

# **User manual of JUMPp Batch**



**Center for Proteomics and Metabolomics**

**The Peng Laboratory**

**St. Jude Children's Research Hospital**

**June 2020**

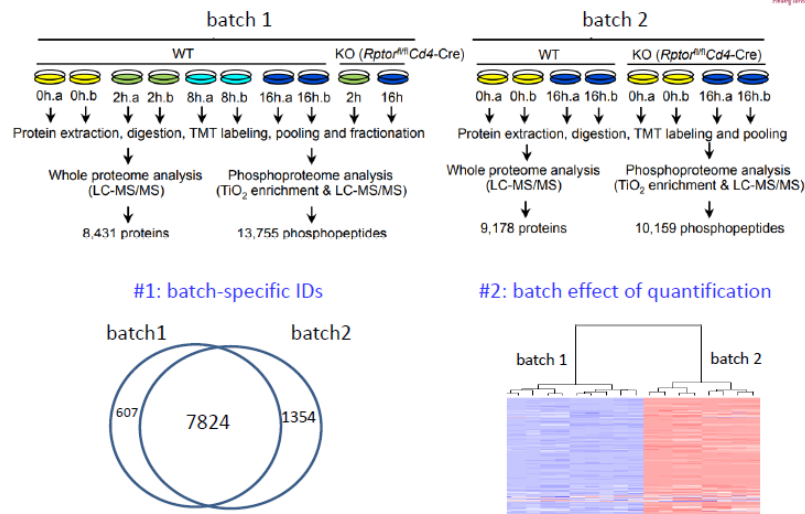
## **Contents**

<b>1. Introduction .....</b>	<b>3</b>
<b>2. How to setup and run JUMPP Batch .....</b>	<b>4</b>
<b>2.1 Run JUMPP Batch .....</b>	<b>4</b>
<b>2.1.1 JUMPP Batch setup.....</b>	<b>4</b>
<b>2.1.2 Edit batch params .....</b>	<b>5</b>
<b>2.2 Troubleshoot.....</b>	<b>8</b>

## 1. Introduction

There are two kinds of challenges in multi-batch TMT analysis as below.

### Multi-batch TMT analysis: two challenges

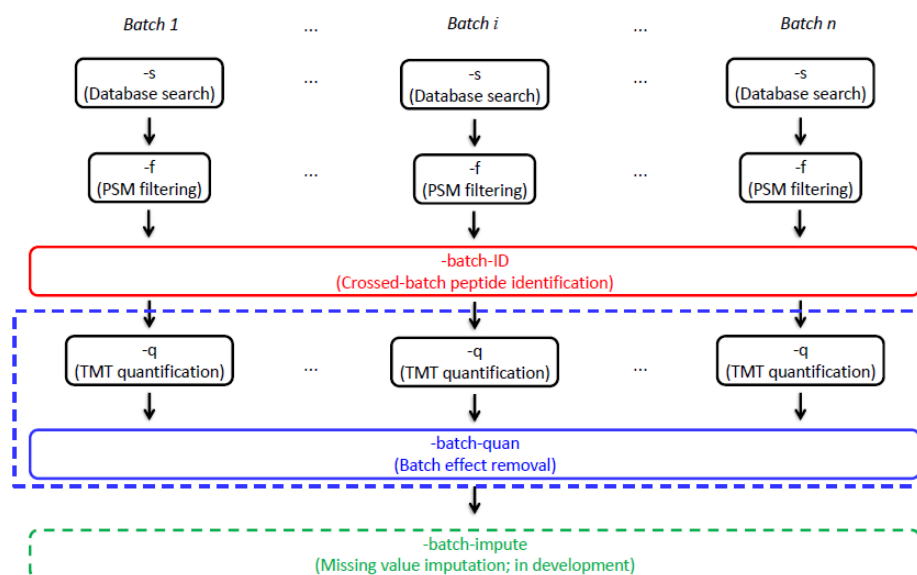


For challenge #1, we developed JUMPp -batch-ID by using peptides identified from other batches to rescue peptides in the current batch.

