BSgenome.Cfamiliaris.UCSC.canFam2

April 6, 2012

Cfamiliaris *Canis lupus familiaris full genome (UCSC version canFam2)*

Description

Canis lupus familiaris full genome as provided by UCSC (canFam2, May 2005) and stored in
Biostrings objects.

Note

This BSgenome data package was made from the following source data files:

```
sequences: chromFa.tar.gz
from http://hgdownload.cse.ucsc.edu/goldenPath/canFam2/bigZips/
AGAPS masks: all the chr*_gap.txt.gz files from ftp://hgdownload.cse.ucsc.edu/goldenPath/canFam2
RM masks: http://hgdownload.cse.ucsc.edu/goldenPath/canFam2/bigZips/chromOut.tar.gz
TRF masks: http://hgdownload.cse.ucsc.edu/goldenPath/canFam2/bigZips/chromTrf.tar.gz
```

See ?BSgenomeForge and the BSgenomeForge vignette (vignette("BSgenomeForge")) in the
BSgenome software package for how to make a BSgenome data package.

Author(s)

H. Pages

See Also

BSgenome-class, DNAString-class, available.genomes, BSgenomeForge

Examples

```
Cfamiliaris
seqlengths(Cfamiliaris)
Cfamiliaris$chr1  # same as Cfamiliaris[["chr1"]]

if ("AGAPS" %in% masknames(Cfamiliaris)) {

  ## Check that the assembly gaps contain only Ns:
  checkOnlyNsInGaps <- function(seq)
```

```
  {
    ## Replace all masks by the inverted AGAPS mask
    masks(seq) <- gaps(masks(seq)["AGAPS"])
    unique_letters <- uniqueLetters(seq)
    if (any(unique_letters != "N"))
        stop("assembly gaps contain more than just Ns")
  }

  ## A message will be printed each time a sequence is removed
  ## from the cache:
  options(verbose=TRUE)

  for (seqname in seqnames(Cfamiliaris)) {
    cat("Checking sequence", seqname, "... ")
    seq <- Cfamiliaris[[seqname]]
    checkOnlyNsInGaps(seq)
    cat("OK\n")
  }
}

## See the GenomeSearching vignette in the BSgenome software
## package for some examples of genome-wide motif searching using
## Biostrings and the BSgenome data packages:
if (interactive())
    vignette("GenomeSearching", package="BSgenome")
```

Index