages uses Python environment. Set with setGlobalMBatchEnv. Curently uses /BEA/gendev

## 2023 10 06 12:31:09.739 DEBUG qcprludev10 Changing LC_COLLATE to C for duration of run
## 2023 10 06 12:31:09.740 INFO qcprludev10 \/ \/ \/ \/ \/ \/ \/ \/ \/ \/ \/ \/ \/
## 2023 10 06 12:31:09.740 INFO qcprludev10 Starting mbatchLoadFiles
## 2023 10 06 12:31:09.741 INFO qcprludev10 MBatch Version: BEA_VERSION_TIMESTAMP
## 2023 10 06 12:31:09.741 INFO qcprludev10 read batch file= /builds/BatchEffects_clean/BatchEffectsPacl
## 2023 10 06 12:31:09.743 INFO qcprludev10 read gene file= /builds/BatchEffects_clean/BatchEffectsPacka
## 2023 10 06 12:31:13.376 INFO qcprludev10 filter samples in batches using gene samples
## 2023 10 06 12:31:13.378 INFO qcprludev10 sort batches by gene file samples
## 2023 10 06 12:31:13.436 INFO qcprludev10 Finishing mbatchLoadFiles
## 2023 10 06 12:31:13.437 INFO qcprludev10 ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
## 2023 10 06 12:31:13.438 DEBUG qcprludev10 Changing LC_COLLATE to C for duration of run
## 2023 10 06 12:31:13.438 INFO qcprludev10 \/ \/ \/ \/ \/ \/ \/ \/ \/ \/ \/ \/ \/
## 2023 10 06 12:31:13.439 INFO qcprludev10 mbatchTrimData Starting
## 2023 10 06 12:31:13.439 INFO qcprludev10 MBatch Version: BEA_VERSION_TIMESTAMP
## 2023 10 06 12:31:21.025 INFO qcprludev10 mbatchTrimData theMaxSize= 1e+05
## 2023 10 06 12:31:21.026 INFO qcprludev10 mbatchTrimData ncol(theMatrix)= 80
## 2023 10 06 12:31:21.027 INFO qcprludev10 mbatchTrimData nrow(theMatrix)= 1250
## 2023 10 06 12:31:21.027 INFO qcprludev10 mbatchTrimData Finishing
## 2023 10 06 12:31:21.027 INFO qcprludev10 ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
## 2023 10 06 12:31:21.029 DEBUG qcprludev10 Changing LC_COLLATE to C for duration of run
## 2023 10 06 12:31:21.029 INFO qcprludev10 \/ \/ \/ \/ \/ \/ \/ \/ \/ \/ \/ \/ \/
## 2023 10 06 12:31:21.030 INFO qcprludev10 mbatchFilterData Starting
## 2023 10 06 12:31:21.030 INFO qcprludev10 MBatch Version: BEA_VERSION_TIMESTAMP
## 2023 10 06 12:31:21.031 DEBUG qcprludev10 rows pre filter  1250
## 2023 10 06 12:31:21.250 DEBUG qcprludev10 rows post filter  1250
## 2023 10 06 12:31:21.250 DEBUG qcprludev10 mbatchFilterData Prefilter, gene data had  1250  while pos
## 2023 10 06 12:31:21.252 DEBUG qcprludev10 mbatchFilterData Prefilter, batch data had  80  while post
## 2023 10 06 12:31:21.252 INFO qcprludev10 mbatchFilterData Finishing
## 2023 10 06 12:31:21.253 INFO qcprludev10 ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
## 2023 10 06 12:31:21.254 DEBUG qcprludev10 Changing LC_COLLATE to C for duration of run
## 2023 10 06 12:31:21.255 DEBUG qcprludev10 make future dir /BEA/BatchEffectsPackage_data/testing_dynar
## 2023 10 06 12:31:21.255 DEBUG qcprludev10 write ALL__CompListDSC.RData /BEA/BatchEffectsPackage_data/
## 2023 10 06 12:31:21.256 DEBUG qcprludev10 createBatchEffectsOutput_pca_dualBatch - start loop
## 2023 10 06 12:31:21.258 DEBUG qcprludev10 doSamplePcaCall - pre-filter matrix size= 1250,80
## 2023 10 06 12:31:21.259 DEBUG qcprludev10 filterBasedOnGeneLimit theMatrixGeneData nrow= 1250
## 2023 10 06 12:31:21.259 DEBUG qcprludev10 filterBasedOnGeneLimit theMatrixGeneData ncol= 80
## 2023 10 06 12:31:21.259 DEBUG qcprludev10 doSamplePcaCall - pre-filter matrix size= 1250,80
## 2023 10 06 12:31:21.260 DEBUG qcprludev10 doSamplePcaCall - no-NA matrix length= 1247
## 2023 10 06 12:31:21.281 DEBUG qcprludev10 doSamplePcaCall - max(theCompPairList)= 3
## 2023 10 06 12:31:21.282 DEBUG qcprludev10 doSamplePcaCall - ncol(pca@scores)= 80
## 2023 10 06 12:31:21.282 DEBUG qcprludev10 doSamplePcaCall - pca scores size= 80,80
```

```
## 2023 10 06 12:31:21.283 DEBUG qcprludev10 checkCreateDir:  /BEA/BatchEffectsPackage_data/testing_dyna
## 2023 10 06 12:31:21.283 DEBUG qcprludev10 checkCreateDir:  /BEA/BatchEffectsPackage_data/testing_dyna
## 2023 10 06 12:31:21.284 DEBUG qcprludev10 checkCreateDir:  /BEA/BatchEffectsPackage_data/testing_dyna
## 2023 10 06 12:31:21.285 DEBUG qcprludev10 checkCreateDir:  /BEA/BatchEffectsPackage_data/testing_dyna
## 2023 10 06 12:31:21.285 DEBUG qcprludev10 checkCreateDir:  /BEA/BatchEffectsPackage_data/testing_dyna
## 2023 10 06 12:31:21.286 DEBUG qcprludev10 pvalueDSC start
## 2023 10 06 12:31:21.286 DEBUG qcprludev10 pvalueDSC component A < 1
## 2023 10 06 12:31:21.287 DEBUG qcprludev10 pvalueDSC call pvalueDSCwithExcerpt
## 2023 10 06 12:31:21.288 DEBUG qcprludev10 pvalueDSCwithExcerpt start
## 2023 10 06 12:31:21.288 DEBUG qcprludev10 nrow(thePcaDataExcerpt)= 80
## 2023 10 06 12:31:21.289 DEBUG qcprludev10 ncol(thePcaDataExcerpt)= 80
## 2023 10 06 12:31:21.289 DEBUG qcprludev10 length(theBatchIdsForSamples)= 80
## 2023 10 06 12:31:21.289 DEBUG qcprludev10 getDSCwithExcerpt before Python
## 2023 10 06 12:31:21.290 DEBUG qcprludev10 getDSCwithExcerpt - import(mbatch.dsc.dsc_calc)
## 2023 10 06 12:31:22.028 DEBUG qcprludev10 getDSCwithExcerpt after Python
## 2023 10 06 12:31:22.029 DEBUG qcprludev10 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2023 10 06 12:31:22.030 DEBUG qcprludev10 pvalueDSCwithExcerpt length(unique(theBatchIdsForSamples))=
## 2023 10 06 12:31:22.030 DEBUG qcprludev10 pvalueDSCwithExcerpt 2222 length(permResults)= 1
## 2023 10 06 12:31:22.031 DEBUG qcprludev10 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2023 10 06 12:31:22.038 DEBUG qcprludev10 openAndWriteDscPermsFile -- start
## 2023 10 06 12:31:22.039 DEBUG qcprludev10 openAndWriteDscPermsFile -- redo name
## 2023 10 06 12:31:22.039 DEBUG qcprludev10 openAndWriteDscPermsFile -- check file
## 2023 10 06 12:31:22.040 DEBUG qcprludev10 openAndWriteDscPermsFile -- no file, write
## 2023 10 06 12:31:22.040 DEBUG qcprludev10 openAndWriteDscPermsFile -- nothing to write
## 2023 10 06 12:31:22.040 DEBUG qcprludev10 Write PCA Files 2  /BEA/BatchEffectsPackage_data/testing_dy
## 2023 10 06 12:31:22.041 DEBUG qcprludev10 write writePcaDataFilesForDataset
## 2023 10 06 12:31:22.041 DEBUG qcprludev10 writePCAValuesTSV
## 2023 10 06 12:31:22.042 DEBUG qcprludev10 writePCAValuesTSV title file
## 2023 10 06 12:31:22.042 DEBUG qcprludev10 writeTitleFile - pre title  Example Title for Test Data / P
## 2023 10 06 12:31:22.042 DEBUG qcprludev10 writeTitleFile - theTitle  Example Title for Test Data / PC
## 2023 10 06 12:31:22.043 DEBUG qcprludev10 writeTitleFile - titleFile  /BEA/BatchEffectsPackage_data/t
## 2023 10 06 12:31:22.044 DEBUG qcprludev10 writePCAValuesTSV outputFile
## 2023 10 06 12:31:22.044 DEBUG qcprludev10 writePCAValuesTSV components
## 2023 10 06 12:31:22.045 DEBUG qcprludev10 writePCAValuesTSV myDataFrame
## 2023 10 06 12:31:22.045 DEBUG qcprludev10 writePCAValuesTSV - values
## 2023 10 06 12:31:22.046 DEBUG qcprludev10 writePCAValuesTSV - matrix
## 2023 10 06 12:31:22.047 DEBUG qcprludev10 writePCAValuesTSV - write
## 2023 10 06 12:31:22.049 DEBUG qcprludev10 writePCAValuesTSV finished  /BEA/BatchEffectsPackage_data/t
## 2023 10 06 12:31:22.050 DEBUG qcprludev10 writePCAAnnotations
## 2023 10 06 12:31:22.050 DEBUG qcprludev10 writePCAAnnotations - header and all data
## 2023 10 06 12:31:22.596 DEBUG qcprludev10 doInternalPca before filter - Number of rownames for pca sc
## 2023 10 06 12:31:22.597 DEBUG qcprludev10 doInternalPca before filter - length of batch ids for sampl
## 2023 10 06 12:31:22.598 DEBUG qcprludev10 doInternalPca after filter - Number of rownames for pca sco
## 2023 10 06 12:31:22.598 DEBUG qcprludev10 doInternalPca after filter - length of batch ids for sample
## 2023 10 06 12:31:22.598 DEBUG qcprludev10 doInternalPca - pvalueDSC
## 2023 10 06 12:31:22.599 DEBUG qcprludev10 pvalueDSC start
## 2023 10 06 12:31:22.599 DEBUG qcprludev10 pvalueDSC transpose
## 2023 10 06 12:31:22.600 DEBUG qcprludev10 pvalueDSC call pvalueDSCwithExcerpt
## 2023 10 06 12:31:22.600 DEBUG qcprludev10 pvalueDSCwithExcerpt start
## 2023 10 06 12:31:22.600 DEBUG qcprludev10 nrow(thePcaDataExcerpt)= 2
## 2023 10 06 12:31:22.601 DEBUG qcprludev10 ncol(thePcaDataExcerpt)= 80
## 2023 10 06 12:31:22.601 DEBUG qcprludev10 length(theBatchIdsForSamples)= 80
## 2023 10 06 12:31:22.602 DEBUG qcprludev10 getDSCwithExcerpt before Python
## 2023 10 06 12:31:22.602 DEBUG qcprludev10 getDSCwithExcerpt - import(mbatch.dsc.dsc_calc)
```

