

Title: DEAF1_CB_Output

Report generated at Sun Sep 30 21:09:46 CDT 2018

Command line arguments

-out_dir DEAF1_CB_Output -extsize_macs2 200 -species hg38 -se -nth 8 -fastq1:1 /gpfs/gpfs2/home/jloupe/Flowcells/HFCF3DMXX/30M_trimmed_fastq/30M.GSLv5-8_i7_82-GSLv5-8_i5_15.merge.fastq.gz -ctl_fastq1:1 /gpfs/gpfs2/home/jloupe/Flowcells/30M_controls/30M_dedup_trim_CB_1224_control.fastq.gz -ctl_fastq2:1 /gpfs/gpfs2/home/jloupe/Flowcells/30M_controls/30M_dedup_trim_CB_1230_control.fastq.gz -fastq2:1 /gpfs/gpfs2/home/jloupe/Flowcells/HFCF3DMXX/30M_trimmed_fastq/30M.GSLv5-8_i7_91-GSLv5-8_i5_06.merge.fastq.gz

Directories and files

[Expand all](#) [Collapse all](#)

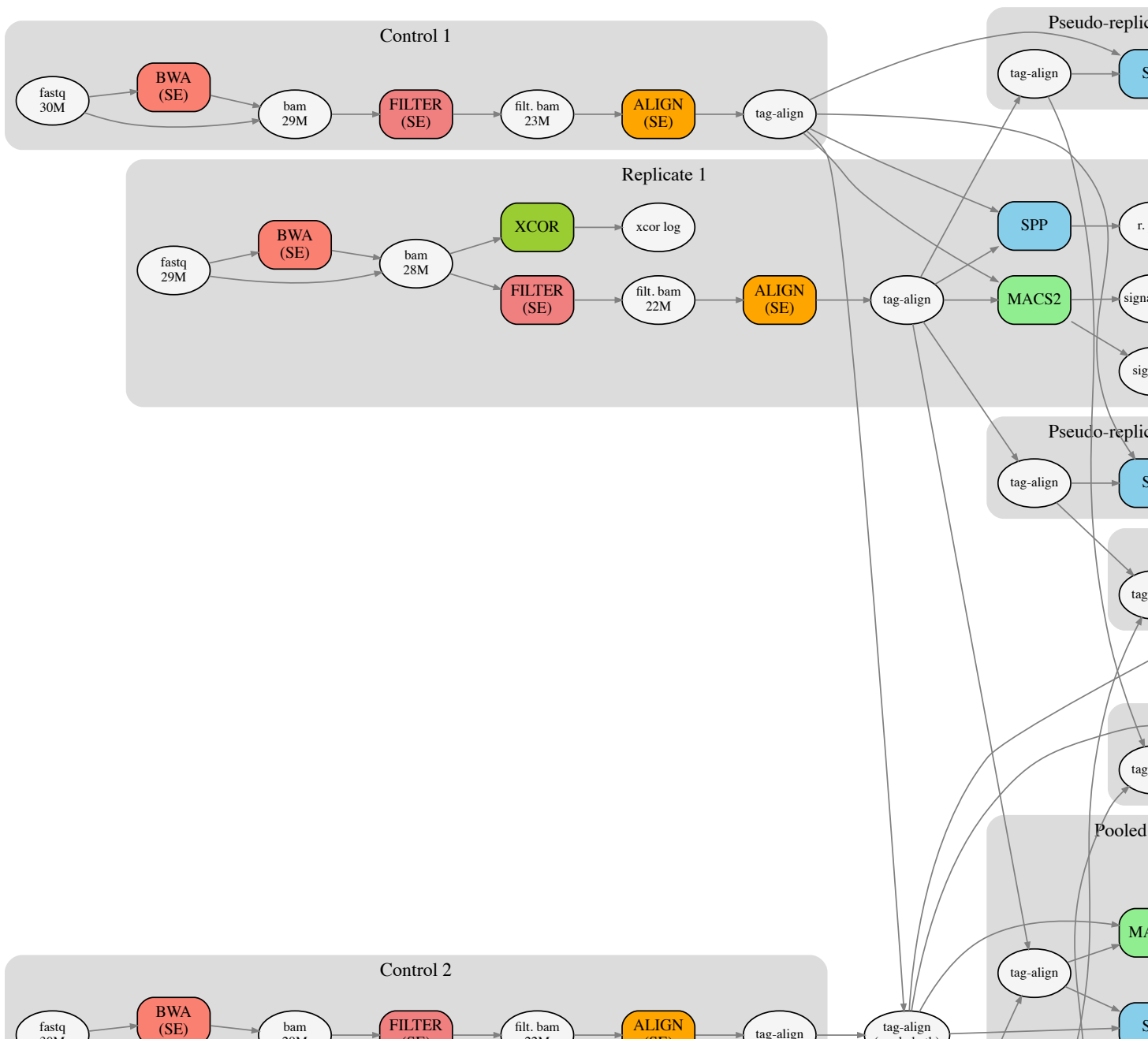
| Files | Path |
|---------------|------|
| Raw reads | |
| Alignment | |
| Signal tracks | |
| Peaks | |
| QC and logs | |

