Package 'rrbgen'

July 13, 2018

Type Package

Title Robbie's lightweight limited functionality R bgen read/write library

Version 0.0.4

Date 2018-07-13

Author Robert William Davies

Maintainer Robert William Davies <robertwilliamdavies@gmail.com>

Description A lightweight R bgen library with limited functionality

Imports Rcpp, RcppArmadillo, parallel

LinkingTo Rcpp, RcppArmadillo

License GPL | file LICENSE

SystemRequirements C++11

NeedsCompilation yes

Suggests testthat

RoxygenNote 6.0.1

R topics documented:

`rrbgen_load` *Load entire bgen file*

Usage

```
rrbgen_load(bgen_file, gp_names_col = "varid", vars_to_get = NULL,
  samples_to_get = NULL)
```

Arguments

bgen_file	Path to bgen file to load
gp_names_col	Which column to use to label the variants in the first dimension of the genotype probabilities. Note that bgen has two variant ID columns, a variant identifier (here labelled varid) and rsid (here labelled rsid)
vars_to_get	Only load these variants from file (from varid column). Note that this requires a complete pass over the data, although this should be efficient, as nothing is decompressed
samples_to_get	Only load these samples from file. Note that this requires a complete pass over the data, although this should be efficient, as nothing is decompressed

Value

gp, sample_names and var_info

Author(s)

Robert Davies

---

rrbgen_load_samples *Load samples names only from a bgen file*

Usage

```
rrbgen_load_samples(bgen_file)
```

Arguments

bgen_file	Path to bgen file to load

Value

sample_names

Author(s)

Robert Davies

rrbgen_load_variant_info *Load variant information only from a bgen file*

Usage

```
rrbgen_load_variant_info(bgen_file)
```

Arguments

bgen_file	Path to bgen file to load

Value

var_info

Author(s)

Robert Davies

rrbgen_write *Write variant information to a bgen file*

Usage

```
rrbgen_write(bgen_file, sample_names = NULL, var_info = NULL, gp = NULL,
  list_of_gp_raw_t = NULL, free = NULL, Layout = 2,
  CompressedSNPBlocks = 1, B_bit_prob = 8, close_bgen_file = TRUE,
  add_to_bgen_connection = FALSE, bgen_file_connection = NULL,
  previous_offset = NULL, header_M = NULL, nCores = 1, verbose = FALSE)
```

Arguments

bgen_file	Path to bgen file to be written to
sample_names	character vector of sample names
var_info	var_info (6 first columns of gen file, or chr, varid, rsid, position (1-based), ref, alt)
gp	genotype probabilities as 3-dimensional array, where dim1 = variants, dim2 = samples, dim3 = hom ref, het, hom alt
list_of_gp_raw_t	list of raw data. bespoke format. to understand, investigate the tests
free	free text for header. Should probably remain NULL. Not read if written in rrbgen_read
Layout	Only supported 2 (see bgen spec)
CompressedSNPBlocks	Only supported 1. Whether to (1) or not to (0) compress genotype probabilities

B_bit_prob	How many bits to use to encode genotype probabilities (only supported 8, 16, 24, 32)
close_bgen_file	Whether to close bgen file once this function has completed
add_to_bgen_connection	Whether to add to previous bgen file connection
bgen_file_connection	Previous bgen file connection for file to add to
previous_offset	Previous length of open bgen file +4
header_M	How many SNPs to specify in the header
nCores	How many cores to use for compression
verbose	Whether to print logging information when writing

Author(s)

Robert Davies

Index