```
## 2023 10 06 12:31:22.604 DEBUG qcprludev10 getDSCwithExcerpt after Python
## 2023 10 06 12:31:22.605 DEBUG qcprludev10 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2023 10 06 12:31:22.606 DEBUG qcprludev10 pvalueDSCwithExcerpt length(unique(theBatchIdsForSamples))=
## 2023 10 06 12:31:22.606 DEBUG qcprludev10 pvalueDSCwithExcerpt 2222 length(permResults)= 1
## 2023 10 06 12:31:22.606 DEBUG qcprludev10 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2023 10 06 12:31:22.607 DEBUG qcprludev10 openAndWriteDscPermsFile -- start
## 2023 10 06 12:31:22.607 DEBUG qcprludev10 openAndWriteDscPermsFile -- redo name
## 2023 10 06 12:31:22.608 DEBUG qcprludev10 openAndWriteDscPermsFile -- check file
## 2023 10 06 12:31:22.608 DEBUG qcprludev10 openAndWriteDscPermsFile -- no file, write
## 2023 10 06 12:31:22.609 DEBUG qcprludev10 openAndWriteDscPermsFile -- nothing to write
## 2023 10 06 12:31:22.611 DEBUG qcprludev10 write PCA file direct 1  /BEA/BatchEffectsPackage_data/tes
## 2023 10 06 12:31:22.614 DEBUG qcprludev10 writePCAAnnotations
## 2023 10 06 12:31:22.615 DEBUG qcprludev10 writePCAAnnotations - data  PC1,PC2
## 2023 10 06 12:31:22.630 DEBUG qcprludev10 writePcaAnalysis_Image - Number of rownames for pca score
## 2023 10 06 12:31:22.630 DEBUG qcprludev10 writePcaAnalysis_Image - length of batch ids for samples  8

## 2023 10 06 12:31:22.732 DEBUG qcprludev10 mbatchStandardLegend - theTitle  Rays: PlateId (80)
## 2023 10 06 12:31:22.733 DEBUG qcprludev10 mbatchStandardLegend - theVersion  MBatch 2.0.3
## 2023 10 06 12:31:22.734 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath  /BEA/BatchEffectsPa
## 2023 10 06 12:31:22.734 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames  A29J (80)
## 2023 10 06 12:31:22.735 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames length  1
## 2023 10 06 12:31:22.735 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors  #B30000
## 2023 10 06 12:31:22.736 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors length  1
## 2023 10 06 12:31:22.736 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols
## 2023 10 06 12:31:22.736 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols length  0
## 2023 10 06 12:31:22.737 DEBUG qcprludev10 mbatchStandardLegend - myColors  #b30000
## 2023 10 06 12:31:22.737 DEBUG qcprludev10 mbatchStandardLegend - theTitle UTF-8 =  Rays: PlateId (80)
## 2023 10 06 12:31:22.737 DEBUG qcprludev10 mbatchStandardLegend - theVersion UTF-8 =  MBatch 2.0.3
## 2023 10 06 12:31:22.738 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath UTF-8 =  /BEA/Batch
## 2023 10 06 12:31:22.738 DEBUG qcprludev10 mbatchStandardLegend before Python
## 2023 10 06 12:31:22.739 DEBUG qcprludev10 mbatchStandardLegend - getGlobalMBatchEnv() =  /BEA/gendev
## 2023 10 06 12:31:22.739 DEBUG qcprludev10 mbatchStandardLegend - import(mbatch.legend.legend)
## 2023 10 06 12:31:23.081 DEBUG qcprludev10 mbatchStandardLegend - after import
## 2023 10 06 12:31:23.082 DEBUG qcprludev10 mbatchStandardLegend - after color list
## 2023 10 06 12:31:23.083 DEBUG qcprludev10 mbatchStandardLegend - after symbol list
## 2023 10 06 12:31:23.083 DEBUG qcprludev10 mbatchStandardLegend - colorList =  #b30000
## 2023 10 06 12:31:23.083 DEBUG qcprludev10 mbatchStandardLegend - symbolList =
## 2023 10 06 12:31:23.084 DEBUG qcprludev10 mbatchStandardLegend - legendNameList =  A29J (80)
## 2023 10 06 12:31:23.084 DEBUG qcprludev10 mbatchStandardLegend - myTitle =  Rays: PlateId (80) MBatch
## 2023 10 06 12:31:23.085 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath =  /BEA/BatchEffects
## 2023 10 06 12:31:23.210 DEBUG qcprludev10 mbatchStandardLegend after Python
## 2023 10 06 12:31:23.212 DEBUG qcprludev10 mbatchStandardLegend - theTitle  Points: TSS (80)
## 2023 10 06 12:31:23.213 DEBUG qcprludev10 mbatchStandardLegend - theVersion  MBatch 2.0.3
## 2023 10 06 12:31:23.213 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath  /BEA/BatchEffectsPa
## 2023 10 06 12:31:23.214 DEBUG qcprludev10 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)
## 2023 10 06 12:31:23.214 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames length  5
## 2023 10 06 12:31:23.215 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors  #B30000, #8FB300, #
## 2023 10 06 12:31:23.215 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors length  5
## 2023 10 06 12:31:23.216 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols  0, 1, 2, 3, 4
## 2023 10 06 12:31:23.216 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols length  5
```

