Package ‘PopGenHelpR’

January 20, 2025

Title Streamline Population Genomic and Genetic Analyses

Version 1.3.2

Description Estimate commonly used population genomic statistics and generate publication quality figures. 'PopGenHelpR' uses vcf, 'geno' (012), and csv files to generate output.

URL https://kfarleigh.github.io/PopGenHelpR/

BugReports https://github.com/kfarleigh/PopGenHelpR/issues

License GPL (>= 3)

Encoding UTF-8

RoxygenNote 7.3.1

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

Config/testthat/edition 3

Imports dplyr, ggplot2, magrittr, methods, reshape2, rlang, scatterpie, stats, spData, spdep, utils, vcfR

Depends R (>= 2.10)

LazyData true

VignetteBuilder knitr

NeedsCompilation no

Author Keaka Farleigh [aut, cph, cre] (<https://orcid.org/0000-0002-9195-121X>), Mason Murphy [aut, cph, ctb] (<https://orcid.org/0000-0002-5317-1386>), Christopher Blair [aut, cph, ctb] (<https://orcid.org/0000-0002-4782-6788>), Tereza Jezkova [aut, cph, ctb] (<https://orcid.org/0000-0003-4114-4564>)

Maintainer Keaka Farleigh <keakafarleigh@gmail.com>

Repository CRAN

Date/Publication 2024-08-02 10:20:04 UTC

Contents

Ancestry_barchart . . . 2
Differentiation . . . 3
Fst_dat . . . 5
Heterozygosity . . . 6
Het_dat . . . 8
HornedLizard_Pop . . . 9
HornedLizard_VCF . . . 10
Network_map . . . 10
Pairwise_heatmap . . . 12
PCA . . . 13
Piechart_map . . . 14
Plot_coordinates . . . 15
Point_map . . . 16
Private.alleles . . . 18
Q_dat . . . 19

Index **20**

`Ancestry_barchart` *Plot an ancestry matrix for individuals and(or) populations.*

Description

Plot an ancestry matrix for individuals and(or) populations.

Usage

```
Ancestry_barchart(
  anc.mat,
  pops,
  K,
  plot.type = "all",
  col,
  ind.order = NULL,
  pop.order = NULL
)
```

Arguments

`anc.mat` Data frame or character string that supplies the input data. If it is a character string, the file should be a csv. The first column should be the names of each sample/population, followed by the estimated contribution of each cluster to that individual/pop.

pops	Data frame or character string that supplies the input data. If it is a character string, the file should be a csv. The columns should be named Sample, containing the sample IDs; Population indicating the population assignment of the individual, population and sample names must be the same type (i.e., both numeric or both characters); Long, indicating the longitude of the sample; Lat, indicating the latitude of the sample.
K	Numeric.The number of genetic clusters in your data set, please contact the package authors if you need help doing this.
plot.type	Character string. Options are all, individual, and population. All is default and recommended, this will plot a barchart for both the individuals and populations.
col	Character vector indicating the colors you wish to use for plotting.
ind.order	Character vector indicating the order to plot the individuals in the individual ancestry bar chart.
pop.order	Chracter vector indicating the order to plot the populations in the population ancesyry bar chart.

Value

A list containing your plots and the data frames used to generate the plots.

Author(s)

Keaka Farleigh

Examples

```
data(Q_dat)
Qmat <- Q_dat[[1]]
rownames(Qmat) <- Qmat[,1]
Loc <- Q_dat[[2]]
Test_all <- Ancestry_barchart(anc.mat = Qmat, pops = Loc, K = 5,
plot.type = 'all',col = c('#d73027', '#fc8d59', '#e0f3f8', '#91bfdb', '#4575b4'))
```

Differentiation — *A function to estimate three measures of genetic differentiation using geno files, vcf files, or vcfR objects. Data is assumed to be bi-allelic.*

Description

A function to estimate three measures of genetic differentiation using geno files, vcf files, or vcfR objects. Data is assumed to be bi-allelic.