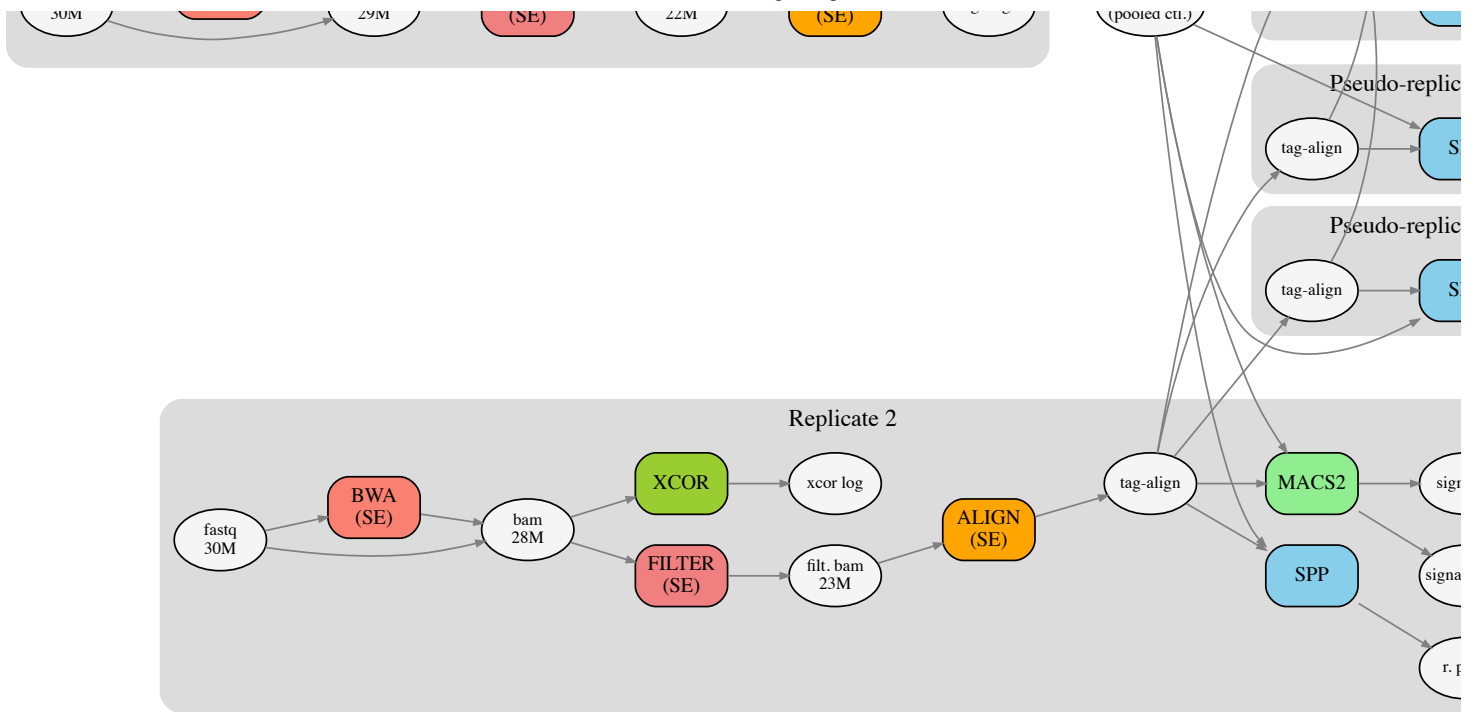
Visualization (add -url_base [URL_ROOT_DIR_FOR_OUT_DIR] to the command line.)