```
## 2023 10 06 12:31:23.216 DEBUG qcprludev10 mbatchStandardLegend - myColors  #b30000,#8fb300,#00b347,#0
## 2023 10 06 12:31:23.217 DEBUG qcprludev10 mbatchStandardLegend - theTitle UTF-8 =  Points: TSS (80)
## 2023 10 06 12:31:23.217 DEBUG qcprludev10 mbatchStandardLegend - theVersion UTF-8 =  MBatch 2.0.3
## 2023 10 06 12:31:23.218 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath UTF-8 =  /BEA/Batchl
## 2023 10 06 12:31:23.218 DEBUG qcprludev10 mbatchStandardLegend before Python
## 2023 10 06 12:31:23.218 DEBUG qcprludev10 mbatchStandardLegend - getGlobalMBatchEnv() =  /BEA/gendev
## 2023 10 06 12:31:23.219 DEBUG qcprludev10 mbatchStandardLegend - import(mbatch.legend.legend)
## 2023 10 06 12:31:23.219 DEBUG qcprludev10 mbatchStandardLegend - after import
## 2023 10 06 12:31:23.219 DEBUG qcprludev10 mbatchStandardLegend - after color list
## 2023 10 06 12:31:23.220 DEBUG qcprludev10 mbatchStandardLegend - after symbol list
## 2023 10 06 12:31:23.220 DEBUG qcprludev10 mbatchStandardLegend - colorList =  #b30000, mbatchStandard
## 2023 10 06 12:31:23.221 DEBUG qcprludev10 mbatchStandardLegend - symbolList =  0, mbatchStandardLeger
## 2023 10 06 12:31:23.221 DEBUG qcprludev10 mbatchStandardLegend - legendNameList =  OR - University o
## (72), mbatchStandardLegend - legendNameList =  OU - Roswell Park (1), mbatchStandardLegend - legendNa
## Research Institute (2), mbatchStandardLegend - legendNameList =  PA - University of Minnesota
## (1), mbatchStandardLegend - legendNameList =  PK - University Health
## Network (4)
## 2023 10 06 12:31:23.221 DEBUG qcprludev10 mbatchStandardLegend - myTitle =  Points: TSS (80) MBatch 1
## 2023 10 06 12:31:23.222 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath =  /BEA/BatchEffects
## 2023 10 06 12:31:23.341 DEBUG qcprludev10 mbatchStandardLegend after Python
## 2023 10 06 12:31:23.342 DEBUG qcprludev10 mbatchStandardLegend - theTitle  Dispersion Metrics
## 2023 10 06 12:31:23.342 DEBUG qcprludev10 mbatchStandardLegend - theVersion  MBatch 2.0.3
## 2023 10 06 12:31:23.343 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath  /BEA/BatchEffectsPa
## 2023 10 06 12:31:23.343 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames  First PCA Component
## 2023 10 06 12:31:23.344 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames length  14
## 2023 10 06 12:31:23.344 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors
## 2023 10 06 12:31:23.344 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors length  0
## 2023 10 06 12:31:23.345 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols
## 2023 10 06 12:31:23.345 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols length  0
## 2023 10 06 12:31:23.345 DEBUG qcprludev10 mbatchStandardLegend - myColors
## 2023 10 06 12:31:23.346 DEBUG qcprludev10 mbatchStandardLegend - theTitle UTF-8 =  Dispersion Metrics
## 2023 10 06 12:31:23.346 DEBUG qcprludev10 mbatchStandardLegend - theVersion UTF-8 =  MBatch 2.0.3
## 2023 10 06 12:31:23.347 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath UTF-8 =  /BEA/Batchl
## 2023 10 06 12:31:23.347 DEBUG qcprludev10 mbatchStandardLegend before Python
## 2023 10 06 12:31:23.347 DEBUG qcprludev10 mbatchStandardLegend - getGlobalMBatchEnv() =  /BEA/gendev
## 2023 10 06 12:31:23.348 DEBUG qcprludev10 mbatchStandardLegend - import(mbatch.legend.legend)
## 2023 10 06 12:31:23.348 DEBUG qcprludev10 mbatchStandardLegend - after import
## 2023 10 06 12:31:23.348 DEBUG qcprludev10 mbatchStandardLegend - after color list
## 2023 10 06 12:31:23.349 DEBUG qcprludev10 mbatchStandardLegend - after symbol list
## 2023 10 06 12:31:23.349 DEBUG qcprludev10 mbatchStandardLegend - colorList =
## 2023 10 06 12:31:23.350 DEBUG qcprludev10 mbatchStandardLegend - symbolList =
## 2023 10 06 12:31:23.350 DEBUG qcprludev10 mbatchStandardLegend - legendNameList =  First PCA Componer
## 2023 10 06 12:31:23.350 DEBUG qcprludev10 mbatchStandardLegend - myTitle =  Dispersion Metrics MBatch
## 2023 10 06 12:31:23.351 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath =  /BEA/BatchEffects
## 2023 10 06 12:31:23.375 DEBUG qcprludev10 mbatchStandardLegend after Python
## 2023 10 06 12:31:23.376 DEBUG qcprludev10 mbatchStandardCombineLegends - theTitle  Example Title for
## 2023 10 06 12:31:23.376 DEBUG qcprludev10 mbatchStandardCombineLegends - theFilenamePath  /BEA/Batchl
## 2023 10 06 12:31:23.377 DEBUG qcprludev10 mbatchStandardCombineLegends - theListOfFiles  /BEA/BatchEf
## 2023 10 06 12:31:23.377 DEBUG qcprludev10 mbatchStandardCombineLegends - theTitle UTF-8 =  Example Ti
## 2023 10 06 12:31:23.377 DEBUG qcprludev10 mbatchStandardCombineLegends - theFilenamePath UTF-8 =  /BE
## 2023 10 06 12:31:23.378 DEBUG qcprludev10 mbatchStandardCombineLegends before Python
## 2023 10 06 12:31:23.378 DEBUG qcprludev10 mbatchStandardCombineLegends - import(mbatch.legend.legend)
## 2023 10 06 12:31:23.442 DEBUG qcprludev10 mbatchStandardCombineLegends after Python
## 2023 10 06 12:31:23.445 DEBUG qcprludev10 doInternalPca before filter - Number of rownames for pca s
```

```
## 2023 10 06 12:31:23.446 DEBUG qcprludev10 doInternalPca before filter - length of batch ids for sampl
## 2023 10 06 12:31:23.446 DEBUG qcprludev10 doInternalPca after filter - Number of rownames for pca sc
## 2023 10 06 12:31:23.446 DEBUG qcprludev10 doInternalPca after filter - length of batch ids for sample
## 2023 10 06 12:31:23.447 DEBUG qcprludev10 doInternalPca - pvalueDSC
## 2023 10 06 12:31:23.447 DEBUG qcprludev10 pvalueDSC start
## 2023 10 06 12:31:23.447 DEBUG qcprludev10 pvalueDSC transpose
## 2023 10 06 12:31:23.448 DEBUG qcprludev10 pvalueDSC call pvalueDSCwithExcerpt
## 2023 10 06 12:31:23.448 DEBUG qcprludev10 pvalueDSCwithExcerpt start
## 2023 10 06 12:31:23.449 DEBUG qcprludev10 nrow(thePcaDataExcerpt)= 2
## 2023 10 06 12:31:23.449 DEBUG qcprludev10 ncol(thePcaDataExcerpt)= 80
## 2023 10 06 12:31:23.449 DEBUG qcprludev10 length(theBatchIdsForSamples)= 80
## 2023 10 06 12:31:23.450 DEBUG qcprludev10 getDSCwithExcerpt before Python
## 2023 10 06 12:31:23.450 DEBUG qcprludev10 getDSCwithExcerpt - import(mbatch.dsc.dsc_calc)
## 2023 10 06 12:31:23.452 DEBUG qcprludev10 getDSCwithExcerpt after Python
## 2023 10 06 12:31:23.453 DEBUG qcprludev10 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2023 10 06 12:31:23.453 DEBUG qcprludev10 pvalueDSCwithExcerpt length(unique(theBatchIdsForSamples))=
## 2023 10 06 12:31:23.454 DEBUG qcprludev10 pvalueDSCwithExcerpt 2222 length(permResults)= 1
## 2023 10 06 12:31:23.454 DEBUG qcprludev10 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2023 10 06 12:31:23.455 DEBUG qcprludev10 openAndWriteDscPermsFile -- start
## 2023 10 06 12:31:23.455 DEBUG qcprludev10 openAndWriteDscPermsFile -- redo name
## 2023 10 06 12:31:23.456 DEBUG qcprludev10 openAndWriteDscPermsFile -- check file
## 2023 10 06 12:31:23.456 DEBUG qcprludev10 openAndWriteDscPermsFile -- no file, write
## 2023 10 06 12:31:23.456 DEBUG qcprludev10 openAndWriteDscPermsFile -- nothing to write
## 2023 10 06 12:31:23.457 DEBUG qcprludev10 write PCA file direct 1  /BEA/BatchEffectsPackage_data/tes
## 2023 10 06 12:31:23.458 DEBUG qcprludev10 writePCAAnnotations
## 2023 10 06 12:31:23.458 DEBUG qcprludev10 writePCAAnnotations - data  PC1,PC3
## 2023 10 06 12:31:23.460 DEBUG qcprludev10 writePcaAnalysis_Image - Number of rownames for pca score
## 2023 10 06 12:31:23.461 DEBUG qcprludev10 writePcaAnalysis_Image - length of batch ids for samples  8

## 2023 10 06 12:31:23.534 DEBUG qcprludev10 mbatchStandardLegend - theTitle  Rays: PlateId (80)
## 2023 10 06 12:31:23.534 DEBUG qcprludev10 mbatchStandardLegend - theVersion  MBatch 2.0.3
## 2023 10 06 12:31:23.535 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath  /BEA/BatchEffectsPa
## 2023 10 06 12:31:23.535 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames  A29J (80)
## 2023 10 06 12:31:23.536 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames length  1
## 2023 10 06 12:31:23.536 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors  #B30000
## 2023 10 06 12:31:23.537 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors length  1
## 2023 10 06 12:31:23.537 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols
## 2023 10 06 12:31:23.537 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols length  0
## 2023 10 06 12:31:23.538 DEBUG qcprludev10 mbatchStandardLegend - myColors  #b30000
## 2023 10 06 12:31:23.538 DEBUG qcprludev10 mbatchStandardLegend - theTitle UTF-8 =  Rays: PlateId (80)
## 2023 10 06 12:31:23.538 DEBUG qcprludev10 mbatchStandardLegend - theVersion UTF-8 =  MBatch 2.0.3
## 2023 10 06 12:31:23.539 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath UTF-8 =  /BEA/Batchl
## 2023 10 06 12:31:23.539 DEBUG qcprludev10 mbatchStandardLegend before Python
## 2023 10 06 12:31:23.539 DEBUG qcprludev10 mbatchStandardLegend - getGlobalMBatchEnv() =  /BEA/gendev
## 2023 10 06 12:31:23.540 DEBUG qcprludev10 mbatchStandardLegend - import(mbatch.legend.legend)
## 2023 10 06 12:31:23.540 DEBUG qcprludev10 mbatchStandardLegend - after import
## 2023 10 06 12:31:23.541 DEBUG qcprludev10 mbatchStandardLegend - after color list
## 2023 10 06 12:31:23.541 DEBUG qcprludev10 mbatchStandardLegend - after symbol list
## 2023 10 06 12:31:23.541 DEBUG qcprludev10 mbatchStandardLegend - colorList =  #b30000
## 2023 10 06 12:31:23.542 DEBUG qcprludev10 mbatchStandardLegend - symbolList =
## 2023 10 06 12:31:23.542 DEBUG qcprludev10 mbatchStandardLegend - legendNameList =  A29J (80)
## 2023 10 06 12:31:23.542 DEBUG qcprludev10 mbatchStandardLegend - myTitle =  Rays: PlateId (80) MBatch
## 2023 10 06 12:31:23.543 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath =  /BEA/BatchEffects
## 2023 10 06 12:31:23.582 DEBUG qcprludev10 mbatchStandardLegend after Python
```

