

Package ‘snpAIMeR’

February 20, 2024

Type Package

Title Assess the Diagnostic Power of Genomic Marker Combinations

Version 2.1.1

Description Population genetics package for designing diagnostic panels. Candidate markers, marker combinations, and different panel sizes are assessed for how well they can predict the source population of known samples. Requires a genotype file of candidate markers in STRUCTURE format. Methods for population cross-validation are described in Jombart (2008) <[doi:10.1093/bioinformatics/btn129](https://doi.org/10.1093/bioinformatics/btn129)>.

License MIT + file LICENSE

Contact Kim Vertacnik <kim.vertacnik@mailbox.org> or Julian Dupuis <julian.dupuis@uky.edu>

URL <https://github.com/OksanaVe/snpAIMeR>

Depends R (>= 2.10)

Imports adegenet, doParallel, dplyr, forcats, foreach, ggplot2, graphics, magrittr, parallel, readr, tidyr, utils, withr, yaml

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

Encoding UTF-8

RoxygenNote 7.3.1

NeedsCompilation no

Author Kim Vertacnik [cre, aut] (<<https://orcid.org/0009-0006-7097-5091>>),
Oksana Vernygora [aut] (<<https://orcid.org/0000-0001-6444-3888>>),
Julian Dupuis [aut] (<<https://orcid.org/0000-0002-6989-9179>>)

Maintainer Kim Vertacnik <kim.vertacnik@mailbox.org>

Repository CRAN

Date/Publication 2024-02-20 20:20:02 UTC

R topics documented:

snpAIMeR 2

Description

Population genetics package for optimizing diagnostic panels. User-selected candidate markers are assessed individually and in combination for how well they can predict the source population of known samples. Requires a genotype file in STRUCTURE format.

Usage

```
snpAIMeR(run_mode, config_file = NULL, verbose = TRUE)
```

Arguments

run_mode	Modes are "interactive", "non-interactive", or "example"; mode must be in quotes.
config_file	Yaml file required for "non-interactive" mode; filename/path must be in quotes.
verbose	Default is TRUE.

Details

Yaml file format for "non-interactive" mode (do not include bullet points):

- min_range: <minimum panel size>
- max_range: <maximum panel size; we recommend no more than 15 markers>
- assignment_rate_threshold: <value from 0 to 1>
- cross_validation_replicates: <we recommend 100 minimum>
- working_directory: <path name in quotes>
- structure_file: <path name in quotes>
- number_of_individuals: <same as adegenet's "n.ind">
- number_of_loci: <same as adegenet's "n.loc">
- one_data_row_per_individual: <TRUE or FALSE>
- column_sample_IDs: <column number>
- column_population_assignments: <column number>
- column_other_info: <column number>
- row_markernames: <row number>
- no_genotype_character: <default is "-9">
- optional_population_info: <optional>
- genotype_character_separator: <optional>

Minimizing run time: Because of the number of possible combinations, we recommend testing no more than 15 markers. For example, testing 15 markers in panel sizes of 1 to 15 (32,767 total combinations) with 1,000 cross-validation replicates on a system with 48 processor cores took about 5 hours and 20 GB RAM. Reducing the number of cross-validation replicates will reduce run time, however, we recommend no less than 100 replicates.

Value

Cross-validation assignment rates for individual markers, marker combinations, and panel sizes. Outputs three .csv and two .pdf files to a user-specified directory.

See Also

<https://github.com/OksanaVe/snpAIMeR>

Examples

```
if (requireNamespace("adegenet", quietly = TRUE)) {  
  data(nancycats, package = "adegenet")  
  snpAIMeR("example", verbose = TRUE)  
}
```

Index

snpAIMeR, [2](#)