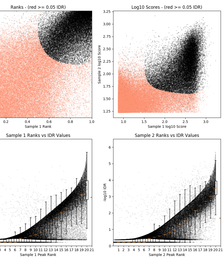
QC Report

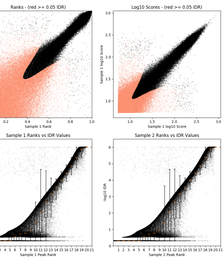
general	
Report generated at	2025-10-23 07:33:08
Title	ENCSR214HRO vs. ENCSR214AHF
Description	Comparison of Chr19, FE2 and ChIP-seq (SK-N_BE)
Pipeline version	1.4.0
Pipeline type	chip
Genome	hg38
Organism	human
Sequencing endedness	rep1: [paired_end?: True], rep2: [paired_end?: True], ctl1: [paired_end?: True], ctl2: [paired_end?: True]
Peak caller	macs2

Replication quality metrics

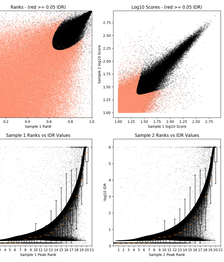
IDR (Irreproducible Discovery Rate) plots



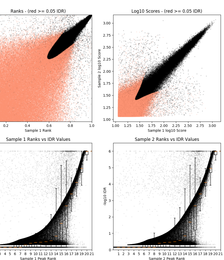
rep1_vs_rep2



rep1_vs_rep1_pp2



rep2_pr1_vs_rep2_pp2



pooled-pr1_vs_pooled-pr2

Reproducibility QC and peak detection statistics

	overlap	idr
N1	86114	200634
N2	86116	92113
N3	105884	40191
N4	105884	92122
N optimal	105884	92122
N conservative	86114	200634
Optimal Set	pooled-pr1_vs_pooled-pr2	pooled-pr1_vs_pooled-pr2
Conservative Set	rep1_vs_rep2	rep1_vs_rep2
Recall Ratio	0.5429688671113	2.2117267297558
IoP Consistency Ratio	4.7679917067411	1.9114815628874
Reproducibility Test	Irreducibe	pass

Reproducibility QC

- N1: Replicate 1 self-consistent peaks (comparing two pseudoreplicates generated by subsampling Rep1 reads)
- N2: Replicate 2 self-consistent peaks (comparing two pseudoreplicates generated by subsampling Rep2 reads)
- N3: True Replicate consistent peaks (comparing two replicates Rep1 vs Rep2 generated by subsampling Rep1 reads)
- N4: True Replicate consistent peaks (comparing two replicates Rep1 vs Rep2)
- Np: Pooled pseudoreplicate consistent peaks (comparing two pseudoreplicates generated by subsampling pooled reads from Rep1 and Rep2)
- Recall Ratio: max(Nc,Nt)/min(Nc,Nt) (Nc: Nc/Nc)
- Recall Ratio: max(Np,Nc)/ min (Np,Nc)
- Reproducibility Test: If IoD consistency Ratio (2 AND Rescue Ratio < 2, then Fail' else 'Pass'

Number of raw peaks

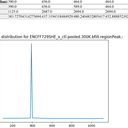
	rep1	rep2
Number of peaks	2701351	2595620

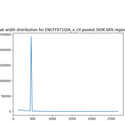
Top 500000 raw peaks from rep1 with FDR 5.01

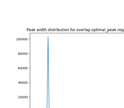
Peak calling statistics

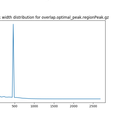
Peak region size

	rep1	rep2	ctl_opt	ctrl0a_opt
Min size	50.0	113.0	126.0	126.0
25 percentile	290.0	454.0	526.0	404.0
50 percentile (median)	590.0	670.0	664.0	494.0
75 percentile	1126.0	1126.0	1126.0	1126.0
Max size	1125.0	1687.0	3693.0	3693.0
Mean	921.7276612279694	977.3556348409256	1091.2404417126472	1032.8888722829648