```
## 2023 10 06 12:31:23.583 DEBUG qcprludev10 mbatchStandardLegend - theTitle  Points: TSS (80)
## 2023 10 06 12:31:23.584 DEBUG qcprludev10 mbatchStandardLegend - theVersion  MBatch 2.0.3
## 2023 10 06 12:31:23.584 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath  /BEA/BatchEffectsPa
## 2023 10 06 12:31:23.585 DEBUG qcprludev10 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)
## 2023 10 06 12:31:23.585 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames length  5
## 2023 10 06 12:31:23.586 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors  #B30000, #8FB300, #
## 2023 10 06 12:31:23.586 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors length  5
## 2023 10 06 12:31:23.586 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols  0, 1, 2, 3, 4
## 2023 10 06 12:31:23.587 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols length  5
## 2023 10 06 12:31:23.587 DEBUG qcprludev10 mbatchStandardLegend - myColors  #b30000,#8fb300,#00b347,#0
## 2023 10 06 12:31:23.588 DEBUG qcprludev10 mbatchStandardLegend - theTitle UTF-8 =  Points: TSS (80)
## 2023 10 06 12:31:23.588 DEBUG qcprludev10 mbatchStandardLegend - theVersion UTF-8 =  MBatch 2.0.3
## 2023 10 06 12:31:23.588 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath UTF-8 =  /BEA/Batchl
## 2023 10 06 12:31:23.589 DEBUG qcprludev10 mbatchStandardLegend before Python
## 2023 10 06 12:31:23.589 DEBUG qcprludev10 mbatchStandardLegend - getGlobalMBatchEnv() =  /BEA/gendev
## 2023 10 06 12:31:23.590 DEBUG qcprludev10 mbatchStandardLegend - import(mbatch.legend.legend)
## 2023 10 06 12:31:23.590 DEBUG qcprludev10 mbatchStandardLegend - after import
## 2023 10 06 12:31:23.591 DEBUG qcprludev10 mbatchStandardLegend - after color list
## 2023 10 06 12:31:23.591 DEBUG qcprludev10 mbatchStandardLegend - after symbol list
## 2023 10 06 12:31:23.591 DEBUG qcprludev10 mbatchStandardLegend - colorList =  #b30000, mbatchStandard
## 2023 10 06 12:31:23.592 DEBUG qcprludev10 mbatchStandardLegend - symbolList =  0, mbatchStandardLeger
## 2023 10 06 12:31:23.592 DEBUG qcprludev10 mbatchStandardLegend - legendNameList =  OR - University of
## (72), mbatchStandardLegend - legendNameList =  OU - Roswell Park (1), mbatchStandardLegend - legendNa
## Research Institute (2), mbatchStandardLegend - legendNameList =  PA - University of Minnesota
## (1), mbatchStandardLegend - legendNameList =  PK - University Health
## Network (4)
## 2023 10 06 12:31:23.593 DEBUG qcprludev10 mbatchStandardLegend - myTitle =  Points: TSS (80) MBatch 2
## 2023 10 06 12:31:23.593 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath =  /BEA/BatchEffects
## 2023 10 06 12:31:23.717 DEBUG qcprludev10 mbatchStandardLegend after Python
## 2023 10 06 12:31:23.718 DEBUG qcprludev10 mbatchStandardLegend - theTitle  Dispersion Metrics
## 2023 10 06 12:31:23.718 DEBUG qcprludev10 mbatchStandardLegend - theVersion  MBatch 2.0.3
## 2023 10 06 12:31:23.719 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath  /BEA/BatchEffectsPa
## 2023 10 06 12:31:23.720 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames  First PCA Component
## 2023 10 06 12:31:23.720 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames length  14
## 2023 10 06 12:31:23.720 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors
## 2023 10 06 12:31:23.721 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors length  0
## 2023 10 06 12:31:23.721 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols
## 2023 10 06 12:31:23.722 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols length  0
## 2023 10 06 12:31:23.722 DEBUG qcprludev10 mbatchStandardLegend - myColors
## 2023 10 06 12:31:23.722 DEBUG qcprludev10 mbatchStandardLegend - theTitle UTF-8 =  Dispersion Metrics
## 2023 10 06 12:31:23.723 DEBUG qcprludev10 mbatchStandardLegend - theVersion UTF-8 =  MBatch 2.0.3
## 2023 10 06 12:31:23.723 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath UTF-8 =  /BEA/Batchl
## 2023 10 06 12:31:23.723 DEBUG qcprludev10 mbatchStandardLegend before Python
## 2023 10 06 12:31:23.724 DEBUG qcprludev10 mbatchStandardLegend - getGlobalMBatchEnv() =  /BEA/gendev
## 2023 10 06 12:31:23.724 DEBUG qcprludev10 mbatchStandardLegend - import(mbatch.legend.legend)
## 2023 10 06 12:31:23.725 DEBUG qcprludev10 mbatchStandardLegend - after import
## 2023 10 06 12:31:23.725 DEBUG qcprludev10 mbatchStandardLegend - after color list
## 2023 10 06 12:31:23.725 DEBUG qcprludev10 mbatchStandardLegend - after symbol list
## 2023 10 06 12:31:23.726 DEBUG qcprludev10 mbatchStandardLegend - colorList =
## 2023 10 06 12:31:23.726 DEBUG qcprludev10 mbatchStandardLegend - symbolList =
```

