

RANKRANKBATCH USER'S GUIDE

OVERVIEW

This section provides an overview of the “rankrankbatch” web page (<http://systems.crump.ucla.edu/~llam/rankrank/rankrankbatch.php>) and user guide.

Rankrankbatch Description

The rankrankbatch web page receives and processes multiple rank rank hypergeometric (RRHO) comparisons to generate a batch of RRHO results for each comparison.

RRHO processing is placed in the server's processing queue once a zipped batch folder containing all data sets and annotations are successfully uploaded from the rankrankbatch web page. An email containing a link to the zipped results are sent to the user once RRHO processing is complete for a submitted zip file.

ZIPPED INPUT FILES

This section provides a description of the types of files that may be included in the zip file.

The following files may be required in the zipped file to successfully upload to the RRHO server.

1. RNKLIST files (*.txt) - A tab delimited text file that is only required when performing the simple version of RRHO. The rnklist file contains the following essential columns:
(Sample file: http://systems.crump.ucla.edu/rankrank/demo/rankrank.94aaa8b7716d7f.ranks_true.txt)
 - a. UniGene ID
 - b. Gene Name
 - c. Rank 1
 - d. Rank 2
 - e. Metric 1
 - f. Metric 2
2. GCT files (*.gct) - Gene cluster text file format is a tab delimited text file that includes expression values. This file is required when performing the regular version of RRHO.
(Sample file: <http://systems.crump.ucla.edu/rankrank/demo/krasla.gct>)
 - a. [http://www.broadinstitute.org/cancer/software/gsea/wiki/index.php/Data_formats#GCT: Gene Cluster Text file format .28.2A.gct.29](http://www.broadinstitute.org/cancer/software/gsea/wiki/index.php/Data_formats#GCT:GeneClusterTextfileformat.28.2A.gct.29)
3. CLS file (*.cls) – Categorical class file format to define the phenotype or class of each sample in the GCT file. This file is required when performing the regular version of RRHO.
(Sample file: <http://systems.crump.ucla.edu/rankrank/demo/krasla.cls>)
 - a. [http://www.broadinstitute.org/cancer/software/gsea/wiki/index.php/Data_formats#CLS: Categorical .28e.g tumor vs normal.29 class file format .28.2A.cls.29](http://www.broadinstitute.org/cancer/software/gsea/wiki/index.php/Data_formats#CLS:Categorical.28e.g.tumorvsnormal.29classfileformat.28.2A.cls.29)
4. Annotation file - Tab delimited text file with each line having the columns listed below. This file is required when performing the regular version of RRHO.

(Sample file: <http://systems.crump.ucla.edu/~llam/rankrank/download/MG-U74Av2.Mm.2007.04.24.acc2ug.combined>)

