BSgenome.Athaliana.TAIR.01222004

October 17, 2011

Athaliana	*Arabidopsis thaliana full genome (TAIR version from January 22, 2004)*

Description

Arabidopsis thaliana full genome as provided by TAIR (snapshot from January 22, 2004) and stored in Biostrings objects.

Note

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

```
ATH1_chr1.1con.01222004
ATH1_chr2.1con.01222004
ATH1_chr3.1con.01222004
ATH1_chr4.1con.01222004
ATH1_chr5.1con.01222004
mitochondrial_genomic_sequence
ATH1_chloroplast.1con.01072002
from ftp://ftp.arabidopsis.org/home/tair/Sequences/whole_chromosomes/OLD/
```

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

1

Examples

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

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

  ## 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(Athaliana)) {
    cat("Checking sequence", seqname, "... ")
    seq <- Athaliana[[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