```
## 2023 10 06 12:31:23.727 DEBUG qcprludev10 mbatchStandardLegend - legendNameList =  First PCA Componer
## 2023 10 06 12:31:23.727 DEBUG qcprludev10 mbatchStandardLegend - myTitle =  Dispersion Metrics MBatch
## 2023 10 06 12:31:23.727 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath =  /BEA/BatchEffects
## 2023 10 06 12:31:23.753 DEBUG qcprludev10 mbatchStandardLegend after Python
## 2023 10 06 12:31:23.753 DEBUG qcprludev10 mbatchStandardCombineLegends - theTitle  Example Title for
## 2023 10 06 12:31:23.754 DEBUG qcprludev10 mbatchStandardCombineLegends - theFilenamePath  /BEA/Batch
## 2023 10 06 12:31:23.754 DEBUG qcprludev10 mbatchStandardCombineLegends - theListOfFiles  /BEA/BatchE
## 2023 10 06 12:31:23.755 DEBUG qcprludev10 mbatchStandardCombineLegends - theTitle UTF-8 =  Example T
## 2023 10 06 12:31:23.755 DEBUG qcprludev10 mbatchStandardCombineLegends - theFilenamePath UTF-8 =  /B
## 2023 10 06 12:31:23.755 DEBUG qcprludev10 mbatchStandardCombineLegends before Python
## 2023 10 06 12:31:23.756 DEBUG qcprludev10 mbatchStandardCombineLegends - import(mbatch.legend.legend
## 2023 10 06 12:31:23.819 DEBUG qcprludev10 mbatchStandardCombineLegends after Python
## 2023 10 06 12:31:23.820 DEBUG qcprludev10 doSamplePcaCall - pre-filter matrix size= 1250,80
## 2023 10 06 12:31:23.821 DEBUG qcprludev10 filterBasedOnGeneLimit theMatrixGeneData nrow= 1250
## 2023 10 06 12:31:23.821 DEBUG qcprludev10 filterBasedOnGeneLimit theMatrixGeneData ncol= 80
## 2023 10 06 12:31:23.821 DEBUG qcprludev10 doSamplePcaCall - pre-filter matrix size= 1250,80
## 2023 10 06 12:31:23.822 DEBUG qcprludev10 doSamplePcaCall - no-NA matrix length= 1247
## 2023 10 06 12:31:23.840 DEBUG qcprludev10 doSamplePcaCall - max(theCompPairList)= 3
## 2023 10 06 12:31:23.841 DEBUG qcprludev10 doSamplePcaCall - ncol(pca@scores)= 80
## 2023 10 06 12:31:23.841 DEBUG qcprludev10 doSamplePcaCall - pca scores size= 80,80
## 2023 10 06 12:31:23.843 DEBUG qcprludev10 checkCreateDir:  /BEA/BatchEffectsPackage_data/testing_dyna
## 2023 10 06 12:31:23.843 DEBUG qcprludev10 checkCreateDir:  /BEA/BatchEffectsPackage_data/testing_dyna
## 2023 10 06 12:31:23.844 DEBUG qcprludev10 checkCreateDir:  /BEA/BatchEffectsPackage_data/testing_dyna
## 2023 10 06 12:31:23.845 DEBUG qcprludev10 checkCreateDir:  /BEA/BatchEffectsPackage_data/testing_dyna
## 2023 10 06 12:31:23.845 DEBUG qcprludev10 checkCreateDir:  /BEA/BatchEffectsPackage_data/testing_dyna
## 2023 10 06 12:31:23.846 DEBUG qcprludev10 pvalueDSC start
## 2023 10 06 12:31:23.846 DEBUG qcprludev10 pvalueDSC component A < 1
## 2023 10 06 12:31:23.846 DEBUG qcprludev10 pvalueDSC call pvalueDSCwithExcerpt
## 2023 10 06 12:31:23.847 DEBUG qcprludev10 pvalueDSCwithExcerpt start
## 2023 10 06 12:31:23.847 DEBUG qcprludev10 nrow(thePcaDataExcerpt)= 80
## 2023 10 06 12:31:23.848 DEBUG qcprludev10 ncol(thePcaDataExcerpt)= 80
## 2023 10 06 12:31:23.848 DEBUG qcprludev10 length(theBatchIdsForSamples)= 80
## 2023 10 06 12:31:23.848 DEBUG qcprludev10 getDSCwithExcerpt before Python
## 2023 10 06 12:31:23.849 DEBUG qcprludev10 getDSCwithExcerpt - import(mbatch.dsc.dsc_calc)
## 2023 10 06 12:31:23.858 DEBUG qcprludev10 getDSCwithExcerpt after Python
## 2023 10 06 12:31:23.859 DEBUG qcprludev10 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2023 10 06 12:31:23.860 DEBUG qcprludev10 pvalueDSCwithExcerpt length(unique(theBatchIdsForSamples))=
## 2023 10 06 12:31:23.860 DEBUG qcprludev10 pvalueDSCwithExcerpt 2222 length(permResults)= 1
## 2023 10 06 12:31:23.861 DEBUG qcprludev10 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2023 10 06 12:31:23.868 DEBUG qcprludev10 openAndWriteDscPermsFile -- start
## 2023 10 06 12:31:23.868 DEBUG qcprludev10 openAndWriteDscPermsFile -- redo name
## 2023 10 06 12:31:23.869 DEBUG qcprludev10 openAndWriteDscPermsFile -- check file
## 2023 10 06 12:31:23.869 DEBUG qcprludev10 openAndWriteDscPermsFile -- no file, write
## 2023 10 06 12:31:23.870 DEBUG qcprludev10 openAndWriteDscPermsFile -- nothing to write
## 2023 10 06 12:31:23.870 DEBUG qcprludev10 Write PCA Files 2  /BEA/BatchEffectsPackage_data/testing_dy
## 2023 10 06 12:31:23.870 DEBUG qcprludev10 write writePcaDataFilesForDataset
## 2023 10 06 12:31:23.871 DEBUG qcprludev10 writePCAValuesTSV
## 2023 10 06 12:31:23.871 DEBUG qcprludev10 writePCAValuesTSV title file
## 2023 10 06 12:31:23.872 DEBUG qcprludev10 writeTitleFile - pre title  Example Title for Test Data / P
## 2023 10 06 12:31:23.872 DEBUG qcprludev10 writeTitleFile - theTitle  Example Title for Test Data / PC
## 2023 10 06 12:31:23.873 DEBUG qcprludev10 writeTitleFile - titleFile  /BEA/BatchEffectsPackage_data/t
## 2023 10 06 12:31:23.873 DEBUG qcprludev10 writePCAValuesTSV outputFile
## 2023 10 06 12:31:23.874 DEBUG qcprludev10 writePCAValuesTSV components
## 2023 10 06 12:31:23.874 DEBUG qcprludev10 writePCAValuesTSV myDataFrame
```

```
## 2023 10 06 12:31:23.874 DEBUG qcprludev10 writePCAValuesTSV - values
## 2023 10 06 12:31:23.875 DEBUG qcprludev10 writePCAValuesTSV - matrix
## 2023 10 06 12:31:23.876 DEBUG qcprludev10 writePCAValuesTSV - write
## 2023 10 06 12:31:23.878 DEBUG qcprludev10 writePCAValuesTSV finished  /BEA/BatchEffectsPackage_data/
## 2023 10 06 12:31:23.878 DEBUG qcprludev10 writePCAAnnotations
## 2023 10 06 12:31:23.879 DEBUG qcprludev10 writePCAAnnotations - header and all data
## 2023 10 06 12:31:24.072 DEBUG qcprludev10 doInternalPca before filter - Number of rownames for pca s
## 2023 10 06 12:31:24.073 DEBUG qcprludev10 doInternalPca before filter - length of batch ids for samp
## 2023 10 06 12:31:24.073 DEBUG qcprludev10 doInternalPca after filter - Number of rownames for pca sc
## 2023 10 06 12:31:24.074 DEBUG qcprludev10 doInternalPca after filter - length of batch ids for sampl
## 2023 10 06 12:31:24.074 DEBUG qcprludev10 doInternalPca - pvalueDSC
## 2023 10 06 12:31:24.075 DEBUG qcprludev10 pvalueDSC start
## 2023 10 06 12:31:24.075 DEBUG qcprludev10 pvalueDSC transpose
## 2023 10 06 12:31:24.076 DEBUG qcprludev10 pvalueDSC call pvalueDSCwithExcerpt
## 2023 10 06 12:31:24.076 DEBUG qcprludev10 pvalueDSCwithExcerpt start
## 2023 10 06 12:31:24.076 DEBUG qcprludev10 nrow(thePcaDataExcerpt)= 2
## 2023 10 06 12:31:24.077 DEBUG qcprludev10 ncol(thePcaDataExcerpt)= 80
## 2023 10 06 12:31:24.077 DEBUG qcprludev10 length(theBatchIdsForSamples)= 80
## 2023 10 06 12:31:24.078 DEBUG qcprludev10 getDSCwithExcerpt before Python
## 2023 10 06 12:31:24.078 DEBUG qcprludev10 getDSCwithExcerpt - import(mbatch.dsc.dsc_calc)
## 2023 10 06 12:31:24.080 DEBUG qcprludev10 getDSCwithExcerpt after Python
## 2023 10 06 12:31:24.081 DEBUG qcprludev10 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2023 10 06 12:31:24.081 DEBUG qcprludev10 pvalueDSCwithExcerpt length(unique(theBatchIdsForSamples))=
## 2023 10 06 12:31:24.082 DEBUG qcprludev10 pvalueDSCwithExcerpt 2222 length(permResults)= 1
## 2023 10 06 12:31:24.082 DEBUG qcprludev10 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2023 10 06 12:31:24.083 DEBUG qcprludev10 openAndWriteDscPermsFile -- start
## 2023 10 06 12:31:24.083 DEBUG qcprludev10 openAndWriteDscPermsFile -- redo name
## 2023 10 06 12:31:24.085 DEBUG qcprludev10 openAndWriteDscPermsFile -- check file
## 2023 10 06 12:31:24.085 DEBUG qcprludev10 openAndWriteDscPermsFile -- no file, write
## 2023 10 06 12:31:24.086 DEBUG qcprludev10 openAndWriteDscPermsFile -- nothing to write
## 2023 10 06 12:31:24.086 DEBUG qcprludev10 write PCA file direct 1  /BEA/BatchEffectsPackage_data/tes
## 2023 10 06 12:31:24.087 DEBUG qcprludev10 writePCAAnnotations
## 2023 10 06 12:31:24.087 DEBUG qcprludev10 writePCAAnnotations - data  PC1,PC2
## 2023 10 06 12:31:24.090 DEBUG qcprludev10 writePcaAnalysis_Image - Number of rownames for pca score
## 2023 10 06 12:31:24.090 DEBUG qcprludev10 writePcaAnalysis_Image - length of batch ids for samples  8

## 2023 10 06 12:31:24.166 DEBUG qcprludev10 mbatchStandardLegend - theTitle  Rays: BatchId (80)
## 2023 10 06 12:31:24.167 DEBUG qcprludev10 mbatchStandardLegend - theVersion  MBatch 2.0.3
## 2023 10 06 12:31:24.168 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath  /BEA/BatchEffectsPa
## 2023 10 06 12:31:24.168 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames  00304 (80)
## 2023 10 06 12:31:24.168 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames length  1
## 2023 10 06 12:31:24.169 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors  #B30000
## 2023 10 06 12:31:24.169 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors length  1
## 2023 10 06 12:31:24.170 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols
## 2023 10 06 12:31:24.170 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols length  0
## 2023 10 06 12:31:24.170 DEBUG qcprludev10 mbatchStandardLegend - myColors  #b30000
## 2023 10 06 12:31:24.171 DEBUG qcprludev10 mbatchStandardLegend - theTitle UTF-8 =  Rays: BatchId (80)
## 2023 10 06 12:31:24.171 DEBUG qcprludev10 mbatchStandardLegend - theVersion UTF-8 =  MBatch 2.0.3
## 2023 10 06 12:31:24.172 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath UTF-8 =  /BEA/Batch
## 2023 10 06 12:31:24.172 DEBUG qcprludev10 mbatchStandardLegend before Python
## 2023 10 06 12:31:24.172 DEBUG qcprludev10 mbatchStandardLegend - getGlobalMBatchEnv() =  /BEA/gendev
## 2023 10 06 12:31:24.173 DEBUG qcprludev10 mbatchStandardLegend - import(mbatch.legend.legend)
## 2023 10 06 12:31:24.173 DEBUG qcprludev10 mbatchStandardLegend - after import
## 2023 10 06 12:31:24.174 DEBUG qcprludev10 mbatchStandardLegend - after color list
```

