# Package 'GGoutlieR'

January 20, 2025

**Type** Package

**Title** Identify Individuals with Unusual Geo-Genetic Patterns

**Version** 1.0.2

**Description**

Identify and visualize individuals with unusual association patterns of genetics and geography using the approach of Chang and Schmid (2023) <doi:10.1101/2023.04.06.535838>. It detects potential outliers that violate the isolation-by-distance assumption using the K-nearest neighbor approach. You can obtain a table of outliers with statistics and visualize unusual geo-genetic patterns on a geographical map. This is useful for landscape genomics studies to discover individuals with unusual geography and genetics associations from a large biological sample.

**License** MIT + file LICENSE

**Depends** R (>= 3.5.0)

**Imports** stats4, FastKNN, foreach, doParallel, parallel, scales, RColorBrewer, ggforce, rlang, stats, tidyr, utils, rnaturalearth, sf, ggplot2, cowplot

**Suggests** rnaturalearthdata

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.2.3

**NeedsCompilation** no

**Author** Che-Wei Chang [aut, cre] (<https://orcid.org/0000-0001-6704-3732>),
Karl Schmid [ths] (<https://orcid.org/0000-0001-5129-895X>)

**Maintainer** Che-Wei Chang <cheweichang92@gmail.com>

**Repository** CRAN

**Date/Publication** 2023-10-15 21:50:18 UTC

# Contents

---

`anc_coeff_to_GeneticSimilarityMatrix`
                    *calculate genetic similarity from ancestry coefficients.*

---

### Description

    calculate genetic similarity from ancestry coefficients.

### Usage

    anc_coeff_to_GeneticSimilarityMatrix(anc_coef)

### Arguments

    anc_coef          a matrix of ancestry coefficients with samples by rows. Each column corre-
                      sponds to an ancestral population.

### Details

    Since ancestry coefficients can be interpreted as a propotion of genome derived from a specific an-
    cestral population, this function calculate genetic similarity as the probability of a random genome
    segment of two individuals derived from the same ancestral population.

---

`get_GGNet_adjacency_matrix`
                    *obtain an adjacency matrix to make a network graph*

---

### Description

    'get_GGNet_adjacency_matrix' calculates p-values based on the KNNs and heuristic Gamma dis-
    tribution obtained in the outlier identification processes. The matrices of p-values are then multi-
    plied with the given genetic similarity matrix to form adjacency matrices.

**Usage**

```
get_GGNet_adjacency_matrix(
  ggoutlier_res,
  geo_coord,
  gen_coord,
  mutual = FALSE,
  adjust_p_value = TRUE
)
```

**Arguments**

| | |
|---|---|
| ggoutlier_res | an output from 'ggoutlier' |
| geo_coord | a two-column matrix or data.frame. the first column is longitude and the second one is latitude. |
| gen_coord | a matrix of "coordinates in a genetic space". Users can provide ancestry coefficients or eigenvectors for calculation. If, for example, ancestry coefficients are given, each column corresponds to an ancestral population. Samples are ordered in rows as in 'geo_coord'. Users have to provide 'pgdM' if 'gen_coord' is not given. |
| mutual | logic. If a multi-stage test is used in the outlier identification, some samples could not be a mutual neighbor with its K nearest neighbors. In this case, setting 'mutual=TRUE' can force those samples to become mutual neighbors in the output adjacency matrix. |
| adjust_p_value | logic. If 'adjust_p_value=TRUE', p values are adjusted for each nearest neighbor of a given sample. |

**Value**

a list consisting of four matrices that can be used in building network graphs. The default is 'TRUE' 'GeoSP_pvalue' is a matrix describing the strength of edges as p values from the empirical Gamma distribution identified by 'detect_outlier_in_GeoSpace' 'GenSP_pvalue' is a matrix describing the strength of edges as p values from the empirical Gamma distribution identified by 'detect_outlier_in_GeneticSpace'

---

ggoutlier                    *Identify outliers with unusual geo-genetic patterns*

