The **DMRcatedata** package user’s guide

Peters TJ, Buckley MJ, Statham AL, Pidsley R, Zotenko E, Clark SJ, Molloy PL

February 13, 2024

**Package Contents**

**DMRcatedata** accompanies the **DMRcate** package, providing data for examples, probe filtering and transcript annotation.

```r
library(DMRcatedata)
data(crosshyb)

#Proximal SNPs to EPICv1 and 450K
data(snpsall)

#Proximal SNPs to EPICv2
data(epicv2snps)

data(hg19.grt)
data(hg19.generanges)

#Betats for EPICv2
data(ALLbetas)
```

Ten objects are contained in **DMRcatedata**. **crosshyb** is a factor listing EPICv1 and 450K probe IDs potentially confounded by cross-hybridisation to other parts of the genome[1][2]. It is used internally by `rmSNPandCH()`.

**snpsall** is a data.frame containing probes from 450K and EPICv1 that are potentially confounded by a SNP or indel variant[1]. It lists the ID, distance (in nucleotides) to the CpG in question, and minor allele frequency for each associated variant. **epicv2snps** contains the same but for the EPICv2 array.

**XY.probes** is a vector of EPICv1 and 450K Illumina probes whose targets are on human sex chromosomes.

Objects named `.g**rt**|generanges` are annotation objects that are needed by `extractRanges()` and `DMR.plot()` respectively. `hg38` and `mm10` objects have been parsed from Release 96 of Ensembl, and `hg19` from Release 75. These are accessed within the environment of the aforementioned functions.
**ALLbetas** is a matrix of EPICv2 beta values from Noguera-Castells et al. (2023)[3] consisting of five B cell acute lymphoblastic leukaemia (BALL) and five T cell acute lymphoblastic leukaemia (TALL) samples for DMR calling.

**Sources**


**References**

