plot scaled BED features

February 19, 2019

Version: 2v1

Title: plot scaled BED features

Description: Calculates the read coverage at a set of BED features with a user defined bp window either side. The resulting data can be plotted as a histogram or box plot.

Author: Benjamin Barwick [cre,aut], Christopher Scharer [cre, aut], Tian Mi [ctb]

Depends R (v3.5.2)

Imports: data.table, chipseq, ShortRead, GenomicAlignments, SummarizedExperiment, DelayedArray, matrixStats, Biobase, BiocParallel, Rsamtools, BSgenome, rtracklayer, GenomicRanges, GenomeInfoDb, Biostrings, XVector, IRanges, S4Vectors, BiocGenerics

Suggests: compiler

Git url: https://github.com/cdschar/plotScaledBEDfeatures

plot.scaled.bedFeatures.R Usage:

Requires the BEDcoverageTools.R package to subset a BAM File for the given Bed coordinates and will calculate the read depth. Using a manifest file the data is summarized by groups and normalized in two manners for plotting.

Files:

BEDcoverageTools.R – R library containing the functions to calculate BED coverage of a BAM file, summarized data, and generate plots.

plot.scaled.bedFeatures.R – Driver script sets the variables and calls different functions. Requires the sample manifest file, bed file, BEDcoverageTools.R functions, and BAM files for each data set. Comments are provided at each step to provide options and explalanations for different parameters.

Sample.manifest.txt – Manifest file that contains a sample column of unique sample IDs and group annotations for each sample. Column headers are fixed and are required for subsetting the data. Other custom columns can be added in addition to the three required columns.

Column description and functions

1. **x** – Numbering column for of all samples.
2. **sample** – Unique sample name that _must_ start with a letter
3. **group** – groups for each sample
4. **Include** - TRUE/FALSE to determine if sample should be plotted.
5. **dir** – path to sample files
6. **bamFile** – bamFile name. '.bai' BAM index must also exist in same directory
7. **frip** – numerical score for fraction of reads in peaks for rppm normalization

BED file – Assumes BED files are generated from HOMER annotatePeaks.pl function. Example of this is provided. This script mostly used motif annotation in ATAC peaks but will work for any other BED feature.