---

**Description**

This function is used to identify outliers with unusual geo-genetic patterns using the KNN approach. For the details of the outlier detection approach, please see Chang and Schmid 2023 (doi:https://doi.org/10.1101/2023.04.06.535838)

## Usage

```
ggoutlier(
  geo_coord,
  gen_coord,
  pgdM = NULL,
  method = c("geneticKNN", "geoKNN", "composite"),
  K = NULL,
  k_geneticKNN = NULL,
  k_geoKNN = NULL,
  klim = c(3, 50),
  make_fig = FALSE,
  plot_dir = ".",
  w_geo = 1,
  w_genetic = 2,
  p_thres = 0.05,
  s = 100,
  min_nn_dist = 100,
  cpu = 1,
  geneticKNN_output = NULL,
  geoKNN_output = NULL,
  verbose = TRUE,
  multi_stages = TRUE,
  maxIter = NULL,
  keep_all_stg_res = FALSE,
  warning_minR2 = 0.9
)
```

## Arguments

| | |
|---|---|
| geo_coord | matrix or data.frame with two columns. The first column is longitude and the second one is latitude. |
| gen_coord | matrix. A matrix of "coordinates in a genetic space". Users can provide ancestry coefficients or eigenvectors for calculation. If, for example, ancestry coefficients are given, each column corresponds to an ancestral population. Samples are ordered in rows as in 'geo_coord'. |
| pgdM | matrix. A pairwise genetic distance matrix. Users can provide a customized genetic distance matrix with this argument. Samples are ordered in rows and columns as in the rows of 'geo_coord'. The default of 'pgdM' is 'NULL'. If 'pgdM' is not provided, a genetic distance matrix will be calculated from 'gen_coord'. |
| method | string. The method to run 'GGoutlieR'. It can be '"composite"', '"geneticKNN"', or '"geoKNN"'. |
| K | integer. Number of the nearest neighbors. If 'K' is not 'NULL', the value will pass to 'k_geneticKNN' and 'k_geoKNN'. |
| k_geneticKNN | integer. Number of the nearest neighbors in a genetic space. The default is 'NULL'. The 'ggoutlier' will search the optimal K if 'k_geneticKNN = NULL'. |

| | |
|---|---|
| k_geoKNN | integer. Number of the nearest neighbors in a geographical space. the default is 'NULL'. The 'ggoutlier' will search the optimal K if 'k_geoKNN = NULL'. |
| klim | vector. A range of K to search for the optimal number of nearest neighbors. The default is 'klim = c(3, 50)' |
| make_fig | logic. If 'make_fig = TRUE', plots for diagnosing GGoutlieR analysis will be generated and saved to 'plot_dir'. The default is 'FALSE' |
| plot_dir | string. The path to save plots |
| w_geo | numeric. A value controlling the power of distance weight in geographical KNN prediction. |
| w_genetic | numeric. A value controlling the power of distance weight in genetic KNN prediction. |
| p_thres | numeric. A significance level |
| s | integer. A scalar of geographical distance. The default 's=100' scales the distance to a unit of 0.1 kilometer. |
| min_nn_dist | numeric. A minimal geographical distance for searching KNNs. Neighbors of a focal sample within this distance will be excluded from the KNN searching procedure. |
| cpu | integer. Number of CPUs to use for searching the optimal K. |
| geneticKNN_output | |
| | output of 'ggoutlier_geneticKNN'. Users can use this argument if running 'ggoutlier_geneticKNN' in advance. |
| geoKNN_output | output of 'ggoutlier_geoKNN'. Users can use this argument if running 'ggoutlier_geoKNN' in advance. |
| verbose | logic. If 'verbose = FALSE', 'ggoutlier' will suppress printout messages. |
| multi_stages | logic. A multi-stage test will be performed if is 'TRUE' (the default is 'TRUE'). |
| maxIter | numeric. Maximal iteration number of multi-stage KNN test. |
| keep_all_stg_res | |
| | logic. Results from all iterations of the multi-stage test will be retained if it is'TRUE'. (the default is 'FALSE') |
| warning_minR2 | numeric. The prediction accuracy of KNN is evaluated as R^2 to assess the violation of isolation-by-distance expectation. If any R^2 is larger than 'warning_minR2', a warning message will be reported at the end of your analysis. |