- a. ID as used in the GCT file (e.g., Affymetrix probeset ID)
 - b. Accession number (e.g., RefSeq)
 - c. UniGene ID
 - d. Gene/protein symbol
 - e. Gene/protein description
5. Manifest.txt – plain text file listing the data set files and dataset information for each RRHO comparison. This file is required for both simple and regular versions of RRHO.
- a. Each RRHO comparison comprise of two data sets described by the following lines
RNKLIST: <FILE NAME OF RNKLIST FILE IN THE ZIP FOLDER | FOR REGULAR RRHO LEAVE WHITESPACE AFTER THE COLON >
GCT1: <FILE NAME OF GCT FILE IN ZIP FOLDER IN THE 1ST DATASET | FOR SIMPLE RRHO LEAVE WHITESPACE AFTER THE COLON>
CLS1: <FILE NAME OF CLS FILE FOR GCT1 IN THE 1ST DATASET | FOR SIMPLE RRHO LEAVE WHITESPACE AFTER THE COLON>
Class1: <1ST CLASS NAME DEFINED IN CLS1 IN THE 1ST DATASET >
Class2: <2ND CLASS NAME DEFINED IN CLS1 IN THE 1ST DATASET >
Annotation1: <FILE NAME OF ANNOTATION FILE FOR GCT1 IN THE 1ST DATASET | FOR SIMPLE RRHO LEAVE WHITESPACE AFTER THE COLON >
Species1: <UNIGENE SPECIES IDENTIFIER e.g., Mm or Hs IN THE 1ST DATASET | FOR SIMPLE RRHO LEAVE WHITESPACE AFTER THE COLON >
Description1: <USER DEFINED DESCRIPTION IN THE 1ST DATASET >
GCT2: <FILE NAME OF GCT FILE IN ZIP FOLDER IN THE 2ND DATASET | FOR SIMPLE RRHO LEAVE WHITESPACE AFTER THE COLON>
CLS2: < FILE NAME OF CLS FILE FOR GCT2 IN THE 2ND DATASET | FOR SIMPLE RRHO LEAVE WHITESPACE AFTER THE COLON>
Class1: <1ST CLASS NAME DEFINED IN CLS2 IN THE 2ND DATASET>
Class2: < 2ND CLASS NAME DEFINED IN CLS2 IN THE 2ND DATASET >
Annotation2: <FILENAME OF ANNOTATION FILE FOR GCT2 IN THE 2ND DATASET | FOR SIMPLE RRHO LEAVE WHITESPACE AFTER THE COLON>
Species2: < UNIGENE SPECIES IDENTIFIER e.g., Mm or Hs IN THE 2ND DATASET | FOR SIMPLE RRHO LEAVE WHITESPACE AFTER THE COLON >
Description2: <USER DEFINED DESCRIPTION IN THE 2ST DATASET >
 - b. Additional RRHO comparisons are preceded by three pound symbols “###”
 - c. A sample file with two regular RRHO comparisons can be found in the following location.
<http://systems.crump.ucla.edu/~llam/rankrank/download/regular/manifest.txt>
 - d. A sample file with two simple RRHO comparisons can be found in the following location.
<http://systems.crump.ucla.edu/~llam/rankrank/download/simple/manifest.txt>

RANKRANKBATCH WEB PAGE PARAMETERS

This section describes each field and options within the rankrankbatch web page.

- Zip File Field – Select the browse button and choose the full path the zip file containing the RRHO data sets.
- Email Address - The user email address used to receive the results notification.
- Step Size – The number of permutations.
- Rank Lists – Check this box to receive the rank list file(s) in the result zip file.
- Bin file – check this box to receive the bin file(s) in the result zip file.
- Heat map – check this box to receive the heat map(s) in the result zip file.
- Rank scatter plot – check this box to receive the rank scatter plot(s) in the result zip file.
- Metric scatter plot – check this box to receive the metric scatter plot(s) in the result zip file.
- Statistics Summary – check this box to receive the statistics report(s) in the result zip file.
- Perform Bejamen-Yekutieli p-value correction – check this box to receive the corrected heat map(s).
- Overlapping gene list – check this box to receive the list of overlapping genes.

SAMPLE REGULAR BATCH RRHO SUBMISSION

This section outlines the steps needed to complete two regular RRHO comparisons in a single submission.

Zip File Parameters

	Dataset 1	Dataset 2
1st RRHO	RNKLIST: GCT1: krasla.gct CLS1: krasla.cls Class1: normal Class2: tumor Annotation1: MG- U74Av2.Mm.2007.04.24.acc2ug.combined Species1: Mm Description1: KrasLA	GCT2:bostonlungsubclasses.gct CLS2:bostonlungsub.cls Class1: nl Class2: ad Annotation2: HG- U95Av2.Hs.2007.04.24.acc2ug.combined Species2: Hs Description2: Boston
2nd RRHO	RNKLIST: GCT1:bostonlungsubclasses.gct CLS1:bostonlungsub.cls Class1: nl Class2: ad Annotation1: HG- U95Av2.Hs.2007.04.24.acc2ug.combined Species1: Hs Description1: Boston2	GCT2: krasla.gct CLS2: krasla.cls Class1: normal Class2: tumor Annotation2: MG- U74Av2.Mm.2007.04.24.acc2ug.combined Species2: Mm Description2: KrasLA2

Selected Batch Options

Rank Lists	Bin File	Heat Map	Rank scatter plot	Metric scatter plot	Statistics Summary	Perform Benjamini-Yekutieli p-value correction	Overlapping gene list
X	x	x	x	x	x	x	x

Step 1

Upload batch zip file.