```
## 2023 10 06 12:31:24.174 DEBUG qcprludev10 mbatchStandardLegend - after symbol list
## 2023 10 06 12:31:24.174 DEBUG qcprludev10 mbatchStandardLegend - colorList =  #b30000
## 2023 10 06 12:31:24.175 DEBUG qcprludev10 mbatchStandardLegend - symbolList =
## 2023 10 06 12:31:24.175 DEBUG qcprludev10 mbatchStandardLegend - legendNameList =  00304 (80)
## 2023 10 06 12:31:24.176 DEBUG qcprludev10 mbatchStandardLegend - myTitle =  Rays: BatchId (80) MBatch
## 2023 10 06 12:31:24.176 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath =  /BEA/BatchEffects
## 2023 10 06 12:31:24.219 DEBUG qcprludev10 mbatchStandardLegend after Python
## 2023 10 06 12:31:24.220 DEBUG qcprludev10 mbatchStandardLegend - theTitle  Points: TSS (80)
## 2023 10 06 12:31:24.220 DEBUG qcprludev10 mbatchStandardLegend - theVersion  MBatch 2.0.3
## 2023 10 06 12:31:24.221 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath  /BEA/BatchEffectsPa
## 2023 10 06 12:31:24.222 DEBUG qcprludev10 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)
## 2023 10 06 12:31:24.222 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames length  5
## 2023 10 06 12:31:24.222 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors  #B30000, #8FB300, #
## 2023 10 06 12:31:24.223 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors length  5
## 2023 10 06 12:31:24.223 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols  0, 1, 2, 3, 4
## 2023 10 06 12:31:24.223 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols length  5
## 2023 10 06 12:31:24.224 DEBUG qcprludev10 mbatchStandardLegend - myColors  #b30000,#8fb300,#00b347,#0
## 2023 10 06 12:31:24.224 DEBUG qcprludev10 mbatchStandardLegend - theTitle UTF-8 =  Points: TSS (80)
## 2023 10 06 12:31:24.225 DEBUG qcprludev10 mbatchStandardLegend - theVersion UTF-8 =  MBatch 2.0.3
## 2023 10 06 12:31:24.225 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath UTF-8 =  /BEA/Batchl
## 2023 10 06 12:31:24.225 DEBUG qcprludev10 mbatchStandardLegend before Python
## 2023 10 06 12:31:24.226 DEBUG qcprludev10 mbatchStandardLegend - getGlobalMBatchEnv() =  /BEA/gendev
## 2023 10 06 12:31:24.226 DEBUG qcprludev10 mbatchStandardLegend - import(mbatch.legend.legend)
## 2023 10 06 12:31:24.226 DEBUG qcprludev10 mbatchStandardLegend - after import
## 2023 10 06 12:31:24.227 DEBUG qcprludev10 mbatchStandardLegend - after color list
## 2023 10 06 12:31:24.227 DEBUG qcprludev10 mbatchStandardLegend - after symbol list
## 2023 10 06 12:31:24.228 DEBUG qcprludev10 mbatchStandardLegend - colorList =  #b30000, mbatchStandard
## 2023 10 06 12:31:24.228 DEBUG qcprludev10 mbatchStandardLegend - symbolList =  0, mbatchStandardLegen
## 2023 10 06 12:31:24.228 DEBUG qcprludev10 mbatchStandardLegend - legendNameList =  OR - University of
## (72), mbatchStandardLegend - legendNameList =  OU - Roswell Park (1), mbatchStandardLegend - legendNa
## Research Institute (2), mbatchStandardLegend - legendNameList =  PA - University of Minnesota
## (1), mbatchStandardLegend - legendNameList =  PK - University Health
## Network (4)
## 2023 10 06 12:31:24.229 DEBUG qcprludev10 mbatchStandardLegend - myTitle =  Points: TSS (80) MBatch 2
## 2023 10 06 12:31:24.229 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath =  /BEA/BatchEffects
## 2023 10 06 12:31:24.347 DEBUG qcprludev10 mbatchStandardLegend after Python
## 2023 10 06 12:31:24.348 DEBUG qcprludev10 mbatchStandardLegend - theTitle  Dispersion Metrics
## 2023 10 06 12:31:24.349 DEBUG qcprludev10 mbatchStandardLegend - theVersion  MBatch 2.0.3
## 2023 10 06 12:31:24.349 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath  /BEA/BatchEffectsPa
## 2023 10 06 12:31:24.350 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames  First PCA Component
## 2023 10 06 12:31:24.350 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames length  14
## 2023 10 06 12:31:24.351 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors
## 2023 10 06 12:31:24.351 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors length  0
## 2023 10 06 12:31:24.351 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols
## 2023 10 06 12:31:24.352 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols length  0
## 2023 10 06 12:31:24.352 DEBUG qcprludev10 mbatchStandardLegend - myColors
## 2023 10 06 12:31:24.353 DEBUG qcprludev10 mbatchStandardLegend - theTitle UTF-8 =  Dispersion Metrics
## 2023 10 06 12:31:24.353 DEBUG qcprludev10 mbatchStandardLegend - theVersion UTF-8 =  MBatch 2.0.3
## 2023 10 06 12:31:24.353 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath UTF-8 =  /BEA/Batchl
## 2023 10 06 12:31:24.354 DEBUG qcprludev10 mbatchStandardLegend before Python
```

```
## 2023 10 06 12:31:24.354 DEBUG qcprludev10 mbatchStandardLegend - getGlobalMBatchEnv() =  /BEA/gendev
## 2023 10 06 12:31:24.355 DEBUG qcprludev10 mbatchStandardLegend - import(mbatch.legend.legend)
## 2023 10 06 12:31:24.355 DEBUG qcprludev10 mbatchStandardLegend - after import
## 2023 10 06 12:31:24.356 DEBUG qcprludev10 mbatchStandardLegend - after color list
## 2023 10 06 12:31:24.356 DEBUG qcprludev10 mbatchStandardLegend - after symbol list
## 2023 10 06 12:31:24.356 DEBUG qcprludev10 mbatchStandardLegend - colorList =
## 2023 10 06 12:31:24.357 DEBUG qcprludev10 mbatchStandardLegend - symbolList =
## 2023 10 06 12:31:24.357 DEBUG qcprludev10 mbatchStandardLegend - legendNameList =  First PCA Componer
## 2023 10 06 12:31:24.358 DEBUG qcprludev10 mbatchStandardLegend - myTitle =  Dispersion Metrics MBatch
## 2023 10 06 12:31:24.358 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath =  /BEA/BatchEffects
## 2023 10 06 12:31:24.383 DEBUG qcprludev10 mbatchStandardLegend after Python
## 2023 10 06 12:31:24.383 DEBUG qcprludev10 mbatchStandardCombineLegends - theTitle  Example Title for
## 2023 10 06 12:31:24.384 DEBUG qcprludev10 mbatchStandardCombineLegends - theFilenamePath  /BEA/Batchl
## 2023 10 06 12:31:24.384 DEBUG qcprludev10 mbatchStandardCombineLegends - theListOfFiles  /BEA/BatchEf
## 2023 10 06 12:31:24.385 DEBUG qcprludev10 mbatchStandardCombineLegends - theTitle UTF-8 =  Example Ti
## 2023 10 06 12:31:24.385 DEBUG qcprludev10 mbatchStandardCombineLegends - theFilenamePath UTF-8 =  /BE
## 2023 10 06 12:31:24.385 DEBUG qcprludev10 mbatchStandardCombineLegends before Python
## 2023 10 06 12:31:24.386 DEBUG qcprludev10 mbatchStandardCombineLegends - import(mbatch.legend.legend)
## 2023 10 06 12:31:24.449 DEBUG qcprludev10 mbatchStandardCombineLegends after Python
## 2023 10 06 12:31:24.450 DEBUG qcprludev10 doInternalPca before filter - Number of rownames for pca sc
## 2023 10 06 12:31:24.451 DEBUG qcprludev10 doInternalPca before filter - length of batch ids for sampl
## 2023 10 06 12:31:24.451 DEBUG qcprludev10 doInternalPca after filter - Number of rownames for pca sco
## 2023 10 06 12:31:24.452 DEBUG qcprludev10 doInternalPca after filter - length of batch ids for sample
## 2023 10 06 12:31:24.452 DEBUG qcprludev10 doInternalPca - pvalueDSC
## 2023 10 06 12:31:24.453 DEBUG qcprludev10 pvalueDSC start
## 2023 10 06 12:31:24.453 DEBUG qcprludev10 pvalueDSC transpose
## 2023 10 06 12:31:24.454 DEBUG qcprludev10 pvalueDSC call pvalueDSCwithExcerpt
## 2023 10 06 12:31:24.454 DEBUG qcprludev10 pvalueDSCwithExcerpt start
## 2023 10 06 12:31:24.455 DEBUG qcprludev10 nrow(thePcaDataExcerpt)= 2
## 2023 10 06 12:31:24.455 DEBUG qcprludev10 ncol(thePcaDataExcerpt)= 80
## 2023 10 06 12:31:24.455 DEBUG qcprludev10 length(theBatchIdsForSamples)= 80
## 2023 10 06 12:31:24.456 DEBUG qcprludev10 getDSCwithExcerpt before Python
## 2023 10 06 12:31:24.456 DEBUG qcprludev10 getDSCwithExcerpt - import(mbatch.dsc.dsc_calc)
## 2023 10 06 12:31:24.458 DEBUG qcprludev10 getDSCwithExcerpt after Python
## 2023 10 06 12:31:24.459 DEBUG qcprludev10 pvalueDSCwithExcerpt after getDSCwithExcerpt
## 2023 10 06 12:31:24.459 DEBUG qcprludev10 pvalueDSCwithExcerpt length(unique(theBatchIdsForSamples))=
## 2023 10 06 12:31:24.460 DEBUG qcprludev10 pvalueDSCwithExcerpt 2222 length(permResults)= 1
## 2023 10 06 12:31:24.460 DEBUG qcprludev10 pvalueDSCwithExcerpt PCA-PVALUE-DSC
## 2023 10 06 12:31:24.461 DEBUG qcprludev10 openAndWriteDscPermsFile -- start
## 2023 10 06 12:31:24.461 DEBUG qcprludev10 openAndWriteDscPermsFile -- redo name
## 2023 10 06 12:31:24.462 DEBUG qcprludev10 openAndWriteDscPermsFile -- check file
## 2023 10 06 12:31:24.462 DEBUG qcprludev10 openAndWriteDscPermsFile -- no file, write
## 2023 10 06 12:31:24.463 DEBUG qcprludev10 openAndWriteDscPermsFile -- nothing to write
## 2023 10 06 12:31:24.464 DEBUG qcprludev10 write PCA file direct 1  /BEA/BatchEffectsPackage_data/test
## 2023 10 06 12:31:24.464 DEBUG qcprludev10 writePCAAnnotations
## 2023 10 06 12:31:24.465 DEBUG qcprludev10 writePCAAnnotations - data  PC1,PC3
## 2023 10 06 12:31:24.467 DEBUG qcprludev10 writePcaAnalysis_Image - Number of rownames for pca score
## 2023 10 06 12:31:24.467 DEBUG qcprludev10 writePcaAnalysis_Image - length of batch ids for samples  8