Usage

```
Differentiation(
  data,
  pops,
  statistic = "all",
  missing_value = NA,
  write = FALSE,
  prefix = NULL,
  population_col = NULL,
  individual_col = NULL
)
```

Arguments

`data`	Character. String indicating the name of the vcf file, geno file or vcfR object to be used in the analysis.
`pops`	Character. String indicating the name of the population assignment file or dataframe containing the population assignment information for each individual in the data. This file must be in the same order as the vcf file and include columns specifying the individual and the population that individual belongs to. The first column should contain individual names and the second column should indicate the population assignment of each individual. Alternatively, you can indicate the column containing the individual and population information using the individual_col and population_col arguments.
`statistic`	Character. String or vector indicating the statistic to calculate. Options are any of: all; all of the statistics; Fst, Weir and Cockerham (1984) Fst; NeisD, Nei's D statistic; JostsD, Jost's D.
`missing_value`	Character. String indicating missing data in the input data. It is assumed to be NA, but that may not be true (is likely not) in the case of geno files.
`write`	Boolean. Whether or not to write the output to files in the current working directory. There will be one or two files for each statistic. Files will be named based on their statistic such as Fst_perpop.csv.
`prefix`	Character. Optional argument. String that will be appended to file output. Please provide a prefix if write is set to TRUE.
`population_col`	Numeric. Optional argument (a number) indicating the column that contains the population assignment information.
`individual_col`	Numeric. Optional argument (a number) indicating the column that contains the individuals (i.e., sample name) in the data.

Value

A list containing the estimated heterozygosity statistics. The per pop values are calculated by taking the average of the per locus estimates.

Author(s)

Keaka Farleigh

References

Fst:

Pembleton, L. W., Cogan, N. O., & Forster, J. W. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. Molecular ecology resources, 13(5), 946-952.doi:10.1111/17550998.12129

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. evolution, 1358-1370.

Nei's D:

Nei, M. (1972). Genetic distance between populations. The American Naturalist, 106(949), 283-292.doi:10.1086/282771

doi:10.1111/17550998.12129 Pembleton, L. W., Cogan, N. O., & Forster, J. W. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. Molecular ecology resources, 13(5), 946-952.

Jost's D:

Jost L (2008). GST and its relatives do not measure differentiation. Molecular Ecology, 17, 4015–4026.doi:10.1111/j.1365294X.2008.03887.x

Examples

```
data("HornedLizard_Pop")
data("HornedLizard_VCF")
Test <- Differentiation(data = HornedLizard_VCF, pops = HornedLizard_Pop, write = FALSE)
```

`Fst_dat` *A genetic differentiation matrix and locality information for each population. This data was generated by subsetting data of Farleigh et al., 2021.*

Description

A symmetric matrix with estimated genetic differentiation (Fst) between 3 populations.

Usage

```
data(Fst_dat)
```

Format

A list with two elements:

Fst_dat Data frame with three rows and three columns

Loc_dat Data frame containing the locality information for each population

...

Source

Farleigh, K., Vladimirova, S. A., Blair, C., Bracken, J. T., Koochekian, N., Schield, D. R., ... & Jezkova, T. (2021). The effects of climate and demographic history in shaping genomic variation across populations of the Desert Horned Lizard (Phrynosoma platyrhinos). Molecular Ecology, 30(18), 4481-4496.

Examples

```
data(Fst_dat)
Fst <- Fst_dat[[1]]
Loc <- Fst_dat[[2]]

 Test <- Network_map(dat = Fst, pops = Loc,
neighbors = 2,col = c('#4575b4', '#91bfdb', '#e0f3f8','#fd8d3c','#fc4e2a'),
statistic = "Fst", Lat_buffer = 1, Long_buffer = 1)

Fstat_plot <- Pairwise_heatmap(dat = Fst, statistic = 'FST')
```

Heterozygosity — *A function to estimate seven measures of heterozygosity using geno files, vcf files, or vcfR objects. Data is assumed to be bi-allelic.*

Description

A function to estimate seven measures of heterozygosity using geno files, vcf files, or vcfR objects. Data is assumed to be bi-allelic.