Enrichment / Signal-to-noise ratio

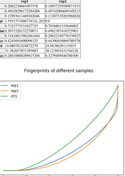
Strand cross-correlation measures (trimmed/filtered SE BAM)

	rep1	rep2
Number of Sequenced Reads	17777172	17777172
Estimated Fragment Length	197	100
Cross-correlation at Estimated Fragment Length	0.28071237168416	0.28070711180577
Phantom Peak	50	50
Cross-correlation at Phantom Peak	0.149931	0.150005
Minimum of Cross-correlation	-0.5387	-0.5387
Minimum of Cross-correlation	0.04120878	0.0412721
NSC (Normalized Strand Cross-correlation coeff.)	1.49963	1.53874
RSC (Relative Strand Cross-correlation coeff.)	1.77232	1.09975

Performed on subsampled (15000000) reads mapped from FASTQs that are trimmed to 50. Such FASTQ trimming and subsampling reads are for cross-correlation analysis only. Untrimmed FASTQs are used for all the other analyses.

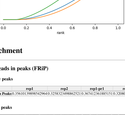
NOTE1: For SE datasets, reads from replicates are randomly subsampled to 15000000.
NOTE2: For PE datasets, the first and R1 of each mate pair is selected and trimmed to 50. The reads are then randomly subsampled to 15000000.

- Normalized strand cross-correlation coefficient (NSC) = coff/ne-coff/ne
- Relative strand cross-correlation coefficient (RSC) = coff/ne-coff/ne
- Estimated fragment length = coff/ne-coff/ne, take the top value



Jensen-Shannon distance (filtered/deduped BAM)

	rep1	rep2
AUC	0.478125000000000	0.478125000000000
Synthetic AUC	0.494296296296296	0.497428888888889
X intercept	0.139761666666667	0.140717333333333
Synthetic X intercept	0.139761666666667	0.140717333333333
Jensen-Shannon	0.675351000000000	0.676468666666667
Synthetic Jsher Peaks	0.711281333333333	0.714833333333333
JS Distance	0.518109333333333	0.519622666666667
Synthetic JS Distance	0.543146666666667	0.545188666666667
N. Genome Enriched	0.208076666666667	0.208196666666667
DHS Enrichment	0.382675333333333	0.383449666666667
CRAMER Divergence	0.281104652996,1,2,0	0.279449666667856



Peak enrichment

Fraction of reads in peaks (FRiP)

FRiP for opp raw peaks

	rep1	rep2	rep1-pr1	rep2-pr1	rep1-pr2	rep2-pr2	pooled	pooled-pr1	pooled-pr2
Fraction of Reads in Peaks	0.596101.5990074129643	0.5761321988062532	0.5671.8638455713	0.5380792.9877777	0.5674517.7772874265	0.5257.5974969743904	0.5671.7969674194	0.5263.7988965183	0.5260.9967966170

FRiP for overlap peaks

	rep1_vs_rep2	rep1-pr1_vs_rep1-pr2	rep1_vs_rep1-pr1	rep1_vs_rep2-pr1	pooled-pr1_vs_pooled-pr2
Fraction of Reads in Peaks	0.6877376142277884	0.7017556337149428	0.7047.5310187749936	0.7714.9777988.7853	0.7714.9777988.7853

FRiP for IDR peaks

	rep1_vs_rep2	rep1-pr1_vs_rep1-pr2	rep2-pr1_vs_rep2-pr2	pooled-pr1_vs_pooled-pr2
Fraction of Reads in Peaks	0.6877876142277884	0.7017556337149428	0.7782.6869588.6253	0.7.2128425733844174

For opp raw peaks:

- rep1N: Peak from true replicate X
- rep1-prN: Peak from self-pseudoreplicates from replicate X
- pooled: Peak from pooled true replicates (pool of rep1, rep2...)
- pooled-pr1: Peak from 1st pooled pseudo replicate (pool of rep1-pr1, rep2-pr1 ...)
- pooled-pr2: Peak from 2nd pooled pseudo replicate (pool of rep1-pr2, rep2-pr2, ...)

For overlap/IDR peaks:

- rep1_vs_rep2: Comparing two peaks from two replicates X and Y
- rep1-pr1_vs_rep1-pr2: Comparing two peaks from both pseudoreplicates from replicate X
- pooled-pr1_vs_pooled-pr2: Comparing two peaks from 1st and 2nd pooled pseudo replicates