

splitfq

↩ input

<input function>

<input function>

output →

workup/splitfq/{sample,{\"\\\".+}\_R1.part\_000}.fastq

workup/splitfq/{sample,{\"\\\".+}\_R1.part\_001}.fastq

workup/splitfq/{sample,{\"\\\".+}\_R2.part\_000}.fastq

workup/splitfq/{sample,{\"\\\".+}\_R2.part\_001}.fastq

compress\_fastq

↩ input

workup/splitfq/{sample}\_R1.part\_{splitid}.fastq

workup/splitfq/{sample}\_R2.part\_{splitid}.fastq

output →

workup/splitfq/{sample,{\"\\\".+}\_R1.part\_{splitid}}.fastq.gz

workup/splitfq/{sample,{\"\\\".+}\_R2.part\_{splitid}}.fastq.gz

adaptor trimming\_pe

↩ input

workup/splitfq/{sample}\_R1.part\_{splitid}.fastq.gz

workup/splitfq/{sample}\_R2.part\_{splitid}.fastq.gz

output →

workup/trimmed/{sample,{\"\\\".+}\_R1.part\_{splitid}}.fastq.gz.trimming\_report.txt

workup/trimmed/{sample,{\"\\\".+}\_R1.part\_{splitid}}\_val\_1.fq.gz

workup/trimmed/{sample,{\"\\\".+}\_R2.part\_{splitid}}.fastq.gz.trimming\_report.txt

workup/trimmed/{sample,{\"\\\".+}\_R2.part\_{splitid}}\_val\_2.fq.gz

barcode\_id

↩ input

workup/trimmed/{sample}\_R1.part\_{splitid}\_val\_1.fq.gz

workup/trimmed/{sample}\_R2.part\_{splitid}\_val\_2.fq.gz

output →

workup/fastqs/{sample,{\"\\\".+}\_R1.part\_{splitid}}.barcoded.fastq.gz

workup/fastqs/{sample,{\"\\\".+}\_R2.part\_{splitid}}.barcoded.fastq.gz

split\_bpm\_dpm

↩ input

workup/fastqs/{sample}\_R1.part\_{splitid}.barcoded.fastq.gz

output →

workup/fastqs/{sample,{\"\\\".+}\_R1.part\_{splitid}}.barcoded\_bpm.fastq.gz

workup/fastqs/{sample,{\"\\\".+}\_R1.part\_{splitid}}.barcoded\_dpm.fastq.gz

workup/fastqs/{sample,{\"\\\".+}\_R1.part\_{splitid}}.barcoded\_other.fastq.gz

workup/fastqs/{sample,{\"\\\".+}\_R1.part\_{splitid}}.barcoded\_short.fastq.gz

cutadapt\_dpm

↩ input

workup/fastqs/{sample}\_R1.part\_{splitid}.barcoded\_dpm.fastq.gz

output →

workup/trimmed/{sample,{\"\\\".+}\_R1.part\_{splitid}}.barcoded\_dpm.RDtrim.fastq.gz

workup/trimmed/{sample,{\"\\\".+}\_R1.part\_{splitid}}.barcoded\_dpm.RDtrim.qc.txt

bowtie2\_align

↩ input

workup/trimmed/{sample}\_R1.part\_{splitid}.barcoded\_dpm.RDtrim.fastq.gz

output →

workup/alignments\_parts/{sample,{\"\\\".+}.part\_{splitid}}.DNA.bowtie2.mapq20.bam

workup/alignments\_parts/{sample,{\"\\\".+}.part\_{splitid}}.unsorted.bam

rename\_and\_filter\_chr

↩ input

workup/alignments\_parts/{sample}.part\_{splitid}.DNA.bowtie2.mapq20.bam

output →

workup/alignments\_parts/{sample,{\"\\\".+}.part\_{splitid}}.DNA.chr.bam

cutadapt\_oligo

↩ input

workup/fastqs/{sample}\_R1.part\_{splitid}.barcoded\_bpm.fastq.gz

output →

workup/trimmed/{sample,{\"\\\".+}\_R1.part\_{splitid}}.barcoded\_bpm.RDtrim.fastq.gz

workup/trimmed/{sample,{\"\\\".+}\_R1.part\_{splitid}}.barcoded\_bpm.RDtrim.qc.txt