## 2023 10 06 12:31:24.540 DEBUG qcprludev10 mbatchStandardLegend - theTitle  Rays: BatchId (80)
## 2023 10 06 12:31:24.541 DEBUG qcprludev10 mbatchStandardLegend - theVersion  MBatch 2.0.3
## 2023 10 06 12:31:24.542 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath  /BEA/BatchEffectsPa
## 2023 10 06 12:31:24.542 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames  00304 (80)
## 2023 10 06 12:31:24.543 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames length  1
```

```
## 2023 10 06 12:31:24.543 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors  #B30000
## 2023 10 06 12:31:24.543 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors length  1
## 2023 10 06 12:31:24.544 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols
## 2023 10 06 12:31:24.544 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols length  0
## 2023 10 06 12:31:24.544 DEBUG qcprludev10 mbatchStandardLegend - myColors  #b30000
## 2023 10 06 12:31:24.545 DEBUG qcprludev10 mbatchStandardLegend - theTitle UTF-8 =  Rays: BatchId (80)
## 2023 10 06 12:31:24.545 DEBUG qcprludev10 mbatchStandardLegend - theVersion UTF-8 =  MBatch 2.0.3
## 2023 10 06 12:31:24.546 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath UTF-8 =  /BEA/Batch
## 2023 10 06 12:31:24.546 DEBUG qcprludev10 mbatchStandardLegend before Python
## 2023 10 06 12:31:24.546 DEBUG qcprludev10 mbatchStandardLegend - getGlobalMBatchEnv() =  /BEA/gendev
## 2023 10 06 12:31:24.547 DEBUG qcprludev10 mbatchStandardLegend - import(mbatch.legend.legend)
## 2023 10 06 12:31:24.547 DEBUG qcprludev10 mbatchStandardLegend - after import
## 2023 10 06 12:31:24.548 DEBUG qcprludev10 mbatchStandardLegend - after color list
## 2023 10 06 12:31:24.548 DEBUG qcprludev10 mbatchStandardLegend - after symbol list
## 2023 10 06 12:31:24.548 DEBUG qcprludev10 mbatchStandardLegend - colorList =  #b30000
## 2023 10 06 12:31:24.549 DEBUG qcprludev10 mbatchStandardLegend - symbolList =
## 2023 10 06 12:31:24.549 DEBUG qcprludev10 mbatchStandardLegend - legendNameList =  00304 (80)
## 2023 10 06 12:31:24.549 DEBUG qcprludev10 mbatchStandardLegend - myTitle =  Rays: BatchId (80) MBatch
## 2023 10 06 12:31:24.550 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath =  /BEA/BatchEffects
## 2023 10 06 12:31:24.590 DEBUG qcprludev10 mbatchStandardLegend after Python
## 2023 10 06 12:31:24.591 DEBUG qcprludev10 mbatchStandardLegend - theTitle  Points: TSS (80)
## 2023 10 06 12:31:24.591 DEBUG qcprludev10 mbatchStandardLegend - theVersion  MBatch 2.0.3
## 2023 10 06 12:31:24.592 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath  /BEA/BatchEffectsPa
## 2023 10 06 12:31:24.592 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames  OR - University of
## (72), OU - Roswell Park (1), P6 - Translational Genomics
## Research Institute (2), PA - University of Minnesota
## (1), PK - University Health
## Network (4)
## 2023 10 06 12:31:24.593 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames length  5
## 2023 10 06 12:31:24.593 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors  #B30000, #8FB300, #
## 2023 10 06 12:31:24.594 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors length  5
## 2023 10 06 12:31:24.594 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols  0, 1, 2, 3, 4
## 2023 10 06 12:31:24.594 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols length  5
## 2023 10 06 12:31:24.595 DEBUG qcprludev10 mbatchStandardLegend - myColors  #b30000,#8fb300,#00b347,#0
## 2023 10 06 12:31:24.595 DEBUG qcprludev10 mbatchStandardLegend - theTitle UTF-8 =  Points: TSS (80)
## 2023 10 06 12:31:24.596 DEBUG qcprludev10 mbatchStandardLegend - theVersion UTF-8 =  MBatch 2.0.3
## 2023 10 06 12:31:24.596 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath UTF-8 =  /BEA/Batch
## 2023 10 06 12:31:24.596 DEBUG qcprludev10 mbatchStandardLegend before Python
## 2023 10 06 12:31:24.597 DEBUG qcprludev10 mbatchStandardLegend - getGlobalMBatchEnv() =  /BEA/gendev
## 2023 10 06 12:31:24.597 DEBUG qcprludev10 mbatchStandardLegend - import(mbatch.legend.legend)
## 2023 10 06 12:31:24.597 DEBUG qcprludev10 mbatchStandardLegend - after import
## 2023 10 06 12:31:24.598 DEBUG qcprludev10 mbatchStandardLegend - after color list
## 2023 10 06 12:31:24.598 DEBUG qcprludev10 mbatchStandardLegend - after symbol list
## 2023 10 06 12:31:24.599 DEBUG qcprludev10 mbatchStandardLegend - colorList =  #b30000, mbatchStandard
## 2023 10 06 12:31:24.599 DEBUG qcprludev10 mbatchStandardLegend - symbolList =  0, mbatchStandardLeger
## 2023 10 06 12:31:24.599 DEBUG qcprludev10 mbatchStandardLegend - legendNameList =  OR - University o
## (72), mbatchStandardLegend - legendNameList =  OU - Roswell Park (1), mbatchStandardLegend - legendNa
## Research Institute (2), mbatchStandardLegend - legendNameList =  PA - University of Minnesota
## (1), mbatchStandardLegend - legendNameList =  PK - University Health
## Network (4)
## 2023 10 06 12:31:24.600 DEBUG qcprludev10 mbatchStandardLegend - myTitle =  Points: TSS (80) MBatch 2
## 2023 10 06 12:31:24.600 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath =  /BEA/BatchEffects
## 2023 10 06 12:31:24.718 DEBUG qcprludev10 mbatchStandardLegend after Python
## 2023 10 06 12:31:24.719 DEBUG qcprludev10 mbatchStandardLegend - theTitle  Dispersion Metrics
```

