Integration with the crlmm package for copy number inference

Robert Schgarpf

October 9, 2012

We load a portion of chromosome 8 from 2 HapMap samples that were processed using the *crlmm* package.

```
> library(oligoClasses)
> library(VanillaICE)
> library2(crlmm)
> library2(SNPchip)
> library2(IRanges)
> data(cnSetExample, package="crlmm")
```

The data `cnSetExample` is an object of class `CNSet`. We coerce the `CNSet` object to a list class that contains information on copy number (log R ratios), genotypes, genotype probabilities, and B allele frequencies.

```
> oligoList <- constructOligoSetListFrom(cnSetExample)
```

The `[[` method can be used to extract a `oligoSnpSet` for the first element in the list:

```
> oligoSet <- oligoList[[1]]
```

Next, we fit a 6-state hidden markov model from estimates of the B allele frequency and log R ratios.

```
> res <- hmm(oligoSet, p.hom=0.1, nupdates=5, TAUP=1e8)
```

The `TAUP` parameter scales the transition probability matrix. Larger values of `TAUP` makes it more expensive to transition from the normal copy number state to states with altered copy number. In the following code chunk, we use a lattice multi-panel display to plot each of the altered stated. We frame each alteration by plotting a genomic interval of 200kb on each side (using the `frame=200e3` argument):

```
> rd <- res[chromosome(res) == "chr8", ]
> rd <- res[!state(res)%in%c(3,4), ]
> if(require(SNPchip)){
+         fig <- xyplotLrrBaf(rd, oligoSet,
+                             frame=200e3,
+                             panel=xypanelBaf,
+                             scales=list(x="free"),
+                             par.strip.text=list(cex=0.9),
+                             cex=0.4,
+                             state.col="black",
+                             state.cex=0.8,
+                             pch=21)
+ }

> print(fig)
```

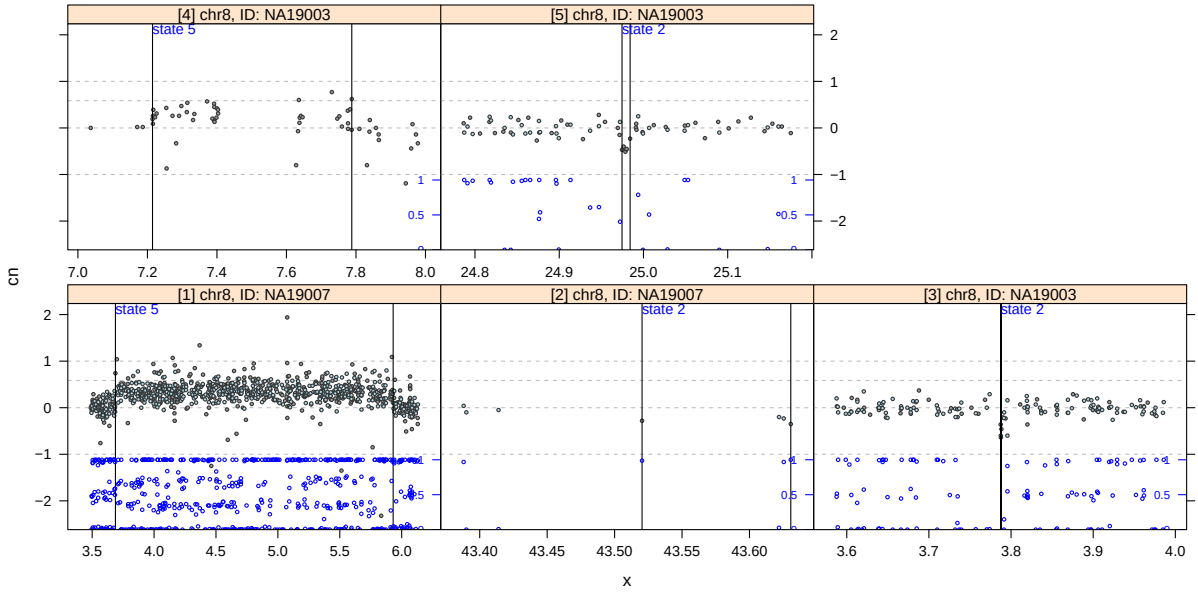

Figure 1: Plot of log R ratios (grey) and B allele frequencies (blue). The B allele frequencies have a range of 0-1 and were rescaled for ease of viewing alongside the log R ratios. Each panel displays one region with a copy number alteration predicted from the 6-state HMM with padding on either side given by the `frame` argument.

1 Session Information

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

- R version 2.15.1 (2012-06-22), `x86_64-unknown-linux-gnu`
- Locale: `LC_CTYPE=en_US.UTF-8`, `LC_NUMERIC=C`, `LC_TIME=en_US.UTF-8`, `LC_COLLATE=C`, `LC_MONETARY=en_US.UTF-8`, `LC_MESSAGES=en_US.UTF-8`, `LC_PAPER=C`, `LC_NAME=C`, `LC_ADDRESS=C`, `LC_TELEPHONE=C`, `LC_MEASUREMENT=en_US.UTF-8`, `LC_IDENTIFICATION=C`
- Base packages: base, datasets, grDevices, graphics, methods, stats, utils
- Other packages: Biobase 2.18.0, BiocGenerics 0.4.0, DBI 0.2-5, IRanges 1.16.2, RColorBrewer 1.0-5, RSQLite 0.11.2, SNPchip 2.4.0, VanillaICE 1.20.3, crlmm 1.16.0, oligo 1.22.0, oligoClasses 1.20.0, pd.mapping50k.hind240 1.8.0, pd.mapping50k.xba240 1.8.0
- Loaded via a namespace (and not attached): AnnotationDbi 1.20.0, BiocInstaller 1.8.2, Biostrings 2.26.1, GenomicRanges 1.10.1, XML 3.95-0.1, affxparser 1.30.0, affyio 1.26.0, annotate 1.36.0, bit 1.1-8, codetools 0.2-8, compiler 2.15.1, ellipse 0.3-7, ff 2.2-7, foreach 1.4.0, genefilter 1.40.0, grid 2.15.1, iterators 1.0.6, lattice 0.20-10, msm 1.1.3, mvtnorm 0.9-9992, parallel 2.15.1, preprocessCore 1.20.0, splines 2.15.1, stats4 2.15.1, survival 2.36-14, tools 2.15.1, xtable 1.7-0, zlibbioc 1.4.0