

HiCDataLymphoblast User Manual

Bori Mifsud and Elodie Darbo

Monday 19th August, 2013

1 Introduction

Chromosome conformation capture (3C) is a cross-link based technique to detect spatial proximity of specific genomic distant loci [1], which has sparked a number of large-scale methods to study multiple interactions simultaneously [2]. Hi-C is a genome wide, unbiased version of 3C [3], in which all interactions are sampled by paired-end high-throughput sequencing.

2 Description

2.1 Data

The data were generated by Lieberman-Aiden *et al.* and published in 2009 [3]. In this study, the authors used lymphoblastoid human cells and two different restriction enzymes (HindIII and NcoI) in the HiC experiments. This package contains the reads mapped to hg18 (using *Bowtie* [4]) of the HindIII replicate SRR027956 for chromosome 20 (with options `-seedmms 2`, `-seedlen 40`, `-solexa1.3-quals`). The two ends of the paired reads were mapped separately. This data package is used in the GOTHiC package, that identifies significant interactions in HiC data.

```
> dirPath = system.file("extdata", package="HiCDataLymphoblast")
> fileName1 = list.files(dirPath, full.names=TRUE)[1]
> fileName2 = list.files(dirPath, full.names=TRUE)[2]
> library(ShortRead)
> alignedReads <- readAligned(fileName1, type="Bowtie")
> alignedReads <- readAligned(fileName2, type="Bowtie")
```

References

- [1] Dekker J, Rippe K and Dekker M and Kleckner N: **Capturing chromosome conformation.** *Science* 2002, **295**:1306–11.
- [2] de Wit E and de Laat W: **A decade of 3C technologies: insights into nuclear organization.** *Genes & development* 2012 **26** (1):11–24.

- [3] Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragozy T, Telling A, Amit I, Lajoie BR, Sabo P, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke Andreas, Stamatoyannopoulos J, Mirny LA, Lander ES and Dekker J: **Comprehensive mapping of long-range interactions reveals folding principles of the human genome.** *Science* 2009 **326**:289–93.
- [4] Langmead B, Trapnell C, Pop M and Salzberg SL: **Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.** *Genome Biology* 2009 **10**:R25.