**Value**

an object of 'list'. If you set 'method = "composite"', 'ggoutlier' will return a nested 'list' with two subsidiary 'list' which are '"geneticKNN_result"' and '"geoKNN_result"'. Each subsidiary list includes five items: 'statistics' is a 'data.frame' consisting of the D_geography ("Dgeo") or D_genetics ("Dg") values, p values and a column of logic values showing if a sample is an outlier or not. 'threshold' is a 'data.frame' recording the significance threshold. 'gamma_parameter' is a vector recording the parameter of the heuristic Gamma distribution. 'knn_index' and 'knn_name' are a 'data.frame' recording the K nearest neighbors of each sample. The subsidiary list 'geneticKNN_result' has an additional item called '"scalar"', which records the value of geographical distance scalar used in the computation. If you set 'method = "geneticKNN"', or 'method = "geoKNN"', 'ggoutlier' will return a 'list' respectively corresponding to '"geneticKNN_result"' or '"geoKNN_result"'.

## Examples

```
library(GGoutlieR)
data("ipk_anc_coef") # get ancestry coefficients
data("ipk_geo_coord") # get geographical coordinates


#DON'T RUN: this analysis will take a few minutes
## Not run:
ggoutlier_example <-
                ggoutlier(geo_coord = ipk_geo_coord,
                          gen_coord = ipk_anc_coef,
                          klim = c(3,6),
                          p_thres = 0.01,
                          cpu = 2,
                          method = "composite",
                          verbose = FALSE,
                          min_nn_dist = 1000,
                          multi_stages = FALSE)

head(summary_ggoutlier(ggoutlier_example)) # get a summary table


## End(Not run)
```

---

ggoutlier_compositeKNN

*GGoutlieR with the composite approach*

---

## Description

perform outlier identification with genetic space KNN and geographical space KNN. For the details of the outlier detection approach, please see the supplementary material of Chang and Schmid 2023 (doi:https://doi.org/10.1101/2023.04.06.535838)

## Usage

```
ggoutlier_compositeKNN(
  geo_coord,
  gen_coord,
  pgdM = NULL,
  k_geneticKNN = NULL,
  k_geoKNN = NULL,
  klim = c(3, 50),
  make_fig = FALSE,
  plot_dir = ".",
  w_geo = 1,
  w_genetic = 2,
  p_thres = 0.05,
  n = 10^6,
```

```
    s = 100,
    min_nn_dist = 1000,
    multi_stages = TRUE,
    maxIter = NULL,
    keep_all_stg_res = FALSE,
    warning_minR2 = 0.9,
    cpu = 1,
    geneticKNN_output = NULL,
    geoKNN_output = NULL,
    verbose = TRUE
)
```

## Arguments

| | |
|---|---|
| geo_coord | matrix or data.frame with two columns. The first column is longitude and the second one is latitude. |
| gen_coord | matrix. A matrix of "coordinates in a genetic space". Users can provide ancestry coefficients or eigenvectors for calculation. If, for example, ancestry coefficients are given, each column corresponds to an ancestral population. Samples are ordered in rows as in 'geo_coord'. |
| pgdM | matrix. A pairwise genetic distance matrix. Users can provide a customized genetic distance matrix with this argument. Samples are ordered in rows and columns as in the rows of 'geo_coord'. The default of 'pgdM' is 'NULL'. If 'pgdM' is not provided, a genetic distance matrix will be calculated from 'gen_coord'. NOTE: the genetic distance matrix is used in the search of KNN and as weights of KNN regression. |
| k_geneticKNN | integer. Number of the nearest neighbors in a genetic space. The default is 'NULL'. The 'ggoutlier' will search the optimal K if 'k_geneticKNN = NULL'. |
| k_geoKNN | integer. Number of the nearest neighbors in a geographical space. the default is 'NULL'. The 'ggoutlier' will search the optimal K if 'k_geoKNN = NULL'. |
| klim | vector. A range of K to search for the optimal number of nearest neighbors. The default is 'klim = c(3, 50)' |
| make_fig | logic. If 'make_fig = TRUE', plots for diagnosing GGoutlieR analysis will be generated and saved to 'plot_dir'. The default is 'FALSE' |
| plot_dir | string. The path to save plots |
| w_geo | numeric. A value controlling the power of distance weight in geographical KNN prediction. |
| w_genetic | numeric. A value controlling the power of distance weight in genetic KNN prediction. |
| p_thres | numeric. A significance level |
| n | numeric. A number of random samples to draw from the null distribution for making a graph. |
| s | integer. A scalar of geographical distance. The default 's=100' scales the distance to a unit of 0.1 kilometer. |