A sample file is available at

http://systems.crump.ucla.edu/~llam/rankrank/download/regular/sample_regular_rrhobatch.zip

Step 2

Enter user email address

Step 3

Enter step size (e.g. 100)

Step 4

Check desired options

Step 5

Select OK

Within 1-3 minutes, a verification email will be sent upon successful submission of the batch file.

Step 6

Download results from emailed download link.

The email will be delivered within 10-15 min.

Batch results will remain on the server for 48 hrs.

SAMPLE SIMPLE BATCH RRHO SUBMISSION

This section outlines the steps needed to complete two simple RRHO comparisons in a single submission.

Zip File Parameters

	Dataset 1	Dataset 2
1st RRHO	RNKLIST: rankrank.00.ranks_true.txt GCT1: CLS1: Class1: normal Class2: tumor Annotation1: Species1: Description1: KrasLA	GCT2: CLS2: Class1: nl Class2: ad Annotation2: Species2: Description2: Boston
2nd RRHO	RNKLIST: rankrank.01.ranks_true.txt GCT1: CLS1: Class1: nl Class2: ad Annotation1: Species1: Description2: Boston2	GCT2: CLS2: Class1: normal Class2: tumor Annotation2: Species2: Description2: KrasLA2

Selected Batch Options

Rank Lists	Bin File	Heat Map	Rank scatter plot	Metric scatter plot	Statistics Summary	Perform Benjamini-Yekutieli p-value correction	Overlapping gene list
X	X	X	X	X	X	X	X

Step 1

Upload batch zip file.

A sample file is available at

http://systems.crump.ucla.edu/~llam/rankrank/download/simple/sample_simple_rrhobatch.zip

Step 2

Enter user email address

Step 3

Enter step size (e.g. 100)

Step 4

Check desired options

Step 5

Select OK

Within 1-3 minutes, a verification email will be sent upon successful submission of the batch file.

Step 6

Download results from emailed download link.

The email will be delivered within 10-15 min.

Batch results will remain on the server for 48 hrs.

LINKS

Regular RRHO webpage: <http://systems.crump.ucla.edu/rankrank/>

Simple RRHO webpage: <http://systems.crump.ucla.edu/rankrank/rankranksimple.php>

RRHO user guide: http://systems.crump.ucla.edu/rankrank/PlaisierSupplementaryData-SupplementaryMethods_UsersGuide.pdf

GSEA Algorithm: <http://www.broadinstitute.org/gsea/>

GCT file description: <http://systems.crump.ucla.edu/rankrank/demo/krasla.gct>

[http://www.broadinstitute.org/cancer/software/gsea/wiki/index.php/Data_formats#GCT: Gene Cluster Text file format .28.2A.gct.29](http://www.broadinstitute.org/cancer/software/gsea/wiki/index.php/Data_formats#GCT:_Gene_Cluster_Text_file_format_.28.2A.gct.29)

CLS file description:

[http://www.broadinstitute.org/cancer/software/gsea/wiki/index.php/Data_formats#CLS: Categorical .28e.g.tumor vs normal.29 class file format .28.2A.cls.29](http://www.broadinstitute.org/cancer/software/gsea/wiki/index.php/Data_formats#CLS:_Categorical_.28e.g.tumor_vs_normal.29_class_file_format_.28.2A.cls.29)

Sample regular batch RRHO zip file:

http://systems.crump.ucla.edu/~llam/rankrank/download/regular/sample_regular_rrhobatch.zip

Sample simple batch RRHO zip file:

http://systems.crump.ucla.edu/~llam/rankrank/download/simple/sample_simple_rrhobatch.zip