

ceu1kg: resources for exploring the 1000 genomes data on individuals of central European ancestry in Bioconductor

VJ Carey

July 4, 2010

1 Introduction

Using results of next generation sequencing experiments, a consortium of geneticists produced calls for SNP at approximately 8 million loci of the genomes of individuals of central European ancestry.

Genotype calls are held in a list of `snp.matrix` instances:

```
> library(ceu1kg)
> if (!exists("ceu1KG.sml")) data(ceu1KG.sml)
> names(ceu1KG.sml)

[1] "chr1" "chr2" "chr3" "chr4" "chr5" "chr6" "chr7" "chr8" "chr9"
[10] "chr10" "chr11" "chr12" "chr13" "chr14" "chr15" "chr16" "chr17" "chr18"
[19] "chr19" "chr20" "chr21" "chr22"

> ceu1KG.sml[[1]]
```

```
A snp.matrix with 60 rows and 605756 columns
Row names: NA06985 ... NA12874
Col names: chr1:533 ... chr1:247196267
```

Metadata about the loci are provided in `GRanges` instances:

```
> if (!exists("ceu1kgMeta_20")) data(ceu1kgMeta_20)
> ceu1kgMeta_20[1:3]
```

`GRanges` with 3 ranges and 5 elementMetadata values

seqnames	ranges	strand	ref	alt	depth
<Rle>	<IRanges>	<Rle>	<character>	<character>	<character>

rs6078030	chr20 [9098, 9098]	*	C	T	192
rs4814683	chr20 [9795, 9795]	*	G	T	332
rs34147676	chr20 [10731, 10731]	*	C	A	228
	ancest alleleCnt				
	<character> <character>				
rs6078030	.				24
rs4814683	.				40
rs34147676	.				14
seqlengths					
chr20					
NA					