Enhancer Identification JD

All Regions

1. Import All BAM files
   1. Do not remove duplicate reads
   2. Treat as HiC data – unchecked
   3. Min mapping quality - 20
   4. Primary alignments only - checked
   5. Treat as RNA-Seq data – checked
   6. Data Type – Paired end
   7. Pair Distance Cutoff (bp) – n/a
2. Merge all files to make a single group track consisting of Veh and KCl samples from each regions using: Edit 🡪 Groups
3. Identify regions of open chromatin (ROCs)
   1. Import ATAC-seq peaks as annotation
      1. AllRegions\_ATACPeaks.bed
   2. Data 🡪 Define Probes 🡪 Feature Probe Generator
   3. Features to design around 🡪 AllRegions\_ATACPeaks.bed
   4. Split into subfeatures 🡪 No
   5. Remove Exact duplicates 🡪 Checked
   6. Ignore feature strand information 🡪 Unchecked
   7. Make Probes 🡪 Over Feature from -500 to +500 bp
   8. Identifies 191,857 probes 🡪 Name “ROCs”
   9. Add ROCs to annotation track
      1. Right click probe list, select “Convert to annotation track”
4. Filtered probes that were within 1kbp of Refseq curated genes
   1. Filtering 🡪 Filter by features
      1. Features to design around – RefSeq\_Curated.bed
      2. Split into subfeatures – No
      3. Make Probes- Over feature From -1000 to +1000 bp
      4. Select probes which are – overlapping
      5. Distance cutoff (bp) – n/a
      6. Use features on strand – Any
      7. 68,856 probes that are overlapping or within 1kb of RefSeq genes
5. Filtered probes that were within 1kbp of UCSC genes
   1. Filtering 🡪 Filter by features
      1. Features to design around – UCSC\_RefSeq.bed
      2. Split into subfeatures – No
      3. Make Probes- Over feature From -1000 to +1000 bp
      4. Select probes which are – overlapping
      5. Distance cutoff (bp) – n/a
      6. Use features on strand – Any
      7. 71,461 probes that are overlapping or within 1kb of UCSC genes
6. Filtered probes that within 1kbp of Ensemble genes (Rn6\_v95\_gtf\_genessonly\_noBS.txt)
   1. Filtering 🡪 Filter by features
      1. Features to design around – Rn6\_v95\_genessonly\_noBS.txt
      2. Split into subfeatures – No
      3. Make Probes- Over feature From -1000 to +1000 bp
      4. Select probes which are – Overlapping
      5. Distance cutoff (bp) – Not used when looking for overlaps
      6. Use features on strand – Any
      7. 76382 probes that overlap exons
7. Filtered probes that overlap Contiguously transcribed regions
   1. Filtering 🡪 Filter by features
      1. Features to design around – Contig 100,100,0.316(100) over 1kbp
      2. Split into subfeatures – No
      3. Make Probes- Over feature From -1000 to +1000 bp
      4. Select probes which are – Overlapping
      5. Distance cutoff (bp) – Not used when looking for overlaps
      6. Use features on strand – Any
      7. 57366 probes that overlap Contigs
8. Repeat filtering for miRNA, misc\_RNA, rRNA, snoRNA, snRNA, tRNA
   * 1. Features to design around – pick from above
     2. Split into subfeatures – No
     3. Make Probes- Over feature From -0 to +0 byp
     4. Select probes which are – Overlapping
     5. Distance cutoff (bp) – Not used when looking for overlaps
     6. Use features on strand – Any
     7. 839 probes overlapping
9. Filtering 🡪 intersect multiple lists…
   1. Include ROCs, exclude all other lists
   2. Intersection leaves 100767 “Intergenic ROCs”
   3. Convert to annotation track
      1. Right click probe list, select “Convert to annotation track”
10. Data 🡪 Define Probes 🡪 Feature Probe Generator
    1. Features to design around 🡪 Intergenic ROCs
    2. Split into subfeatures 🡪 No
    3. Remove Exact duplicates 🡪 Checked
    4. Ignore feature strand information 🡪 Unchecked
    5. Make Probes 🡪 Over Feature from -0 to +0 bp
11. Probe quantitation
    1. Difference Quantitation
       1. Calculate Forward only as a percentage of All reads
       2. Min count = 0
       3. Ignore duplicates – unchecked
12. Filtering 🡪 Filter on Values
    1. RNA-seq group value must be between 5 and 95%
    2. This excludes probes with unidirectional transcription
    3. Identifies 28492 “Transcriptionally active putative enahncers” – TAPEs
    4. Convert to annotation track
       1. Right click probe list, select “Convert to annotation track”
       2. Name “TAPEs”
13. Define probes 🡪 Feature Probe Generator
    1. Features to design around 🡪 “TAPEs”
    2. Split into subfeatures 🡪 No
    3. Remove exact duplicates 🡪 Checked
    4. Ignore strand feature information 🡪 Unchecked
    5. Make probes 🡪 Over feature -0 to +0 bp
14. Read count quantitation
    1. Count reads on strand 🡪 All reads
    2. Correct for total read count 🡪 Checked
    3. Correct to what? 🡪 per million reads
    4. Count total only in probes? 🡪 Unchecked
    5. Correct for probe length? 🡪 Checked
    6. Log transform count 🡪 Unchecked
    7. Count duplicate reads only once 🡪 Unchecked
    8. Gives CPKM value
15. Filtering 🡪 Probe Values filter
    1. RNA-seq group value must be between 0.05 and 10000
    2. Identifies 1916 highly expressed TAPEs
16. Create report for TAPEs
    1. Reports 🡪 Annotated Probe Report
    2. Annotate with overlapping Rn6\_v95\_gtf\_genessonly\_noBS.txt
    3. Annotation distance cutoff – 1Mbp
    4. Include – unannotated probes
    5. Include – data for currently visible stores