Usage

```
Heterozygosity(
  data,
  pops,
  statistic = "all",
  missing_value = NA,
  write = FALSE,
  prefix = NULL,
  population_col = NULL,
  individual_col = NULL
)
```

Arguments

`data`	Character. String indicating the name of the vcf file, geno file or vcfR object to be used in the analysis.

`pops` Character. String indicating the name of the population assignment file or dataframe containing the population assignment information for each individual in the data. This file must be in the same order as the vcf file and include columns specifying the individual and the population that individual belongs to. The first column should contain individual names and the second column should indicate the population assignment of each individual. Alternatively, you can indicate the column containing the individual and population information using the individual_col and population_col arguments.

`statistic` Character. String or vector indicating the statistic to calculate. Options are any of: all; all of the statistics; Ho, observed heterozygosity; He, expected heterozygosity; PHt, proportion of heterozygous loci; Hs_exp, heterozygosity standardized by the average expected heterozygosity; Hs_obs, heterozygosity standardized by the average observed heterozygosity; IR, internal relatedness; HL, homozygosity by locus.

`missing_value` Character. String indicating missing data in the input data. It is assumed to be NA, but that may not be true (is likely not) in the case of geno files.

`write` Boolean. Whether or not to write the output to files in the current working directory. There will be one or two files for each statistic. Files will be named based on their statistic such as Ho_perpop.csv or Ho_perloc.csv.

`prefix` Character. Optional argument. String that will be appended to file output. Please provide a prefix if write is set to TRUE.

`population_col` Numeric. Optional argument (a number) indicating the column that contains the population assignment information.

`individual_col` Numeric. Optional argument (a number) indicating the column that contains the individuals (i.e., sample name) in the data.

Value

A list containing the estimated heterozygosity statistics. The per pop values are calculated by taking the average of the per locus estimates.

Author(s)

Keaka Farleigh

References

Expected (He) and observed heterozygosity (Ho):

Nei, M. (1987) Molecular Evolutionary Genetics. Columbia University Press

Homozygosity by locus (HL) and internal relatedness (IR):

Alho, J. S., Välimäki, K., & Merilä, J. (2010). Rhh: an R extension for estimating multilocus heterozygosity and heterozygosity–heterozygosity correlation. Molecular ecology resources, 10(4), 720-722.

Amos, W., Worthington Wilmer, J., Fullard, K., Burg, T. M., Croxall, J. P., Bloch, D., & Coulson, T. (2001). The influence of parental relatedness on reproductive success. Proceedings of the Royal Society of London. Series B: Biological Sciences, 268(1480), 2021-2027.doi:10.1098/rspb.2001.1751

Aparicio, J. M., Ortego, J., & Cordero, P. J. (2006). What should we weigh to estimate heterozygosity, alleles or loci?. Molecular Ecology, 15(14), 4659-4665.

Heterozygosity standardized by expected (Hs_exp) and observed heterozygosity (Hs_obs):

Coltman, D. W., Pilkington, J. G., Smith, J. A., & Pemberton, J. M. (1999). Parasite-mediated selection against Inbred Soay sheep in a free-living island populaton. Evolution, 53(4), 1259-1267.

Examples

```
data("HornedLizard_Pop")
data("HornedLizard_VCF")
Test <- Heterozygosity(data = HornedLizard_VCF, pops = HornedLizard_Pop, write = FALSE)
```

`Het_dat` *A data frame of hypothetical heterozygosity data produced by Heterozygosity.*

Description

Data frame containing 5 columns and 3 rows

Usage

```
data(Het_dat)
```

Format

A data frame with 5 columns and 3 rows:

Heterozygosity Estimated heterozygosity

Pop Population assignment

Standard.Deviation standard deviation

Longitude Longitude

Latitude Latitude

...

