# BSgenome.Drerio.UCSC.danRer5

October 16, 2009

---

Drerio                    *Danio rerio full genome (UCSC version danRer5)*

---

**Description**

Danio rerio full genome as provided by UCSC (danRer5, Jul. 2007) and stored in Biostrings objects.

**Note**

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

```
sequences: danRer5.fa.gz, upstream1000.fa.gz, upstream2000.fa.gz, upstream5000.f
from http://hgdownload.cse.ucsc.edu/goldenPath/danRer5/bigZips/
AGAPS masks: http://hgdownload.cse.ucsc.edu/goldenPath/danRer5/database/gap.txt.
RM masks: http://hgdownload.cse.ucsc.edu/goldenPath/danRer5/bigZips/danRer5.fa.o
TRF masks: http://hgdownload.cse.ucsc.edu/goldenPath/danRer5/bigZips/danRer5.trf
```

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**

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

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

  ## Check that the assembly gaps contain only Ns:
  checkOnlyNsInGaps <- function(seq)
  {
    ## Replace all masks by the inverted AGAPS mask
```

1

```
    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(Drerio)) {
    cat("Checking sequence", seqname, "... ")
    seq <- Drerio[[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