```
## 2023 10 06 12:31:24.719 DEBUG qcprludev10 mbatchStandardLegend - theVersion  MBatch 2.0.3
## 2023 10 06 12:31:24.720 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath  /BEA/BatchEffectsPa
## 2023 10 06 12:31:24.721 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames  First PCA Component
## 2023 10 06 12:31:24.721 DEBUG qcprludev10 mbatchStandardLegend - theLegendNames length  14
## 2023 10 06 12:31:24.721 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors
## 2023 10 06 12:31:24.722 DEBUG qcprludev10 mbatchStandardLegend - theLegendColors length  0
## 2023 10 06 12:31:24.722 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols
## 2023 10 06 12:31:24.723 DEBUG qcprludev10 mbatchStandardLegend - theLegendSymbols length  0
## 2023 10 06 12:31:24.723 DEBUG qcprludev10 mbatchStandardLegend - myColors
## 2023 10 06 12:31:24.723 DEBUG qcprludev10 mbatchStandardLegend - theTitle UTF-8 =  Dispersion Metrics
## 2023 10 06 12:31:24.724 DEBUG qcprludev10 mbatchStandardLegend - theVersion UTF-8 =  MBatch 2.0.3
## 2023 10 06 12:31:24.724 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath UTF-8 =  /BEA/Batchl
## 2023 10 06 12:31:24.724 DEBUG qcprludev10 mbatchStandardLegend before Python
## 2023 10 06 12:31:24.725 DEBUG qcprludev10 mbatchStandardLegend - getGlobalMBatchEnv() =  /BEA/gendev
## 2023 10 06 12:31:24.725 DEBUG qcprludev10 mbatchStandardLegend - import(mbatch.legend.legend)
## 2023 10 06 12:31:24.726 DEBUG qcprludev10 mbatchStandardLegend - after import
## 2023 10 06 12:31:24.726 DEBUG qcprludev10 mbatchStandardLegend - after color list
## 2023 10 06 12:31:24.726 DEBUG qcprludev10 mbatchStandardLegend - after symbol list
## 2023 10 06 12:31:24.727 DEBUG qcprludev10 mbatchStandardLegend - colorList =
## 2023 10 06 12:31:24.727 DEBUG qcprludev10 mbatchStandardLegend - symbolList =
## 2023 10 06 12:31:24.728 DEBUG qcprludev10 mbatchStandardLegend - legendNameList =  First PCA Componer
## 2023 10 06 12:31:24.728 DEBUG qcprludev10 mbatchStandardLegend - myTitle =  Dispersion Metrics MBatch
## 2023 10 06 12:31:24.728 DEBUG qcprludev10 mbatchStandardLegend - theFilenamePath =  /BEA/BatchEffects
## 2023 10 06 12:31:24.753 DEBUG qcprludev10 mbatchStandardLegend after Python
## 2023 10 06 12:31:24.754 DEBUG qcprludev10 mbatchStandardCombineLegends - theTitle  Example Title for
## 2023 10 06 12:31:24.754 DEBUG qcprludev10 mbatchStandardCombineLegends - theFilenamePath  /BEA/Batchl
## 2023 10 06 12:31:24.755 DEBUG qcprludev10 mbatchStandardCombineLegends - theListOfFiles  /BEA/BatchEf
## 2023 10 06 12:31:24.755 DEBUG qcprludev10 mbatchStandardCombineLegends - theTitle UTF-8 =  Example Ti
## 2023 10 06 12:31:24.755 DEBUG qcprludev10 mbatchStandardCombineLegends - theFilenamePath UTF-8 =  /BE
## 2023 10 06 12:31:24.756 DEBUG qcprludev10 mbatchStandardCombineLegends before Python
## 2023 10 06 12:31:24.756 DEBUG qcprludev10 mbatchStandardCombineLegends - import(mbatch.legend.legend)
## 2023 10 06 12:31:24.861 DEBUG qcprludev10 mbatchStandardCombineLegends after Python
##  [1] "ALL__CompListDSC.RData"
##  [2] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_Comp1_Comp2_DSC.txt"
##  [3] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_Comp1_Comp2_DSC.txt__Com
##  [4] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_Comp1_Comp3_DSC.txt"
##  [5] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_Comp1_Comp3_DSC.txt__Com
##  [6] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_DSC.txt"
##  [7] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_DSC.txt__OverallDSC.RDat
##  [8] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_FVE.txt"
##  [9] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_Scores.txt"
## [10] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_Weights.txt"
## [11] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/BatchIdwithTSS_Comp
## [12] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/BatchIdwithTSS_Comp
## [13] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/BatchIdwithTSS_Comp
## [14] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/BatchIdwithTSS_Comp
## [15] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/BatchIdwithTSS_Comp
## [16] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/BatchIdwithTSS_Comp
## [17] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/BatchIdwithTSS_Comp
## [18] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/BatchIdwithTSS_Comp
## [19] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/BatchIdwithTSS_Comp
## [20] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/BatchIdwithTSS_Comp
## [21] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCAAnnotations.tsv"
## [22] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCAValues.tsv"
```

```
## [23] "BatchIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA_Title.txt"
## [24] "DSC/DATA_2022-09-09-1600/TEST_2022-10-10-1300/DSCOverview.tsv"
## [25] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_Comp1_Comp2_DSC.txt"
## [26] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_Comp1_Comp2_DSC.txt__Co
## [27] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_Comp1_Comp3_DSC.txt"
## [28] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_Comp1_Comp3_DSC.txt__Co
## [29] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_DSC.txt"
## [30] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_DSC.txt__OverallDSC.RDa
## [31] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_FVE.txt"
## [32] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_Scores.txt"
## [33] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/ANY_Weights.txt"
## [34] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/PlateIdwithTSS_Comp
## [35] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/PlateIdwithTSS_Comp
## [36] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/PlateIdwithTSS_Comp
## [37] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/PlateIdwithTSS_Comp
## [38] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/PlateIdwithTSS_Comp
## [39] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/PlateIdwithTSS_Comp
## [40] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/PlateIdwithTSS_Comp
## [41] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/PlateIdwithTSS_Comp
## [42] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/PlateIdwithTSS_Comp
## [43] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA-Plus/PlateIdwithTSS_Comp
## [44] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCAAnnotations.tsv"
## [45] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCAValues.tsv"
## [46] "PlateIdwithTSS/DualBatch/DATA_2022-09-09-1600/TEST_2022-10-10-1300/PCA_Title.txt"
```

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.

```
/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.

It then has a Data Version and a Test Version Directory

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.

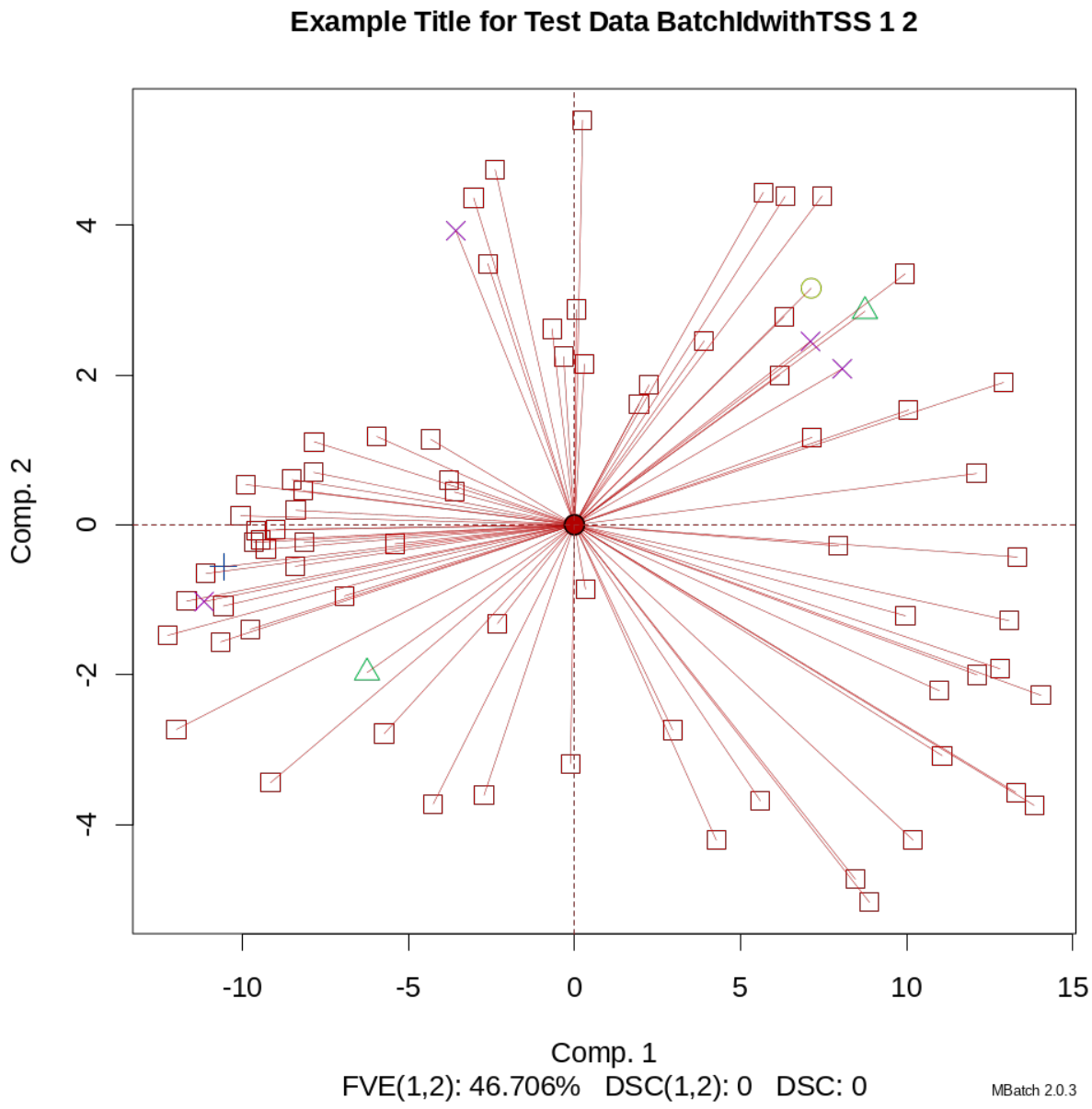

Figure 2: PCAPlus Output