For challenge #2, we developed JUMPp -batch-quant by using internal standard or linear model fitting to normalize crossed-batch signals.

Here is the overall workflow for the multi-batch TMT analysis.

### Quantitative Proteomics JUMP Software Suite for Multi-batch analysis



## 2. How to setup and run JUMPp Batch

### 2.1 Run JUMPp Batch

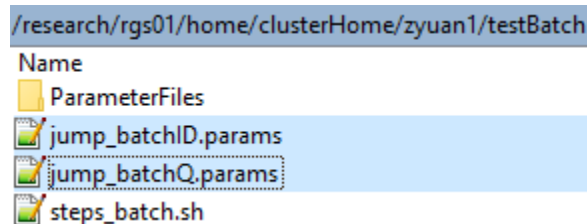
#### 2.1.1 JUMPp Batch setup

Login HPC (PuTTY and WinSCP) and go to a work path (e.g. /home/zyuan1/testBatch).

- a. In the work path, load batch params in PuTTY.  
module load jump/1.13.003  
jump -params

```
[zyuan1@splprhpc05 testBatch]$ module load jump/1.13.003  
[zyuan1@splprhpc05 testBatch]$ jump -params
```

- b. Edit batch params (i.e. jump\_batchID.params, jump\_batchQ.params) in WinSCP. Skip this step if you want to run the example data.



- c. Run batch-ID in PuTTY (by default the results are in folder 'batch\_id').  
jump -batch-id jump\_batchID.params

```
[zyuan1@splprhpc06 testBatch]$ jump -batch-id jump_batchID.params
```

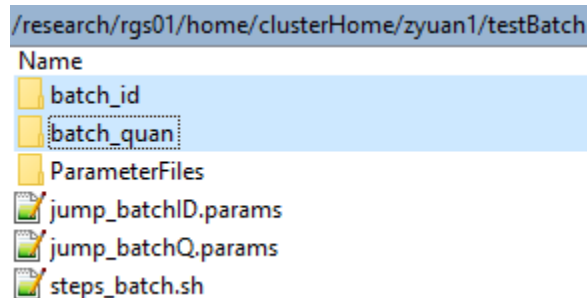
- d. Run batch-quan in PuTTY (by default the results are in folder 'batch\_quan').  
jump -batch-q jump\_batchQ.params

```
[zyuan1@splprhpc06 testBatch]$ jump -batch-q jump_batchQ.params
```

Alternatively, step c & d can be run in one command: bash steps\_batch.sh

```
[zyuan1@splprhpc06 testBatch]$ bash steps_batch.sh
```

Here show the results folders.



## 2.1.2 Edit batch params

```
jump_batchID.params
1 # Inputs: absolute path of publication tables from JUMP -f results (IDwDecoy.txt)
2 input_path_batch1 = /hpcf/authorized_apps/teomics_apps/pipeline/release/SampleData/batch_test/b1/sum_HH_tmt/IDwDecoy.txt
3 input_path_batch2 = /hpcf/authorized_apps/teomics_apps/pipeline/release/SampleData/batch_test/b2/sum_HH_tmt/IDwDecoy.txt
4
5 # thresholds:
6 min_Jscore = 10 # minimum Jscore cutoff to be considered (10 for JUMP; 1 for Comet)
7 multiHit_max_dN = 0.1 # for considering non-top hits for a PSM: PSMs within such dN range will be considered in a re
8
9 enable_group_specific_Jscore = 1 # enable group specific Jscore filtering. each group defined by charge states and pepti
10 score_cutoff_quantile = 0 # Jscore threshold for each group, defined as the lower quantile (default 5%) of Jscores in eac
11
12 # modifications
13 mods = 0 # Display modified peptides and their unmodified (0:Off, K:Lys, STY: Phosphorylation, ...); same as -f
14 #mods = STY # Display modified peptides and their unmodified (0:Off, K:Lys, STY: Phosphorylation, ...); same as -f
15
16 # Output:
17 output_folder = batch_id # output folder name
18
19 # other parameters:
20 jump_f_path = /research/rgs01/applications/hpcf/authorized_apps/teomics_apps/pipeline/release/version1.13.003/JUMP/bin/b
21 # (in JUMP, pit_file = 0; in Comet, set pit_file according to search)
22 # pit_file = /hpcf/authorized_apps/teomics_apps/database/20150201/human_ft_mc2_c57_TMT_R229.pit
23 pit_file = 0
24 database = 0
25
26 # HPC parameters
27 dispatch = localhost
```

jump\_batchID.params: pay attention to the below params, and change others if you need.