[Visualize](#) [JSON \(datahub\)](#)

Toggle workflow

Workflow diagram (g./b. peak : gapped/broad peak from MACS2, n. peak : narrow peak from MACS2, r. peak : relaxed peak from SPP)





Flagstat (raw) QC

| | rep1 | rep2 | ctl1 | ctl2 |
|----------------------------|----------|----------|----------|----------|
| Total | 29577400 | 30000000 | 30000000 | 30000000 |
| Total(QC-failed) | 0 | 0 | 0 | 0 |
| Dupes | 0 | 0 | 0 | 0 |
| Dupes(QC-failed) | 0 | 0 | 0 | 0 |
| Mapped | 28605917 | 28982084 | 29357542 | 29118134 |
| Mapped(QC-failed) | 0 | 0 | 0 | 0 |
| % Mapped | 96.72 | 96.61 | 97.86 | 97.06 |
| Paired | 0 | 0 | 0 | 0 |
| Paired(QC-failed) | 0 | 0 | 0 | 0 |
| Read1 | 0 | 0 | 0 | 0 |
| Read1(QC-failed) | 0 | 0 | 0 | 0 |
| Read2 | 0 | 0 | 0 | 0 |
| Read2(QC-failed) | 0 | 0 | 0 | 0 |
| Properly Paired | 0 | 0 | 0 | 0 |
| Properly Paired(QC-failed) | 0 | 0 | 0 | 0 |
| % Properly Paired | -nan | -nan | -nan | -nan |
| With itself | 0 | 0 | 0 | 0 |
| With itself(QC-failed) | 0 | 0 | 0 | 0 |
| Singletons | 0 | 0 | 0 | 0 |
| Singletons(QC-failed) | 0 | 0 | 0 | 0 |
| % Singleton | -nan | -nan | -nan | -nan |
| Diff. Chrms | 0 | 0 | 0 | 0 |
| Diff. Chrms (QC-failed) | 0 | 0 | 0 | 0 |

Flagstat (filtered) QC

| | rep1 | rep2 | ctl1 | ctl2 |
|----------------------------|----------|----------|----------|----------|
| Total | 22038276 | 23017171 | 23521900 | 22756670 |
| Total(QC-failed) | 0 | 0 | 0 | 0 |
| Dupes | 0 | 0 | 0 | 0 |
| Dupes(QC-failed) | 0 | 0 | 0 | 0 |
| Mapped | 22038276 | 23017171 | 23521900 | 22756670 |
| Mapped(QC-failed) | 0 | 0 | 0 | 0 |
| % Mapped | 100.00 | 100.00 | 100.00 | 100.00 |
| Paired | 0 | 0 | 0 | 0 |
| Paired(QC-failed) | 0 | 0 | 0 | 0 |
| Read1 | 0 | 0 | 0 | 0 |
| Read1(QC-failed) | 0 | 0 | 0 | 0 |
| Read2 | 0 | 0 | 0 | 0 |
| Read2(QC-failed) | 0 | 0 | 0 | 0 |
| Properly Paired | 0 | 0 | 0 | 0 |
| Properly Paired(QC-failed) | 0 | 0 | 0 | 0 |
| % Properly Paired | N/A | N/A | N/A | N/A |
| With itself | 0 | 0 | 0 | 0 |
| With itself(QC-failed) | 0 | 0 | 0 | 0 |
| Singletons | 0 | 0 | 0 | 0 |
| Singletons(QC-failed) | 0 | 0 | 0 | 0 |
| % Singleton | N/A | N/A | N/A | N/A |
| Diff. Chrms | 0 | 0 | 0 | 0 |
| Diff. Chrms (QC-failed) | 0 | 0 | 0 | 0 |