Source

Coordinates and population names taken from Farleigh, K., Vladimirova, S. A., Blair, C., Bracken, J. T., Koochekian, N., Schield, D. R., ... & Jezkova, T. (2021). The effects of climate and demographic history in shaping genomic variation across populations of the Desert Horned Lizard (Phrynosoma platyrhinos). Molecular Ecology, 30(18), 4481-4496.

Examples

```
data(Het_dat)
Test <- Point_map(Het_dat, statistic = "Heterozygosity")
```

`HornedLizard_Pop` *A population assignment data frame to be used in* `Heterozygosity` *and* `Differentiation`.

Description

Data frame containing 4 columns and 72 rows

Usage

```
data(HornedLizard_Pop)
```

Format

A data frame with 4 columns and 72 rows:

Sample Sample Name

Population Population assignment according to sNMF results (see citation)

Longitude Longitude

Latitude Latitude

...

Source

Coordinates and population names taken from Farleigh, K., Vladimirova, S. A., Blair, C., Bracken, J. T., Koochekian, N., Schield, D. R., ... & Jezkova, T. (2021). The effects of climate and demographic history in shaping genomic variation across populations of the Desert Horned Lizard (Phrynosoma platyrhinos). Molecular Ecology, 30(18), 4481-4496.

Examples

```
data("HornedLizard_Pop")
data("HornedLizard_VCF")
Test <- Differentiation(data = HornedLizard_VCF, pops = HornedLizard_Pop, write = FALSE)
```

HornedLizard_VCF — *A vcfR object to be used in* `Heterozygosity` *and* `Differentiation`.

Description

Data frame containing 4 columns and 72 rows

Usage

```
data(HornedLizard_Pop)
```

Format

A vcfR object

vcfR object A vcfR object containing genotype and sample informaiton for 72 individuals.

...

Source

Farleigh, K., Vladimirova, S. A., Blair, C., Bracken, J. T., Koochekian, N., Schield, D. R., ... & Jezkova, T. (2021). The effects of climate and demographic history in shaping genomic variation across populations of the Desert Horned Lizard (Phrynosoma platyrhinos). Molecular Ecology, 30(18), 4481-4496.

Examples

```
data("HornedLizard_Pop")
data("HornedLizard_VCF")
Test <- Heterozygosity(data = HornedLizard_VCF, pops = HornedLizard_Pop, write = FALSE)
```

Network_map — *A function to map statistics (i.e., genetic differentiation) between points as a network on a map.*

Description

A function to map statistics (i.e., genetic differentiation) between points as a network on a map.

Usage

```
Network_map(
  dat,
  pops,
  neighbors,
  col,
  statistic = NULL,
  breaks = NULL,
  Lat_buffer = 1,
  Long_buffer = 1,
  Latitude_col = NULL,
  Longitude_col = NULL
)
```

Arguments

dat	Data frame or character string that supplies the input data. If it is a character string, the file should be a csv. If it is a csv, the 1st row should contain the individual/population names. The columns should also be named in this fashion.
pops	Data frame or character string that supplies the input data. If it is a character string, the file should be a csv. The columns should be named Sample, containing the sample IDs; Population indicating the population assignment of the individual; Long, indicating the longitude of the sample; Lat, indicating the latitude of the sample. Alternatively, see the Longitude_col and Latitude_col arguments.
neighbors	Numeric or character. The number of neighbors to plot connections with, or the specific relationship that you want to visualize. Names should match those in the population assignment file and be seperated by an underscore. If I want to visualize the relationship between East and West, for example, I would set neighbors = "East_West".
col	Character vector indicating the colors you wish to use for plotting.
statistic	Character indicating the statistic being plotted. This will be used to title the legend. The legend title will be blank if left as NULL.
breaks	Numeric. The breaks used to generate the color ramp when plotting. Users should supply 3 values if custom breaks are desired.
Lat_buffer	Numeric. A buffer to customize visualization.
Long_buffer	Numeric. A buffer to customize visualization.
Latitude_col	Numeric. The number of the column indicating the latitude for each sample. If this is not null, PopGenHelpR will use this column instead of looking for the Lat column.
Longitude_col	Numeric. The number of the column indicating the longitude for each sample. If this is not null, PopGenHelpR will use this column instead of looking for the Long column.

