A new approach to SNP location metadata

VJ Carey

April 14, 2011

1 Introduction

Versions of GG tools prior to 2.3.x have a complicated approach to SNP location metadata, involving a specially constructed SQLite database. In versions of GGB up to 3.11, some SNP location metadata support was provided. In GGB as 3.11 and beyond, users are responsible for managing their own location metadata for SNPs and genes. This can be accomplished using the SNP locs. Hsapiens.* package and other annotation resources. The GG tools cisProxScores function shows some of the possibilities.

This vignette is retained as legacy for those who may want to retrieve earlier versions of R/Bioconductor to employ formerly supported facilities.

One class and two methods are supported.

```
> getClass("snpLocs")
```

- > data(hsSnpLocs)
- > hsSnpLocs

The chromosome-specific locations are generated reasonably efficiently:

```
> snpLocs.Hs(chrnum(20), rsid("rs6060535"))
```

2 Construction of serialized reference container

First, unify the name and location information from the SNPlocs package.

```
> humanSNPlocs = list()
> library(SNPlocs.Hsapiens.dbSNP.20071016)
> if (file.exists("humanSNPlocs.rda")) load("humanSNPlocs.rda") else {
    for (i in c(as.character(1:22), "X", "Y")) {
        curc = getSNPlocs(paste("chr", i, sep = ""))
        rsid.int = as.integer(curc[, 1])
        loc.int = as.integer(curc[, 3])
```

```
+ humanSNPlocs[[i]] = rbind(rsid = rsid.int, loc = loc.int)
+ }
+ }
```

Now get offsets for computing the chromosome-wide location values.

```
> require(org.Hs.eg.db)
> chrl = org.Hs.egCHRLENGTHS
> offs = c(0, cumsum(as.double(chrl[1:22])))
```

Now we create the environment-based container instance:

```
> el = new.env()
> getv = function(x) installed.packages()[x, "Version"]
> for (i in names(humanSNPlocs)) assign(i, humanSNPlocs[[i]], el)
> hsSnpLocs = new("snpLocs", locEnv = el, offsets = offs, organism = "Hs",
+ versions = c(org.Hs.eg.db = getv("org.Hs.eg.db"), SNPlocs.Hsapiens.dbSNP.200710
```

This object will be saved in GGBase.