Dup. QC

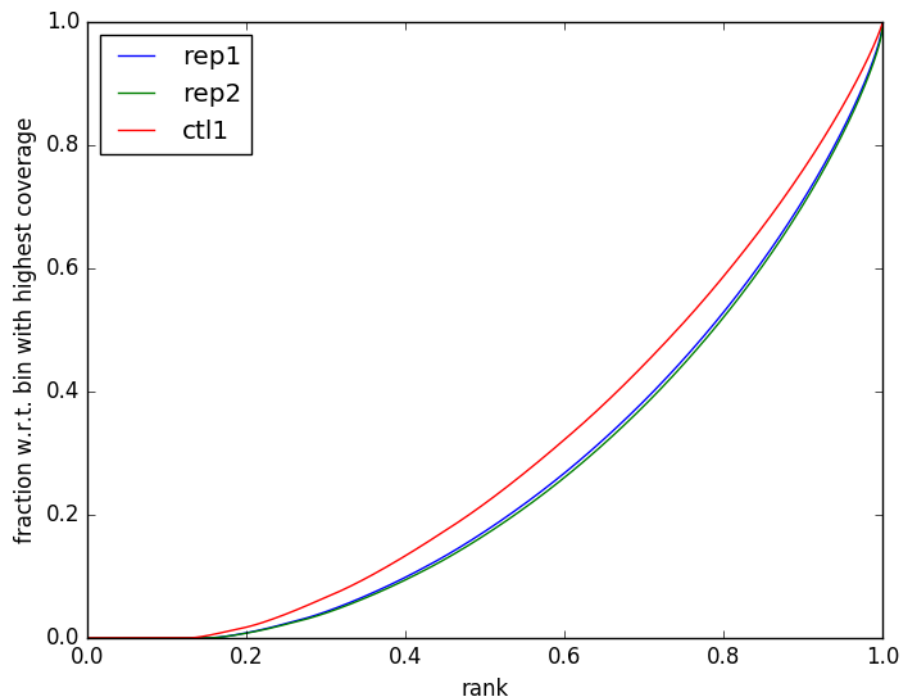
| | rep1 | rep2 | ctl1 | ctl2 |
|-------------------|----------|----------|----------|----------|
| Unpaired Reads | 24899446 | 25169921 | 25321181 | 24502343 |
| Paired Reads | 0 | 0 | 0 | 0 |
| Unmapped Reads | 0 | 0 | 0 | 0 |
| Unpaired Dupes | 2861170 | 2152750 | 1799281 | 1745673 |
| Paired Dupes | 0 | 0 | 0 | 0 |
| Paired Opt. Dupes | 0 | 0 | 0 | 0 |
| % Dupes/100 | 0.114909 | 0.085529 | 0.071058 | 0.071245 |

Fingerprint JS metric

| | rep1 | rep2 |
|-------------------|----------------|----------------|
| % genome enriched | 0.267256782878 | 0.262609918705 |
| AUC | 0.491778891849 | 0.491959622683 |

| | | |
|-----------------------|--------------------|--------------------|
| CHANCE divergence | 0.148066472867 | 0.148248409553 |
| Elbow Point | 3.33834224958e-126 | 5.48866174975e-132 |
| JS Distance | 0.576135504844 | 0.581319700744 |
| Synthetic AUC | 0.501335321485 | 0.493063525656 |
| Synthetic Elbow Point | 0.0649779436337 | 0.077329672717 |
| Synthetic JS Distance | 0.26726312395 | 0.277092425357 |
| Synthetic X-intercept | 36.3675440947 | 35.4996461231 |
| X-intercept | 19.9695673449 | 20.603817931 |
| diff. enrichment | 0.173294339473 | 0.178463494492 |

Fingerprints of different samples



Fingerprint plot

Library Complexity QC

| | rep1 | rep2 | ctl1 | ctl2 |
|-------------------------|----------|-----------|-----------|-----------|
| Total Reads | 24886609 | 25142330 | 25302381 | 24485306 |
| Distinct Reads | 22138069 | 23136238 | 23552323 | 22789943 |
| One Read | 19663035 | 21281681 | 21919959 | 21208560 |
| Two Reads | 2232857 | 1719641 | 1544927 | 1493366 |
| NRF = Distinct/Total | 0.889557 | 0.920211 | 0.930834 | 0.930760 |
| PBC1 = OneRead/Distinct | 0.888200 | 0.919842 | 0.930692 | 0.930610 |
| PBC2 = OneRead/TwoReads | 8.806222 | 12.375653 | 14.188346 | 14.201850 |

NRF (non redundant fraction)

PBC1 (PCR Bottleneck coefficient 1)

PBC2 (PCR Bottleneck coefficient 2)

PBC1 is the primary measure. Provisionally

- 0-0.5 is severe bottlenecking
- 0.5-0.8 is moderate bottlenecking
- 0.8-0.9 is mild bottlenecking
- 0.9-1.0 is no bottlenecking

Enrichment QC (strand cross-correlation measures)