Value

A list containing the map and the matrix used to plot the map.

Author(s)

Keaka Farleigh

Examples

```
data(Fst_dat)
Fst <- Fst_dat[[1]]
Loc <- Fst_dat[[2]]
Test <- Network_map(dat = Fst, pops = Loc,
neighbors = 2,col = c('#4575b4', '#91bfdb', '#e0f3f8','#fd8d3c','#fc4e2a'),
statistic = "Fst", Lat_buffer = 1, Long_buffer = 1)
```

Pairwise_heatmap *A function to plot a heatmap from a symmetric matrix.*

Description

A function to plot a heatmap from a symmetric matrix.

Usage

```
Pairwise_heatmap(dat, statistic, col = NULL)
```

Arguments

dat	Data frame or character string that supplies the input data. If it is a character string, the file should be a csv. If it is a csv, the 1st row should contain the individual/population names. The columns should also be named in this fashion.
statistic	Character indicating the statistic represented in the matrix, this will be used to label the plot.
col	Character vector indicating the colors to be used in plotting. The vector should contain two colors, the first will be the low value, the second will be the high value.

Value

A heatmap plot

Examples

```
#' data(Fst_dat)
Fst <- Fst_dat[[1]]
Fstat_plot <- Pairwise_heatmap(dat = Fst, statistic = 'FST')
```

PCA — *A function to perform principal component analysis (PCA) on genetic data. Loci with missing data will be removed prior to PCA.*

Description

A function to perform principal component analysis (PCA) on genetic data. Loci with missing data will be removed prior to PCA.

Usage

```
PCA(
  data,
  center = TRUE,
  scale = FALSE,
  missing_value = NA,
  write = FALSE,
  prefix = NULL
)
```

Arguments

data	Character. String indicating the name of the vcf file, geno file or vcfR object to be used in the analysis.
center	Boolean. Whether or not to center the data before principal component analysis.
scale	Boolean. Whether or not to scale the data before principal component analysis.
missing_value	Character. String indicating missing data in the input data. It is assumed to be NA, but that may not be true (is likely not) in the case of geno files.
write	Boolean. Whether or not to write the output to files in the current working directory. There will be two files, one for the individual loadings and the other for the percent variance explained by each axis.
prefix	Character. Optional argument. String that will be appended to file output. Please provide a prefix if write is set to TRUE.

Value

A list containing two elements: the loadings of individuals on each principal component and the variance explained by each principal component.

Author(s)

Keaka Farleigh

Examples

```
data("HornedLizard_VCF")
Test <- PCA(data = HornedLizard_VCF)
```

`Piechart_map` *Plot a map of ancestry pie charts.*

Description

Plot a map of ancestry pie charts.

Usage

```
Piechart_map(
  anc.mat,
  pops,
  K,
  plot.type = "all",
  col,
  piesize = 0.35,
  Lat_buffer,
  Long_buffer,
  Latitude_col = NULL,
  Longitude_col = NULL
)
```

Arguments

`anc.mat`	Data frame or character string that supplies the input data. If it is a character string, the file should be a csv. The first column should be the names of each sample/population, followed by the estimated contribution of each cluster to that individual/pop.
`pops`	Data frame or character string that supplies the input data. If it is a character string, the file should be a csv. The columns should be named Sample, containing the sample IDs; Population indicating the population assignment of the individual, population and sample names must be the same type (i.e., both numeric or both characters); Long, indicating the longitude of the sample; Lat, indicating the latitude of the sample. Alternatively, see the Longitude_col and Latitude_col arguments.
`K`	Numeric.The number of genetic clusters in your data set, please contact the package authors if you need help doing this.
`plot.type`	Character string. Options are all, individual, and population. All is default and recommended, this will plot a piechart map for both the individuals and populations.
`col`	Character vector indicating the colors you wish to use for plotting.
`piesize`	Numeric. The radius of the pie chart for ancestry mapping.
`Lat_buffer`	Numeric. A buffer to customize visualization.
`Long_buffer`	Numeric. A buffer to customize visualization.