| | |
|---|---|
| min_nn_dist | numeric. A minimal geographical distance for searching KNNs. Neighbors of a focal sample within this distance will be excluded from the KNN searching procedure. |
| multi_stages | logic. A multi-stage test will be performed if is 'TRUE' (the default is 'TRUE'). |
| maxIter | numeric. Maximal iteration number of multi-stage KNN test. |
| keep_all_stg_res | |
| | logic. Results from all iterations of the multi-stage test will be retained if it is'TRUE'. (the default is 'FALSE') |
| warning_minR2 | numeric. The prediction accuracy of KNN is evaluated as R^2 to assess the violation of isolation-by-distance expectation. If any R^2 is larger than 'warning_minR2', a warning message will be reported at the end of your analysis. |
| cpu | integer. Number of CPUs to use for searching the optimal K. |
| geneticKNN_output | |
| | output of 'ggoutlier_geneticKNN'. Users can use this argument if running 'ggoutlier_geneticKNN' in advance. |
| geoKNN_output | output of 'ggoutlier_geoKNN'. Users can use this argument if running 'ggoutlier_geoKNN' in advance. |
| verbose | logic. If 'verbose = FALSE', 'ggoutlier' will suppress printout messages. |

### Value

an object of nested 'list' with two subsidiary 'list' which are '"geneticKNN_result"' and '"geoKNN_result"'. Each subsidiary list includes five items: 'statistics' is a 'data.frame' consisting of the D_geography ("Dgeo") or D_genetics ("Dg") values, p values and a column of logic values showing if a sample is an outlier or not. 'threshold' is a 'data.frame' recording the significance threshold. 'gamma_parameter' is a vector recording the parameter of the heuristic Gamma distribution. 'knn_index' and 'knn_name' are a 'data.frame' recording the K nearest neighbors of each sample. The subsidiary list 'geneticKNN_result' has an additional item called '"scalar"', which records the value of geographical distance scalar used in the computation.

---

| | |
|---|---|
| ggoutlier_example | *Example of 'ggoutlier' output* |

---

### Description

Example of 'ggoutlier' output

### Format

list

---

ggoutlier_geneticKNN    *GGoutlieR with the genetic KNN approach*

---

**Description**

identify samples geographically remote from K genetically nearest neighbors (genetic KNN). For the details of the outlier detection approach, please see the supplementary material of Chang and Schmid 2023 (doi:https://doi.org/10.1101/2023.04.06.535838)

**Usage**

```
ggoutlier_geneticKNN(
  geo_coord,
  gen_coord = NULL,
  pgdM = NULL,
  k = NULL,
  klim = c(3, 50),
  make_fig = FALSE,
  plot_dir = ".",
  w_power = 2,
  p_thres = 0.05,
  n = 10^6,
  s = 100,
  multi_stages = TRUE,
  maxIter = NULL,
  keep_all_stg_res = FALSE,
  warning_minR2 = 0.9,
  cpu = 1,
  verbose = TRUE
)
```

