Integration with the crlmm package for copy number inference

Robert Schgarpf

June 28, 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, prOutlierBAF=list(initial=1e-4, max=1e-2, maxROH=1e-4))
```

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)
```

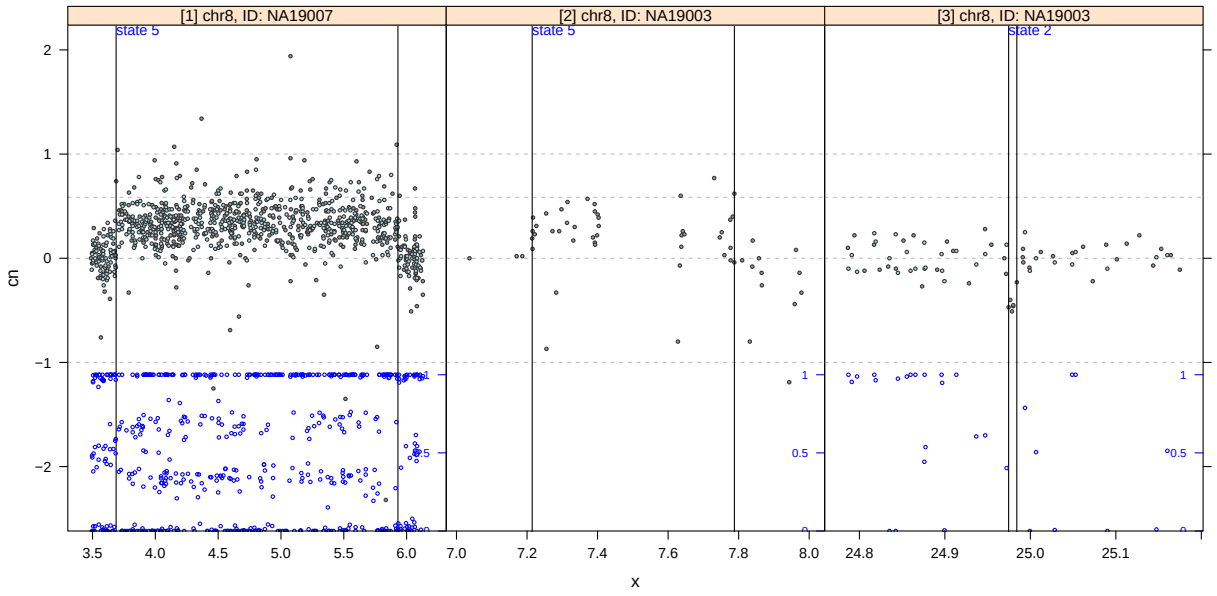

Figure 1: Plot of log R ratios (grey) and B allele frequencies (blue). The B allele frequencies have a range of 0-1 and were rescaled for ease of viewing alongside the log R ratios. Each panel displays one region with a copy number alteration predicted from the 6-state HMM with padding on either side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