`Latitude_col` Numeric. The number of the column indicating the latitude for each sample. If this is not null, PopGenHelpR will use this column instead of looking for the Lat column.

`Longitude_col` Numeric. The number of the column indicating the longitude for each sample. If this is not null, PopGenHelpR will use this column instead of looking for the Long column.

Value

A list containing your plots and the data frames used to generate the plots.

Author(s)

Keaka Farleigh

Examples

```
data(Q_dat)
Qmat <- Q_dat[[1]]
rownames(Qmat) <- Qmat[,1]
Loc <- Q_dat[[2]]
Test_all <- Piechart_map(anc.mat = Qmat, pops = Loc, K = 5,
plot.type = 'all', col = c('#d73027', '#fc8d59', '#e0f3f8', '#91bfdb', '#4575b4'), piesize = 0.35,
Lat_buffer = 1, Long_buffer = 1)
```

`Plot_coordinates` *A function to plot coordinates on a map.*

Description

A function to plot coordinates on a map.

Usage

```
Plot_coordinates(
  dat,
  col = c("#A9A9A9", "#000000"),
  size = 3,
  Lat_buffer = 1,
  Long_buffer = 1,
  Latitude_col = NULL,
  Longitude_col = NULL
)
```

Arguments

`dat`	Data frame or character string that supplies the input data. If it is a character string, the file should be a csv. The coordinates of each row should be indicated by columns named Longitude and Latitude. Alternatively, see the Latitude_col and Longitude_col arugments.
`col`	Character vector indicating the colors you wish to use for plotting, two colors are allowed. The first color will be the fill color, the second is the outline color. For example, if I want red points with a black outline I would set col to col = c("#FF0000", "#000000").
`size`	Numeric. The size of the points to plot.
`Lat_buffer`	Numeric. A buffer to customize visualization.
`Long_buffer`	Numeric. A buffer to customize visualization.
`Latitude_col`	Numeric. The number of the column indicating the latitude for each sample. If this is not null, PopGenHelpR will use this column instead of looking for the Latitude column.
`Longitude_col`	Numeric. The number of the column indicating the longitude for each sample. If this is not null, PopGenHelpR will use this column instead of looking for the Longitude column.

Value

A ggplot object.

Author(s)

Keaka Farleigh

Examples

```
data("HornedLizard_Pop")
Test <- Plot_coordinates(HornedLizard_Pop)
```

`Point_map` *A function to map statistics as colored points on a map.*

Description

A function to map statistics as colored points on a map.

Usage

```
Point_map(
  dat,
  statistic,
  size = 3,
  breaks = NULL,
  col,
  out.col = NULL,
  Lat_buffer = 1,
  Long_buffer = 1,
  Latitude_col = NULL,
  Longitude_col = NULL
)
```

Arguments

dat	Data frame or character string that supplies the input data. If it is a character string, the file should be a csv. The first column should be the statistic to be plotted. The coordinates of each row should be indicated by columns named Longitude and Latitude. Alternatively, see the Longitude_col and Latitude_col arguments.
statistic	Character string. The statistic to be plotted.
size	Numeric. The size of the points to plot.
breaks	Numeric. The breaks used to generate the color ramp when plotting. Users should supply 3 values if custom breaks are desired.
col	Character vector indicating the colors you wish to use for plotting, three colors are allowed (low, mid, high). The first color will be the low color, the second the middle, the third the high.
out.col	Character. A color for outlining points on the map. There will be no visible outline if left as NULL.
Lat_buffer	Numeric. A buffer to customize visualization.
Long_buffer	Numeric. A buffer to customize visualization.
Latitude_col	Numeric. The number of the column indicating the latitude for each sample. If this is not null, PopGenHelpR will use this column instead of looking for the Latitude column.
Longitude_col	Numeric. The number of the column indicating the longitude for each sample. If this is not null, PopGenHelpR will use this column instead of looking for the Longitude column.