**Arguments**

| | |
|---|---|
| geo_coord | matrix or data.frame with two columns. The first column is longitude and the second one is latitude. |
| gen_coord | matrix. A matrix of "coordinates in a genetic space". Users can provide ancestry coefficients or eigenvectors for calculation. If, for example, ancestry coefficients are given, each column corresponds to an ancestral population. Samples are ordered in rows as in 'geo_coord'. |
| pgdM | matrix. A pairwise genetic distance matrix. Users can provide a customized genetic distance matrix with this argument. Samples are ordered in rows and columns as in the rows of 'geo_coord'. The default of 'pgdM' is 'NULL'. If 'pgdM' is not provided, a genetic distance matrix will be calculated from 'gen_coord'. |
| k | integer. Number of the nearest neighbors. |

| klim | vector. A range of K to search for the optimal number of nearest neighbors. The default is 'klim = c(3, 50)' |
|---|---|
| make_fig | logic. If 'make_fig = TRUE', plots for diagnosing GGoutlieR analysis will be generated and saved to 'plot_dir'. The default is 'FALSE' |
| plot_dir | string. The path to save plots |
| w_power | numanceric. A value controlling the power of distance weight in genetic KNN prediction. |
| p_thres | numeric. A significe level |
| n | numeric. A number of random samples to draw from the null distribution for making a graph. |
| s | integer. A scalar of geographical distance. The default 's=100' scales the distance to a unit of 0.1 kilometer. |
| multi_stages | logic. A multi-stage test will be performed if is 'TRUE' (the default is 'TRUE'). |
| maxIter | numeric. Maximal iteration number of multi-stage KNN test. |
| keep_all_stg_res | |
| | logic. Results from all iterations of the multi-stage test will be retained if it is'TRUE'. (the default is 'FALSE') |
| warning_minR2 | numeric. The prediction accuracy of KNN is evaluated as R^2 to assess the violation of isolation-by-distance expectation. If any R^2 is larger than 'warning_minR2', a warning message will be reported at the end of your analysis. |
| cpu | integer. Number of CPUs to use for searching the optimal K. |
| verbose | logic. If 'verbose = FALSE', 'ggoutlier' will suppress printout messages. |

## Value

an object of 'list' including six items. 'statistics' is a 'data.frame' consisting of the 'D geography' "Dgeo" values, p values and a column of logic values showing if a sample is an outlier or not. 'threshold' is a 'data.frame' recording the significance threshold. 'gamma_parameter' is a vector recording the parameter of the heuristic Gamma distribution. 'knn_index' and 'knn_name' are a 'data.frame' recording the K nearest neighbors of each sample. 'scalar' is the value of geographical distance scalar used in the computation.

---

| ggoutlier_geoKNN | *GGoutlieR with the geographical KNN approach* |
|---|---|

## Description

identify samples genetically different from K nearest geographical neighbors (geographical KNN). For the details of the outlier detection approach, please see the supplementary material of Chang and Schmid 2023 (doi:https://doi.org/10.1101/2023.04.06.535838)

## Usage

```
ggoutlier_geoKNN(
  geo_coord,
  gen_coord,
  min_nn_dist = 100,
  k = NULL,
  klim = c(3, 50),
  s = 100,
  make_fig = FALSE,
  plot_dir = ".",
  w_power = 1,
  p_thres = 0.05,
  n = 10^6,
  multi_stages = TRUE,
  maxIter = NULL,
  keep_all_stg_res = FALSE,
  warning_minR2 = 0.9,
  cpu = 1,
  verbose = TRUE
)
```