merge\_mask

output →

workup/mask\_merge.bed

fastq\_to\_bam

↩ input

workup/trimmed/{sample}\_R1.part\_{splitid}.barcoded\_bpm.RDtrim.fastq.gz

workup/alignments\_parts/{sample,{\"\\\".+}.part\_{splitid}}.BPM.bam

workup/alignments\_parts/{sample,{\"\\\".+}.part\_{splitid}}.BPM.unsorted.bam

output →

repeat\_mask

↩ input

workup/alignments\_parts/{sample}.part\_{splitid}.DNA.chr.bam

workup/mask\_merge.bed

output →

workup/alignments\_parts/{sample,{\"\\\".+}.part\_{splitid}}.DNA.chr.masked.bam

make\_clusters

↩ input

workup/alignments\_parts/{sample}.part\_{splitid}.BPM.bam

workup/alignments\_parts/{sample}.part\_{splitid}.DNA.chr.masked.bam

output →

workup/clusters\_parts/{sample,{\"\\\".+}.part\_{splitid}}.clusters

workup/clusters\_parts/{sample,{\"\\\".+}.part\_{splitid}}.unsorted.clusters

merge\_clusters

↩ input

workup/clusters\_parts/{sample}.part\_000.clusters

workup/clusters\_parts/{sample}.part\_001.clusters

output →

workup/clusters/{sample,{\"\\\".+}.clusters

workup/clusters/{sample,{\"\\\".+}.duplicated.clusters

merge\_dna

↩ input

workup/alignments\_parts/{sample}.part\_000.DNA.chr.masked.bam

workup/alignments\_parts/{sample}.part\_001.DNA.chr.masked.bam

output →

workup/alignments/{sample,{\"\\\".+}.DNA.merged.bam

merge\_beads

↩ input

workup/alignments\_parts/{sample}.part\_000.BPM.bam

workup/alignments\_parts/{sample}.part\_001.BPM.bam

output →

workup/alignments/{sample,{\"\\\".+}.merged.BPM.bam

thresh\_and\_split

↩ input

workup/alignments/{sample}.DNA.merged.bam

workup/clusters/{sample}.clusters

output →

workup/alignments/{sample,{\"\\\".+}.DNA.merged.labeled.bam

pipeline\_counts

↩ input

workup/alignments/sample1.DNA.merged.bam

workup/alignments/sample1.DNA.merged.labeled.bam

workup/alignments/sample1.merged.BPM.bam

workup/alignments/sample2.DNA.merged.bam

workup/alignments/sample2.DNA.merged.labeled.bam

workup/alignments/sample2.merged.BPM.bam

workup/alignments\_parts/sample1.part\_000.BPM.bam

workup/alignments\_parts/sample1.part\_000.DNA.bowtie2.mapq20.bam

workup/alignments\_parts/sample1.part\_000.DNA.chr.bam

workup/alignments\_parts/sample1.part\_000.DNA.chr.masked.bam

workup/alignments\_parts/sample1.part\_001.BPM.bam

workup/alignments\_parts/sample1.part\_001.DNA.bowtie2.mapq20.bam

workup/alignments\_parts/sample1.part\_001.DNA.chr.bam

workup/alignments\_parts/sample1.part\_001.DNA.chr.masked.bam

workup/alignments\_parts/sample2.part\_000.BPM.bam

workup/alignments\_parts/sample2.part\_000.DNA.bowtie2.mapq20.bam

workup/alignments\_parts/sample2.part\_000.DNA.chr.bam

workup/alignments\_parts/sample2.part\_000.DNA.chr.masked.bam

workup/alignments\_parts/sample2.part\_001.BPM.bam

workup/alignments\_parts/sample2.part\_001.DNA.bowtie2.mapq20.bam

workup/alignments\_parts/sample2.part\_001.DNA.chr.bam

workup/alignments\_parts/sample2.part\_001.DNA.chr.masked.bam

workup/clusters/sample1.clusters

workup/clusters\_parts/sample1.part\_000.clusters

workup/clusters\_parts/sample1.part\_001.clusters

workup/clusters\_parts/sample2.part\_000.clusters

workup/clusters\_parts/sample2.part\_001.clusters

workup/fastqs/sample1\_R1.part\_000.barcoded.fastq.gz

workup/fastqs/sample1\_R1.part\_000.barcoded\_bpm.fastq.gz

workup/fastqs/sample1\_R1.part\_000.barcoded\_dpm.fastq.gz

workup/fastqs/sample1\_R1.part\_001.barcoded.fastq.gz

workup/fastqs/sample1\_R1.part\_001.barcoded\_bpm.fastq.gz

workup/fastqs/sample1\_R1.part\_001.barcoded\_dpm.fastq.gz

workup/fastqs/sample1\_R2.part\_000.barcoded.fastq.gz

workup/fastqs/sample1\_R2.part\_001.barcoded.fastq.gz

workup/fastqs/sample2\_R1.part\_000.barcoded.fastq.gz