Value

A list containing maps and the data frames used to generate them.

Author(s)

Keaka Farleigh

Examples

```
data(Het_dat)
Test <- Point_map(Het_dat, statistic = "Heterozygosity")
```

`Private.alleles` *A function to estimate the number of private alleles in each population.*

Description

A function to estimate the number of private alleles in each population.

Usage

```
Private.alleles(
  data,
  pops,
  write = FALSE,
  prefix = NULL,
  population_col = NULL,
  individual_col = NULL
)
```

Arguments

`data`	Character. String indicating the name of the vcf file or vcfR object to be used in the analysis.
`pops`	Character. String indicating the name of the population assignment file or dataframe containing the population assignment information for each individual in the data. This file must be in the same order as the vcf file and include columns specifying the individual and the population that individual belongs to. The first column should contain individual names and the second column should indicate the population assignment of each individual. Alternatively, you can indicate the column containing the individual and population information using the individual_col and population_col arguments.
`write`	Boolean. Optional argument indicating Whether or not to write the output to a file in the current working directory. This will output to files; 1) the table of private allele counts per population (named prefix_PrivateAlleles_countperpop) and 2) metadata associated with the private alleles (named prefix_PrivateAlleles_metadata). Please supply a prefix it you write files to your working directory as a best practice.
`prefix`	Character. Optional argument indicating a string that will be appended to file output. Please set a prefix if write is TRUE.
`population_col`	Numeric. Optional argument (a number) indicating the column that contains the population assignment information.
`individual_col`	Numeric. Optional argument (a number) indicating the column that contains the individuals (i.e., sample name) in the data.

Value

A list containing the count of private alleles in each population and the metadata for those alleles. The metadata is a list that contains the private allele and locus name for each population.

Author(s)

Keaka Farleigh

Examples

```
data("HornedLizard_Pop")
data("HornedLizard_VCF")
Test <- Private.alleles(data = HornedLizard_VCF, pops = HornedLizard_Pop, write = FALSE)
```

Q_dat — *A list representing a q-matrix and the locality information associated with the qmatrix*

Description

List with two elements

Usage

```
data(Q_dat)
```

Format

A list with two elements:

Qmat A q-matrix with 6 columns and 30 rows, the first column lists the sample name and the remaining 5 represent the contribution a genetic cluster to that individuals ancestry

Loc_dat The locality information for each individual in the q-matrix

...

Source

Data was generated by package authors.

Examples

```
data(Q_dat)
Qmat <- Q_dat[[1]]
rownames(Qmat) <- Qmat[,1]
Loc <- Q_dat[[2]]
Test_all <- Ancestry_barchart(anc.mat = Qmat, pops = Loc, K = 5,
plot.type = 'all',col = c('#d73027', '#fc8d59', '#e0f3f8', '#91bfdb', '#4575b4'))
```

Index

∗ **datasets**
    Fst_dat, 5
    Het_dat, 8
    HornedLizard_Pop, 9
    HornedLizard_VCF, 10
    Q_dat, 19

Ancestry_barchart, 2

Differentiation, 3

Fst_dat, 5

Het_dat, 8
Heterozygosity, 6
HornedLizard_Pop, 9
HornedLizard_VCF, 10

Network_map, 10

Pairwise_heatmap, 12
PCA, 13
Piechart_map, 14
Plot_coordinates, 15
Point_map, 16
Private.alleles, 18

Q_dat, 19