## Arguments

| | |
|---|---|
| geo_coord | matrix or data.frame with two columns. The first column is longitude and the second one is latitude. |
| gen_coord | matrix. A matrix of "coordinates in a genetic space". Users can provide ancestry coefficients or eigenvectors for calculation. If, for example, ancestry coefficients are given, each column corresponds to an ancestral population. Samples are ordered in rows as in 'geo_coord'. |
| min_nn_dist | numeric. A minimal geographical distance for searching KNNs. Neighbors of a focal sample within this distance will be excluded from the KNN searching procedure. |
| k | integer. Number of the nearest neighbors. |
| klim | vector. A range of K to search for the optimal number of nearest neighbors. The default is 'klim = c(3, 50)' |
| s | integer. A scalar of geographical distance. The default 's=100' scales the distance to a unit of 0.1 kilometer. |
| make_fig | logic. If 'make_fig = TRUE', plots for diagnosing GGoutlieR analysis will be generated and saved to 'plot_dir'. The default is 'FALSE' |
| plot_dir | string. The path to save plots |
| w_power | numeric. A value controlling the power of distance weight in geographical KNN prediction. |
| p_thres | numeric. A significance level |
| n | numeric. A number of random samples to draw from the null distribution for making a graph. |

multi_stages    logic. A multi-stage test will be performed if is 'TRUE' (the default is 'TRUE').

maxIter         numeric. Maximal iteration number of multi-stage KNN test.

keep_all_stg_res

                logic. Results from all iterations of the multi-stage test will be retained if it
                is'TRUE'. (the default is 'FALSE')

warning_minR2   numeric. The prediction accuracy of KNN is evaluated as R^2 to assess the
                violation of isolation-by-distance expectation. If any R^2 is larger than 'warn-
                ing_minR2', a warning message will be reported at the end of your analysis.

cpu             integer. Number of CPUs to use for searching the optimal K.

verbose         logic. If 'verbose = FALSE', 'ggoutlier' will suppress printout messages.

## Value

an object of 'list' including five items. 'statistics' is a 'data.frame' consisting of the 'D genetics'
("Dg") values, p values and a column of logic values showing if a sample is an outlier or not.
'threshold' is a 'data.frame' recording the significance threshold. 'gamma_parameter' is a vector
recording the parameter of the heuristic Gamma distribution. 'knn_index' and 'knn_name' are a
'data.frame' recording the K nearest neighbors of each sample.

---

ipk_anc_coef                    *Ancestry coefficients of IPK barley landraces*

---

## Description

Ancestry coefficients of IPK barley landraces

## Format

matrix

---

ipk_geo_coord                   *Geographical origins of IPK barley landraces*

---

## Description

Geographical origins of IPK barley landraces

## Format

data.frame

---

plot_ggoutlier                  *Visualize GGoutlieR results on a geographical map*

---

#### Description

Visualize geo-genetic patterns of outliers with their K nearest neighbors

#### Usage

```
plot_ggoutlier(
  ggoutlier_res,
  geo_coord,
  anc_coef = NULL,
  gen_coord = NULL,
  pie_color = NULL,
  map_color = "black",
  p_thres = NULL,
  color_res = 10,
  dot_cex = NULL,
  map_type = c("geographic_knn", "genetic_knn", "both"),
  select_xlim = c(-180, 180),
  select_ylim = c(-90, 90),
  plot_xlim = NULL,
  plot_ylim = NULL,
  only_edges_in_xylim = TRUE,
  pie_r_scale = 1,
  map_resolution = "medium",
  show_knn_pie = FALSE,
  which_sample = NULL,
  add_benchmark_graph = TRUE,
  adjust_p_value_projection = FALSE,
  linewidth_range = c(0.5, 3),
  plot_labels = "auto"
)
```

#### Arguments

| | |
|---|---|
| ggoutlier_res | output of 'ggoutlier' |
| geo_coord | matrix or data.frame with two columns. The first column is longitude and the second one is latitude. |
| anc_coef | matrix. A matrix of ancestry coefficients with samples ordered by rows. Ancestry coefficients are used to make pie charts on a geographical map. This argument is optional. |
| gen_coord | matrix. A matrix of "coordinates in a genetic space". It should be identical to the 'gen_coord' used for running 'ggoutlier' |