(1) input\_path\_batch1, ..., input\_path\_batchn: absolute path of publication tables from JUMP -f results (IDwDecoy.txt)

(2) min\_Jscore: minimum Jscore cutoff to be considered (10 for JUMP; 1 for Comet)

(3) output\_folder: output folder name (the default is batch\_id). If it is changed here, remember to change it in 'path\_batch\_id' of 'jump\_batchQ.params'.

(4) pit\_file: in JUMP, pit\_file = 0; in Comet, set pit\_file according to search

```
jump_batchQ.params
1 # JUMP batch correction parameter file
2 input_mode = 1 # 1: proteins (for whole proteome); 2: pho site; 3: peptides (from either phosphor- or whole
3
4 # output path of -batch-id results
5 path_batch_id = /research/rgs01/home/clusterHome/zyuan1/testBatch/batch_id
6
7 # ATTENTION: jump -q will run automatically for all batches using DEFAULT parameters
8 # Otherwise, User can edit filtering options within ParameterFiles/TMThh/jump_qj_HH_tmt10_human.params to customize jump -q r
9
10 # Specify TMT-plex for each batch (that match to jump -batch-id results)
11 input_n_batch1 = 10
12 input_n_batch2 = 10
13
14 # Outputs:
15 output_folder = quan # output folder suffix name; prefix always 'batch'
16
17 # Parameters:
18 normalization_method = 1 # 0: None (i.e., just combine publication tables); 1: using internal standard; 2: using 1
19 isoform_rescue = 1 # 0: turn off; 1: turn on function. Suppose for a gene, there are two isoforms (say a and b)
20
21 # Internal standards for each batch. The most comprehensive batch should be put first
22 internal_standard_batch1 = sig126 # internal standard for each batch
23 internal_standard_batch2 = sig126 # internal standard for each batch
24
25 # -i parameters
26 jump_i_path = /research/rgs01/applications/hpcf/authorized_apps/teomics_apps/pipeline/release/version1.13.003/JUMP/bin/bat
27
28 # jump -q parameters (the values here will overwrite the default values copied from the ParameterFiles/ folder above)
29 ppi_filter = 50 # precursor peak intensity percentage threshold
30 impurity_correction = 1 # 1 = Yes; 0 = No; if only a part of reporters are used, it should be set to
31 loading_bias_correction = 1 # 1 = Yes; 0 = No;
32 interference_removal = 0 # 1 = Yes; 0 = No;
```

jump\_batchQ.params: pay attention to the below params, and change others if you need.

(1) path\_batch\_id: output path of -batch-id results, the same folder name as 'output\_folder' of 'jump\_batchID.params'.

(2) jump -q will run automatically for all batches using DEFAULT parameters. Users can edit filtering options within ParameterFiles/TMThh/jump\_qj\_HH\_tmt10\_human.params to customize jump -q runs.

(3) input\_n\_batch1, ..., input\_n\_batchn: specify TMT-plex for each batch (that match to jump -batch-id results).

(4) internal\_standard\_batch1, ..., internal\_standard\_batchn: internal standards for each batch if normalization\_method = 1 (using internal standard).

(5) jump -q parameters (the values here will overwrite the default values copied from the ParameterFiles/ folder above)

Here we show 4 cases to edit parameters.

