Introducing the diffHic package

Aaron Lun

First edition 2 March 2015 Last revised 3 March 2015

The diffHic package is designed for the detection of differential interactions from Hi-C data. It partitions the genome into bins, and counts the number of read pairs mapping to each pair of bins. Each bin pair is then tested for significant differences between libraries, using the methods in the edgeR package. diffHic can be applied to any Hi-C data set containing multiple biological conditions, though the inclusion of biological replicates in the experimental design is strongly recommended.

On a technical level, the diffHic package provides methods for read alignment and pre-processing into HDF5 files for counting. It provides methods for filtering and normalization prior to the statistical analysis. Functions are also defined for consolidating results from multiple bin sizes, and for visualizing the interaction space. The full user's guide is available as part of the online documentation and can be obtained by typing:

> library(diffHic)

> diffHicUsersGuide()

at the R prompt to open the user's guide in a pdf viewer.