workup/fastqs/sample2\_R1.part\_000.barcoded\_bpm.fastq.gz

workup/fastqs/sample2\_R1.part\_001.barcoded.fastq.gz

workup/fastqs/sample2\_R1.part\_001.barcoded\_dpm.fastq.gz

workup/fastqs/sample2\_R2.part\_000.barcoded.fastq.gz

workup/fastqs/sample2\_R2.part\_001.barcoded.fastq.gz

workup/splitfq/sample1\_R1.part\_000.fastq.gz

workup/splitfq/sample1\_R1.part\_001.fastq.gz

workup/splitfq/sample1\_R2.part\_000.fastq.gz

workup/splitfq/sample1\_R2.part\_001.fastq.gz

workup/splitfq/sample2\_R1.part\_000.fastq.gz

workup/splitfq/sample2\_R1.part\_001.fastq.gz

workup/splitfq/sample2\_R2.part\_000.fastq.gz

workup/splitfq/sample2\_R2.part\_001.fastq.gz

workup/trimmed/sample1\_R1.part\_000.barcoded\_bpm.RDtrim.fastq.gz

workup/trimmed/sample1\_R1.part\_000.barcoded\_dpm.RDtrim.fastq.gz

workup/trimmed/sample1\_R1.part\_001.barcoded\_bpm.RDtrim.fastq.gz

workup/trimmed/sample1\_R1.part\_001.barcoded\_dpm.RDtrim.fastq.gz

workup/trimmed/sample1\_R2.part\_000\_val\_1.fq.gz

workup/trimmed/sample1\_R2.part\_001\_val\_1.fq.gz

workup/trimmed/sample1\_R2.part\_000\_val\_2.fq.gz

workup/trimmed/sample1\_R2.part\_001\_val\_2.fq.gz

workup/trimmed/sample2\_R1.part\_000.barcoded\_bpm.RDtrim.fastq.gz

workup/trimmed/sample2\_R1.part\_000.barcoded\_dpm.RDtrim.fastq.gz

workup/trimmed/sample2\_R1.part\_001.barcoded\_bpm.RDtrim.fastq.gz

workup/trimmed/sample2\_R1.part\_001.barcoded\_dpm.RDtrim.fastq.gz

workup/trimmed/sample2\_R2.part\_000\_val\_1.fq.gz

workup/trimmed/sample2\_R2.part\_000\_val\_2.fq.gz

workup/trimmed/sample2\_R2.part\_001\_val\_2.fq.gz

output →

workup/pipeline\_counts.txt

workup/qc/pipeline\_counts.csv

generate\_splitbam\_statistics

↩ input

workup/alignments/sample1.DNA.merged.labeled.bam

workup/alignments/sample2.DNA.merged.labeled.bam

output →

workup/splitbams/splitbam\_statistics.txt

generate\_cluster\_ecdfs

↩ input

workup/clusters/sample1.clusters

workup/clusters/sample2.clusters

output →

workup/clusters/Max\_representation\_counts.pdf

workup/clusters/Max\_representation\_ecdf.pdf

get\_size\_distribution

↩ input

workup/clusters/sample1.clusters

workup/clusters/sample2.clusters

output →

workup/clusters/BPM\_cluster\_distribution.pdf

workup/clusters/BPM\_read\_distribution.pdf

workup/clusters/DPM\_cluster\_distribution.pdf

workup/clusters/DPM\_read\_distribution.pdf

generate\_bigwigs

↩ input

workup/logs/splitbams\_all.log

workup/mask\_merge.bed

multiqc

↩ input

workup/clusters/sample1.clusters

workup/clusters/sample2.clusters

output →

workup/qc/multiqc\_report.html

log\_config

output →

workup/logs/config\_2024.05.16.json

get\_ligation\_efficiency

↩ input

workup/fastqs/{sample}\_R1.part\_{splitid}.barcoded.fastq.gz

output →

workup/{sample,{\"\\\".+}.part\_{splitid}}.ligation\_efficiency.txt

cat\_ligation\_efficiency

↩ input

workup/sample1.part\_000.ligation\_efficiency.txt

workup/sample1.part\_001.ligation\_efficiency.txt

workup/sample2.part\_000.ligation\_efficiency.txt

workup/sample2.part\_001.ligation\_efficiency.txt

output →

workup/ligation\_efficiency.txt

all

↩ input

workup/alignments/sample1.merged.BPM.bam

workup/alignments/sample2.merged.BPM.bam

workup/clusters/BPM\_cluster\_distribution.pdf

workup/clusters/BPM\_read\_distribution.pdf

workup/clusters/DPM\_cluster\_distribution.pdf

workup/clusters/DPM\_read\_distribution.pdf

workup/clusters/Max\_representation\_counts.pdf

workup/clusters/Max\_representation\_ecdf.pdf

workup/clusters/cluster\_statistics.txt

workup/ligation\_efficiency.txt

workup/logs/bigwigs.log

workup/logs/config\_2024.05.16.json

workup/pipeline\_counts.txt

workup/qc/multiqc\_report.html