

contiBAIT: Improving Genome Assemblies Using Strand-seq Data

Kieran O'Neill and Mark Hills

March 10, 2015

koneill@bccrc.ca

Contents

1	Licensing	2
2	Introduction	2
3	Correcting misorientations within scaffolds	3
4	Assigning contigs or scaffolds to chromosomes	3
5	Ordering contigs or scaffolds within chromosomes	3

1 Licensing

Under the Artistic License, you are free to use and redistribute this software.

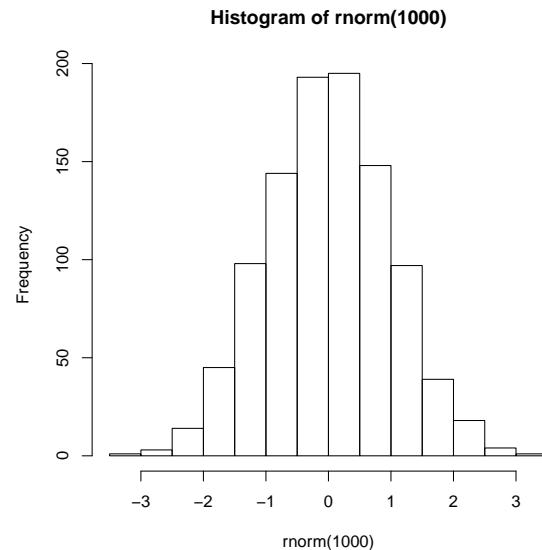
2 Introduction

Strand-seq is a method for determining template strand inheritance in single cells. When strand-seq is collected for many cells from the same organism, spatially close genomic regions show similar patterns of template strand inheritance. ContiBAIT allows users to leverage this property to carry out three tasks to improve early-build genomes. Firstly, misorientations within scaffolds can be found and corrected. Secondly, in genomes made up entirely of contigs or scaffolds not yet assigned to chromosomes, these contigs can be clustered into chromosomes. Thirdly, in genomes wherein scaffolds have been assigned to chromosomes, but not yet placed on those chromosomes, those scaffolds can be placed in order relative to each other.

```
> #An example of some R code within Sweave
> library(contiBAIT)
> print("Hello world")
```

```
[1] "Hello world"
```

```
> plot(hist(rnorm(1000)))
```



3 Correcting misorientations within scaffolds

4 Assigning contigs or scaffolds to chromosomes

5 Ordering contigs or scaffolds within chromosomes

First we load our artificial murine chromosome data. Some words about the meaning of `linkage.group`, `animal.tab`, and `reorientedTable` or whatever it's called.

```
> #data call doesn't seem to work in vignette
> data(contigOrderingExample)
> #load('C:/Users/mike/Dropbox/Rotation/contibait-code/data/contigOrderingExample.rda')
> reorientedTable <- vignetteTestData[[1]]
> animal.tab <- vignetteTestData[[2]][[1]]
> linkage.group <- vignetteTestData[[3]][[1]]
>
```

Next we sort the artificial mouse contigs using `contiBAIT`.

```
> ordering <- orderWithinGroup(linkage.group, reorientedTable, animal.tab)[[1]]
```

The resulting ordering can then be plotted.

```
> plotContigOrder(ordering)
>
```