**Case1:** phos data, input\_mode = 1

jump\_batchID.params:

```
input_path_batch1                                     =  
/hpcf/authorized_apps/proteomics_apps/pipeline/release/SampleData/phos_test/b1/sum  
_HH_tmt/IDwDecoy.txt  
  
input_path_batch2                                     =  
/hpcf/authorized_apps/proteomics_apps/pipeline/release/SampleData/phos_test/b2/sum  
_HH_tmt/IDwDecoy.txt  
  
mods = 0
```

It can be edited: /ParameterFiles/TMThh/jump\_qj\_HH\_tmt10\_human.params

jump\_batchQ.params:

```
input_mode = 1          # 1: proteins (for whole proteome)  
input_n_batch1 = 11  
input_n_batch2 = 11
```

**Case2:** phos data, input\_mode = 2

jump\_batchID.params:

```
input_path_batch1                                     =  
/hpcf/authorized_apps/proteomics_apps/pipeline/release/SampleData/phos_test/b1/sum  
_HH_tmt_mod/IDwDecoy_mod.txt
```

```
input_path_batch2                                     =  
/hpcf/authorized_apps/proteomics_apps/pipeline/release/SampleData/phos_test/b2/sum  
_HH_tmt_mod/IDwDecoy_mod.txt  
mods = STY
```

It can be edited: /ParameterFiles/jump\_l.params

It can be edited: /ParameterFiles/TMThhpho/jump\_qj\_HH\_pho\_tmt10\_human.params

jump\_batchQ.params:

```
input_mode = 2           # 2: pho site  
input_n_batch1 = 11  
input_n_batch2 = 11
```

**Case3:** phos data, input\_mode = 3

jump\_batchID.params:

```
input_path_batch1                                     =  
/hpcf/authorized_apps/proteomics_apps/pipeline/release/SampleData/phos_test/b1/sum  
_HH_tmt_mod/IDwDecoy_mod.txt  
input_path_batch2                                     =  
/hpcf/authorized_apps/proteomics_apps/pipeline/release/SampleData/phos_test/b2/sum  
_HH_tmt_mod/IDwDecoy_mod.txt  
mods = STY
```

It can be edited: /ParameterFiles/TMThhpho/jump\_qj\_HH\_pho\_tmt10\_human.params

jump\_batchQ.params:

```
input_mode = 3           # 3: peptides (from either phosphor- or whole proteome)  
input_n_batch1 = 11  
input_n_batch2 = 11
```

**Case4:** comet data, input\_mode = 1

jump\_batchID.params:

```
input_path_batch1                                     =  
/hpcf/authorized_apps/proteomics_apps/pipeline/release/SampleData/comet_test/b1/su  
m_HH_tmt/IDwDecoy.txt
```

```
input_path_batch2                                     =  
/hpcf/authorized_apps/proteomics_apps/pipeline/release/SampleData/comet_test/b2/su  
m_HH_tmt/IDwDecoy.txt  
min_Jscore = 1  
mods = 0  
pit_file                                              =  
/home/yli4/database/MS_proteomics/2017Feb_human_comprehensiveDB/humanComp  
rehensive_v1_ft_mc2_c57_TMT_K229.pit
```

It can be edited: /ParameterFiles/TMThh/jump\_qj\_HH\_tmt10\_human.params

jump\_batchQ.params:

```
input_mode = 1          # 1: proteins (for whole proteome)
```

## 2.2 Troubleshoot

(1) If you cannot use jump as below, run 'module load jump/1.13.003' first.

```
[zyuanl@splprhpc05 testBatch]$ jump -batch-id jump_batchID.params  
-bash: jump: command not found
```

(2) If batch numbers of 'input\_path\_batch' in 'jump\_batchID.params', 'input\_n\_batch' in 'jump\_batchQ.params', and 'internal\_standard\_batch' in 'jump\_batchQ.params' are not the same, the workflow will stop and ask you to make them consistent.

(3) If 'path\_batch\_id' of 'jump\_batchQ.params' is not consistent with 'output\_folder' of 'jump\_batchID.params', it will stop and ask you to make them consistent.

(4) If you make changes in jump\_batchID.params and jump\_batchQ.params in the work path, jump\_qj\_HH\_tmt10\_human.params in /ParameterFiles/TMThh, the changes will be replaced if you rerun 'jump -params'.

(5) Please be patient when applying a compute node with required RAM as sometimes there is long waiting time.