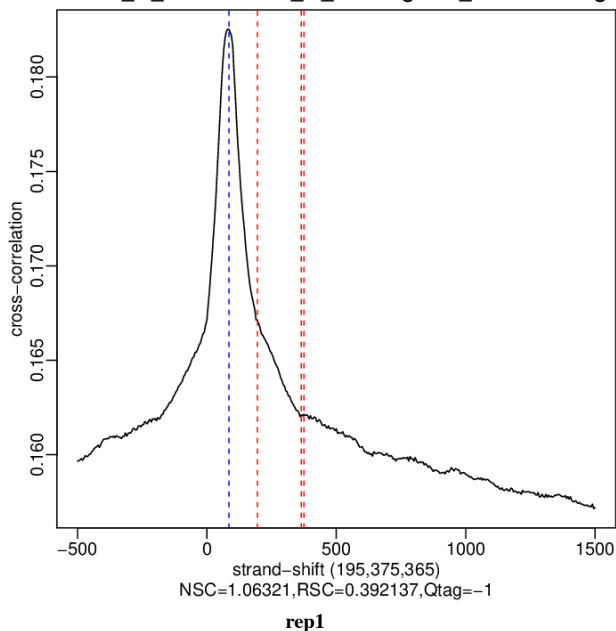
| | rep1 | rep2 |
|--------------------------|-------------------|-------------------|
| Reads | 15000000 | 15000000 |
| Est. Fragment Len. | 195 | -155 |
| Corr. Est. Fragment Len. | 0.167132010658164 | 0.167569073980666 |
| Phantom Peak | 85 | 85 |
| Corr. Phantom Peak | 0.1825343 | 0.191337 |
| Argmin. Corr. | 1500 | 1500 |
| Min. Corr. | 0.1571959 | 0.1621751 |
| NSC | 1.063209 | 1.03326 |

RSC | 0.392137 | 0.1849652

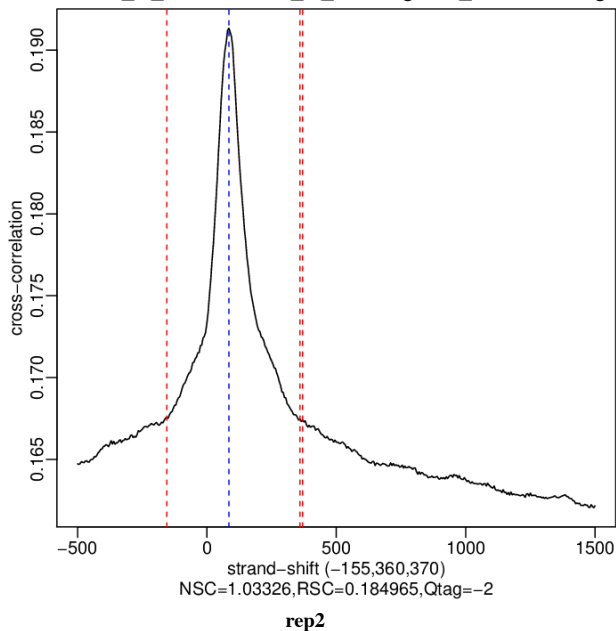
NOTE: Reads from replicates are subsampled to a max of 15M

- Normalized strand cross-correlation coefficient (NSC) = col9 in outFile
- Relative strand cross-correlation coefficient (RSC) = col10 in outFile
- Estimated fragment length = col3 in outFile, take the top value

30M.GSLv5-8_i7_82-GSLv5-8_i5_15.merge.no_chrM.15M.tagAl



30M.GSLv5-8_i7_91-GSLv5-8_i5_06.merge.no_chrM.15M.tagAl



Number of peaks

| | |
|----------|--------|
| rep1 | 300000 |
| rep1-pr1 | 300000 |
| rep1-pr2 | 300000 |
| rep2 | 300000 |
| rep2-pr1 | 300000 |
| rep2-pr2 | 300000 |
| pooled | 300000 |
| ppr1 | 300000 |
| ppr2 | 300000 |

- ppr1: Raw peaks called on the first pooled pseudoreplicates
- ppr2: Raw peaks called on the second pooled pseudoreplicates
- repi: Raw peaks called on true replicate i
- repi-pr1 : Raw peaks called on the first pseudoreplicate from replicate i
- repi-pr2 : Raw peaks called on the second pseudoreplicates from replicate i
- overlap : Overlapped peaks (filtered if blacklist exists)

Enrichment QC (Fraction of reads in raw peaks)

| | pooled_rep | rep1 | rep2 |
|----------------------------------|----------------------------|----------------------|----------------------|
| Fraction of Reads in Peak | 0.11608 | 0.139145 | 0.0938979 |

Enrichment QC (Fraction of reads in overlapping peaks)

| | rep1-pr | rep1-rep2 | rep2-pr |
|----------------------------------|-------------------------|---------------------------|-------------------------|
| Fraction of Reads in Peak | 0.0348961 | 0.0320476 | 0.0305856 |

- ppr: Overlapping peaks comparing pooled pseudo replicates
- rep1-pr: Overlapping peaks comparing pseudoreplicates from replicate 1
- rep2-pr: Overlapping peaks comparing pseudoreplicates from replicate 2
- repi-repj: Overlapping peaks comparing true replicates (rep i vs. rep j)

