

Introduction to BreakPointGenes

Stef van Lieshout, Evert van den Broek

November 18, 2014

1 Running BreakPointGenes

This is a short tutorial on how to use the `BreakPointGenes` package. It covers an example run using the included data set of chromosomes 7–10 of a low grade glioma (LGG) sample. First step is naturally to load the package.

```
> library(BreakPointGenes)
```

1.1 Loading data from a `cghCall` object

Then we need to obtain copy number data. These are available pre-calculated for genome build hg19 and bin sizes 1, 5, 10, 15, 30, 50, 100, 500, and 1000 kbp. They can be downloaded for example with:

```
> bins <- getBinAnnotations(binSize=15)
Downloading bin annotations for genome hg19, bin size 15kbp,
and experiment type SR50 ...
```

After downloading, the bin annotations can be saved locally with `saveRDS`, and in the future be read from the local file with `loadRDS` instead of relying on downloading.

If you are working with another genome build (or another species), see the section on generating the bin annotations.

1.2 Processing bam files

Next step is to load the copy number data from a `cghCall` file or from a text file. This can be done for example with one of the commands below.

```
> breakPoints <- getBP( cghCallObject )
> # all files ending in .bam from the current working directory
>
> # or
>
> breakPoints <- getBP( "file.txt", type='txt' )
> # file 'file.txt' from the current working directory
>
```

This will return an object of class `CNBPgene`.

For the purpose of this tutorial, we load an example data set of chromosomes 7–10 of low grade glioma sample LGG150.

```
> breakPoints <- "example_code"  
> breakPoints
```

```
[1] "example_code"
```

Finally, for other downstream analyses, such as running `CGHregions`, it might be useful to convert to a `cghCall` object.

```
> cgh <- matrix()
```

2 Session Information

The version number of R and packages loaded for generating the vignette were:

R version 3.1.1 (2014-07-10)

Platform: x86_64-pc-linux-gnu (64-bit)

locale:

```
[1] LC_CTYPE=en_US.UTF-8
[2] LC_NUMERIC=C
[3] LC_TIME=en_US.UTF-8
[4] LC_COLLATE=en_US.UTF-8
[5] LC_MONETARY=en_US.UTF-8
[6] LC_MESSAGES=en_US.UTF-8
[7] LC_PAPER=en_US.UTF-8
[8] LC_NAME=C
[9] LC_ADDRESS=C
[10] LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_US.UTF-8
[12] LC_IDENTIFICATION=C
```

attached base packages:

```
[1] stats      graphics  grDevices
[4] utils      datasets  methods
[7] base
```

other attached packages:

```
[1] BreakPointGenes_0.0.1
```

loaded via a namespace (and not attached):

```
[1] Biobase_2.24.0
[2] BiocGenerics_0.10.0
[3] Biostrings_2.32.1
[4] bitops_1.0-6
[5] CGHbase_1.24.0
[6] CGHcall_2.26.0
[7] devtools_1.6
[8] digest_0.6.4
[9] DNACopy_1.38.1
[10] GenomeInfoDb_1.0.2
[11] GenomicRanges_1.16.4
[12] impute_1.38.1
[13] IRanges_1.22.10
[14] limma_3.20.9
[15] marray_1.42.0
[16] matrixStats_0.10.0
```

```
[17] parallel_3.1.1
[18] Rcpp_0.11.3
[19] R.methodsS3_1.6.1
[20] R.oo_1.18.0
[21] roxygen2_4.0.2
[22] Rsamtools_1.16.1
[23] rstudio_0.98.1062
[24] rstudioapi_0.1
[25] R.utils_1.33.0
[26] stats4_3.1.1
[27] stringr_0.6.2
[28] tools_3.1.1
[29] XVector_0.4.0
[30] zlibbioc_1.10.0
```