| pie_color | string. Colors of pie charts. colors are automatically assigned if 'pie_color = NULL' (which is the default). This argument is optional. |
|---|---|
| map_color | string. Colors of map contours. The default is 'map_color = "black"' |
| p_thres | numeric. A value of significant level. Only outliers (p values less than 'p_thres') are mapped on a geographical map if 'p_thres' is provided (the default is 'NULL'). This argument is optional. |
| color_res | integer. The resolution of color scale. |
| dot_cex | numeric. The size of dots denoting the positions of samples on a geographical map. |
| map_type | string. The type of plot to draw. It can be '"geographic_knn"', '"genetic_knn"' and '"both"'. |
| select_xlim | vector. Values controlling longitude boundaries of a window to select outliers to present on a geographical map. The default is 'select_xlim = c(-180,180)'. |
| select_ylim | vector. Values controlling latitude boundaries of a window to select outliers to present on a geographical map. The default is 'select_ylim = c(-90,90)'. |
| plot_xlim | vector. Values controlling longitude boundaries of a map. |
| plot_ylim | vector. Values controlling latitude boundaries of a map. |
| only_edges_in_xylim | |
| | logic. only the edges with starting points within the given 'select_xlim' and 'select_ylim' will display on a geographical map. If 'FALSE', the edges out of the given boundaries will be removed from your plot. The default is 'TRUE'. |
| pie_r_scale | numeric. A scale controlling the radius of pie charts |
| map_resolution | a character string. The resolution of the geographical map. See details of the 'scale' argument in the manual of 'rnaturalearth::ne_countries()'. The default is 'map_resolution = "medium"' |
| show_knn_pie | logic. If 'TRUE', the ancestry coefficients of K nearest neighbors of significant samples will display on the map. The default is 'FALSE'. |
| which_sample | a string vector of sample ID(s). If users want to only show specific sample(s) |
| add_benchmark_graph | |
| | logic. If 'TRUE', a benchmark graph with only pie charts of ancestry coefficients for comparison with the outlier graph. |
| adjust_p_value_projection | |
| | logic. If 'TRUE', the function will perform KNN prediction by forcing K=1 and compute new p-values for visualization. |
| linewidth_range | |
| | numeric. A vector of two values. It is used to control the minimal and maximal width of KNN network on the geographical map. |
| plot_labels | character. A string of labels for plots. The default is 'plot_labels = "auto"'. This parameter is for 'cowplot::plot_grid'. |

**Details**

Red links on the map denote individual pairs that are genetically similar but geographically remote. The color depth and thickness of red links are proportional to -log10(p) based on the empirical Gamma distribution obtained from 'detect_outlier_in_GeneticSpace'. Blue links on the map denote individual pairs that are genetically different but geographically close. The color depth and thickness of blue links are proportional to -log10(p) based on the empirical Gamma distribution obtained from 'detect_outlier_in_GeoSpace'

**Value**

ggplot object. The plot is geographical map(s) with colored lines showing sample pairs with unusual geo-genetic associations.

**Examples**

```
library(GGoutlieR)
data("ipk_anc_coef") # get ancestry coefficients
data("ipk_geo_coord") # get geographical coordinates
data(ggoutlier_example) # get an example output of ggoutlier
## Not run:
plot_ggoutlier(ggoutlier_res = ggoutlier_example,
               gen_coord = ipk_anc_coef,
               geo_coord = ipk_geo_coord,
               p_thres = 0.025,
               map_type = "both",
               select_xlim = c(-20,140),
               select_ylim = c(10,62),
               plot_xlim = c(-20,140),
               plot_ylim = c(10,62),
               pie_r_scale = 1.8,
               map_resolution = "medium")

## End(Not run)
```

---

summary_ggoutlier         *Summarize GGoutlieR results*

---

**Description**

Get a summary table from the 'ggoutlier' output

**Usage**

```
summary_ggoutlier(ggoutlier_res)
```

**Arguments**

ggoutlier_res     output from the function 'ggoutlier'

**Value**

a table contains IDs of outliers and p-values.

**Examples**

```
library(GGoutlieR)
data(ggoutlier_example) # get an example output of ggoutlier
head(summary_ggoutlier(ggoutlier_example))
```

# Index