Reproducibility QC and Peak Detection Statistics (Overlapping peaks)

| | |
|-------------------------------|--------------------|
| Nt | 47946 |
| N1 | 46515 |
| N2 | 67445 |
| Np | 51477 |
| N optimal | 51477 |
| N conservative | 47946 |
| Optimal Set | pooled_pseudo_rep |
| Conservative Set | rep1-rep2 |
| Rescue Ratio | 1.0736453510198973 |
| Self Consistency Ratio | 1.4499623777276147 |
| Reproducibility Test | pass |

- N1: Replicate 1 self-consistent overlapping peaks (comparing two pseudoreplicates generated by subsampling Rep1 reads)
- N2: Replicate 2 self-consistent overlapping peaks (comparing two pseudoreplicates generated by subsampling Rep2 reads)
- Nt: True Replicate consistent overlapping peaks (comparing true replicates Rep1 vs Rep2)
- Np: Pooled-pseudoreplicate consistent overlapping peaks (comparing two pseudoreplicates generated by subsampling pooled reads from Rep1 and Rep2)
- Self-consistency Ratio: $\max(N1,N2) / \min(N1,N2)$
- Rescue Ratio: $\max(Np,Nt) / \min(Np,Nt)$
- Reproducibility Test: If Self-consistency Ratio >2 AND Rescue Ratio > 2, then 'Fail' else 'Pass'

Enrichment QC (Fraction of reads in IDR peaks)

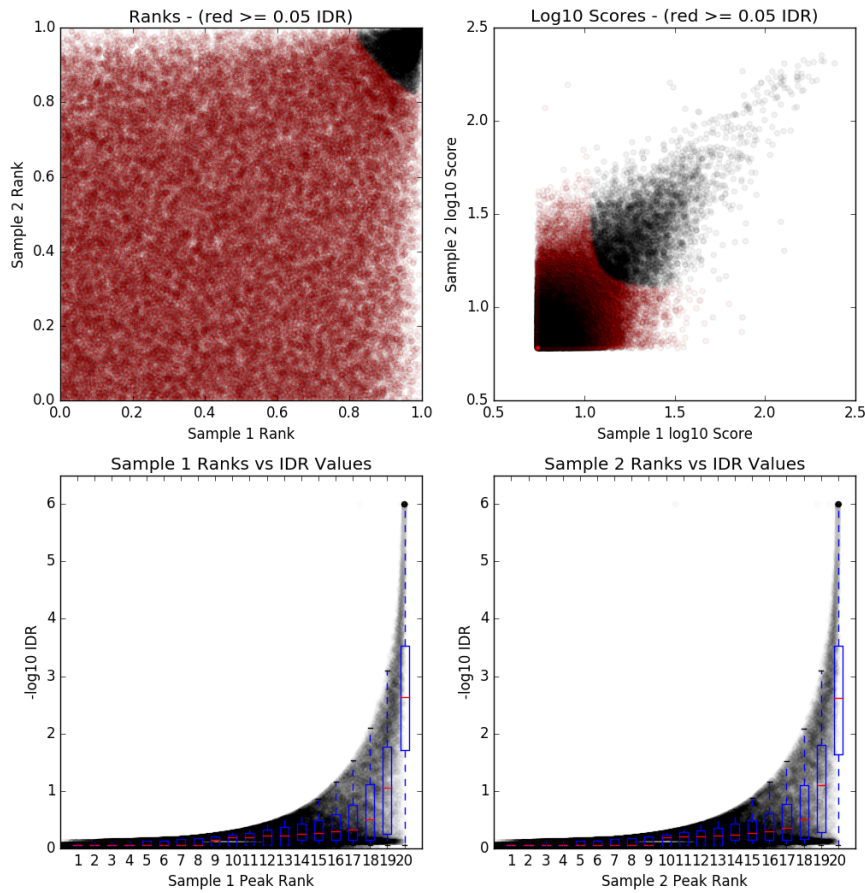
| | rep1-pr | rep1-rep2 | rep2-pr |
|----------------------------------|-------------------------|---------------------------|-------------------------|
| Fraction of Reads in Peak | 0.00842223 | 0.0114135 | 0.00770072 |

- ppr: IDR peaks comparing pooled pseudo replicates
- rep1-pr: IDR peaks comparing pseudoreplicates from replicate 1
- rep2-pr: IDR peaks comparing pseudoreplicates from replicate 2
- repi-repj: IDR peaks comparing true replicates (rep i vs. rep j)

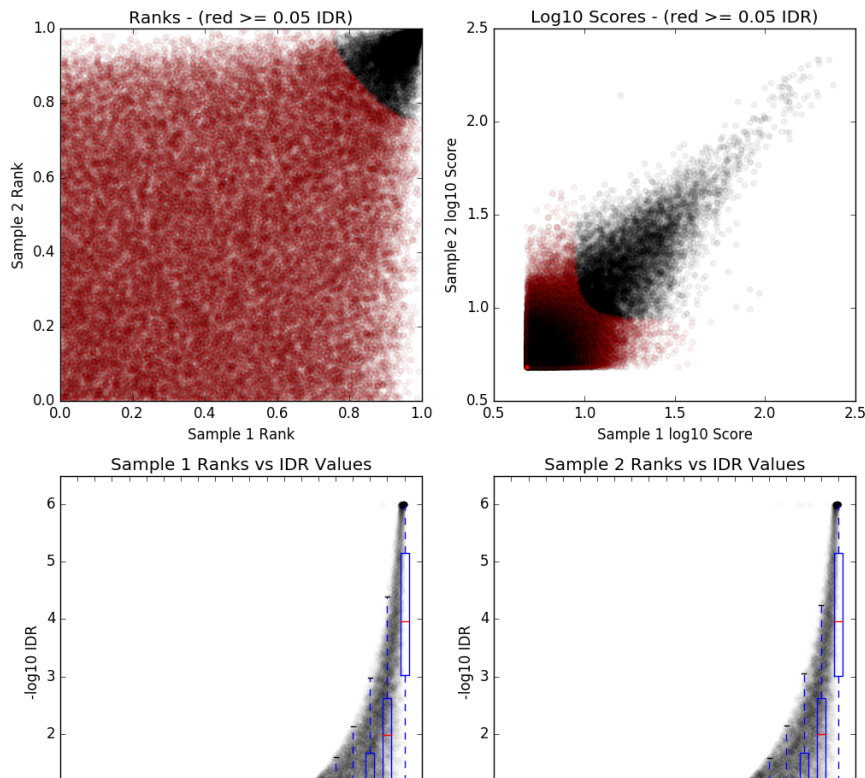
Reproducibility QC and Peak Detection Statistics (Irreproducible Discovery Rate)

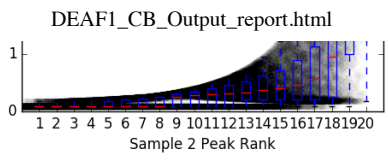
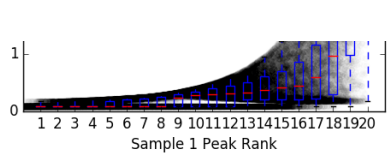
| | |
|-------------------------------|--------------------|
| Nt | 5219 |
| N1 | 3412 |
| N2 | 4013 |
| Np | 7978 |
| N optimal | 7978 |
| N conservative | 5219 |
| Optimal Set | pooled_pseudo_rep |
| Conservative Set | rep1-rep2 |
| Rescue Ratio | 1.5286453343552404 |
| Self Consistency Ratio | 1.176143024618992 |
| Reproducibility Test | pass |

- N1: Replicate 1 self-consistent IDR 0.05 peaks (comparing two pseudoreplicates generated by subsampling Rep1 reads)
- N2: Replicate 2 self-consistent IDR 0.05 peaks (comparing two pseudoreplicates generated by subsampling Rep2 reads)
- Nt: True Replicate consistent IDR 0.05 peaks (comparing true replicates Rep1 vs Rep2)
- Np: Pooled-pseudoreplicate consistent IDR 0.05 peaks (comparing two pseudoreplicates generated by subsampling pooled reads from Rep1 and Rep2)
- Self-consistency Ratio: $\max(N1,N2) / \min(N1,N2)$
- Rescue Ratio: $\max(Np,Nt) / \min(Np,Nt)$
- Reproducibility Test: If Self-consistency Ratio >2 AND Rescue Ratio > 2, then 'Fail' else 'Pass'

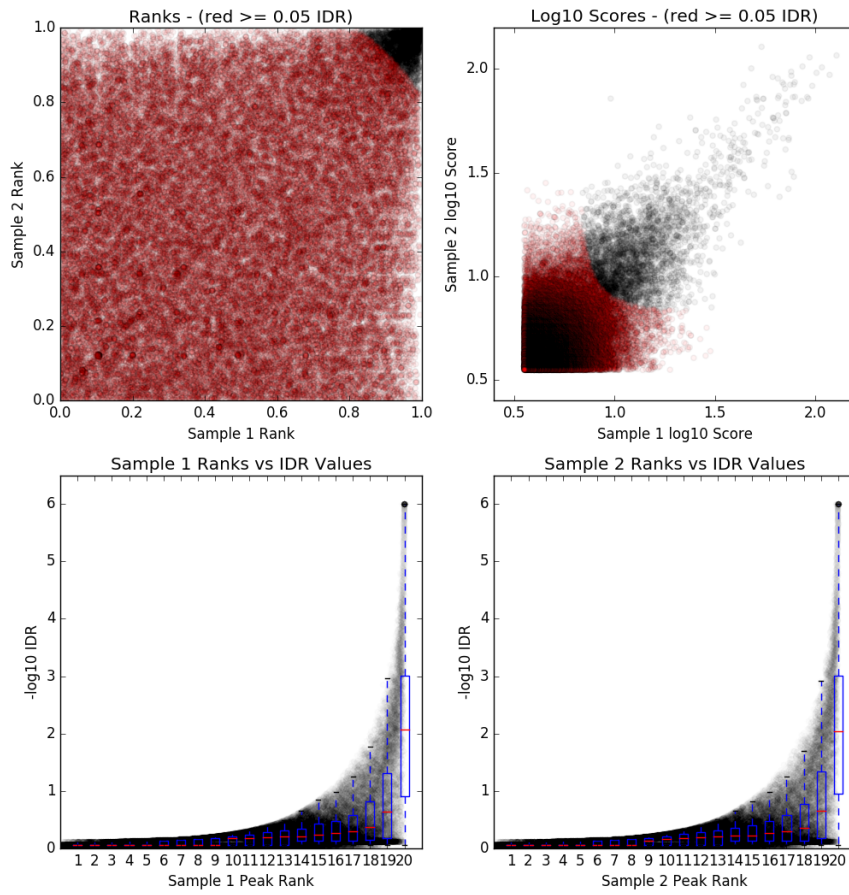


true reps (rep1-rep2)

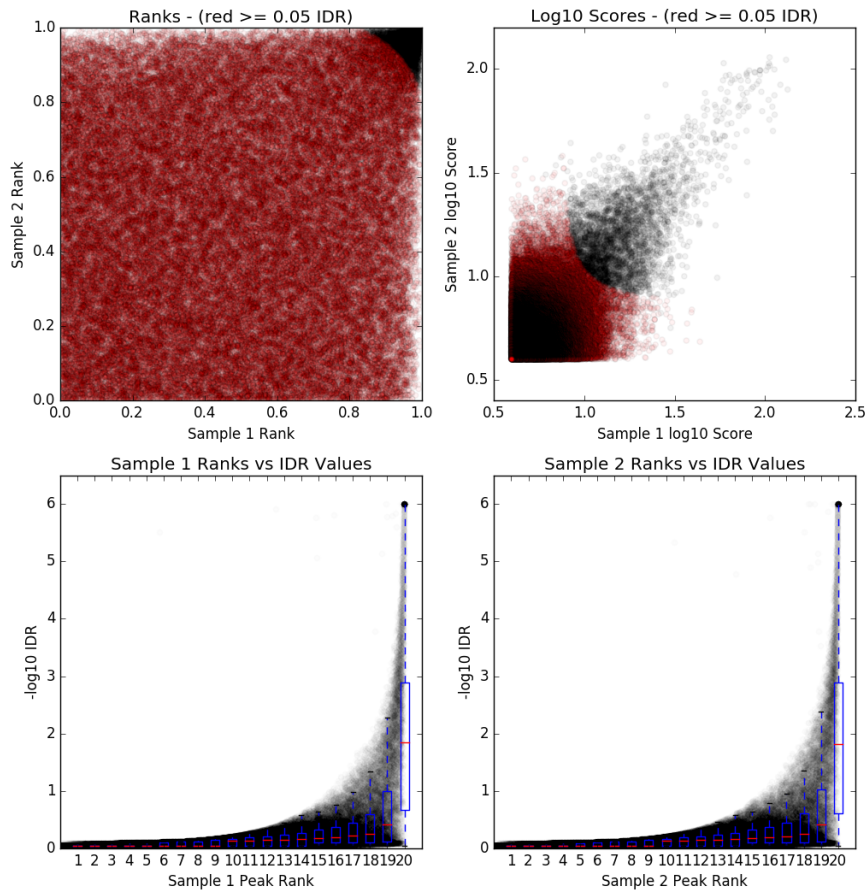




pooled pseudo-reps



rep1 pseudo-reps



rep2 pseudo-reps