Package ‘Orangutan’

February 4, 2026

Title Automated Analysis of Phenotypic Data

Version 2.0.0

Description Provides functions to analyze and visualize meristic and mensural phenotypic data in a comparative framework. The package implements an automated pipeline that summarizes traits, identifies diagnostic variables among groups, performs multivariate and univariate statistical analyses, and produces publication-ready graphics. An earlier implementation (v1.0.0) is described in Torres (2025) <doi:10.64898/2025.12.18.695244>.

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.3.3

Imports adegenet, dplyr, dunn.test, ggplot2, multcompView, RColorBrewer, rlang, tidyr, vegan, withr

Suggests ragg

NeedsCompilation no

Author Javier Torres [aut, cre]

Maintainer Javier Torres <`metalofis@gmail.com`>

Repository CRAN

Date/Publication 2026-02-04 19:50:07 UTC

Contents

`multivariate_tests` *Run multivariate statistical tests*

Description

Performs beta-dispersion and PERMANOVA analyses.

Usage

```
multivariate_tests(df, output_dir, seed_disp = NULL, seed_perm = NULL)
```

Arguments

`df`	A cleaned data frame containing morphometric traits.
`output_dir`	Directory where results will be written.
`seed_disp`	Optional integer; if provided, sets the random seed immediately before beta-dispersion permutation tests to ensure reproducibility.
`seed_perm`	Optional integer; if provided, sets the random seed immediately before PERMANOVA permutation tests to ensure reproducibility.

Value

A list containing multivariate test results.

`run_orangutan` *Run Orangutan*

Description

Runs the full Orangutan morphometric analysis pipeline.

Usage

```
run_orangutan(
  data_path,
  output_dir = file.path(dirname(data_path), "analysis_outputs"),
  apply_allometry = FALSE,
  allometry_var = NULL,
  remove_outliers = FALSE,
  outlier_vars = NULL,
  outlier_tail_pct = 0.05,
  palette_name = "Paired",
  species_to_encircle = character(0),
  seeds = list(betadisper = 123, permanova = 456),
 point_aes = list(point_size = 3.5, jitter_width = 0.1, jitter_alpha = 0.8, jitter_shape
```

```
    = 21, jitter_color = "black", jitter_stroke = 0.35),
  mean_aes = list(size = 1.8, shape = 21, fill = "white", color = "black", stroke = 0.6),
  violin_aes = list(alpha = 0.4),
  box_aes = list(alpha = 0.4, width = 0.15),
  label_aes = list(text_size = 6, axis_text_size = 10, title_size = 12, label_offset =
    0.05),
  label_templates = NULL,
  verbose = FALSE
)
```

Arguments

`data_path`	Path to input CSV file
`output_dir`	Output directory for results
`apply_allometry`	Logical; apply allometric correction
`allometry_var`	Character; size variable for allometry
`remove_outliers`	Logical; remove outliers
`outlier_vars`	Variables used for outlier detection
`outlier_tail_pct`	Tail proportion for Tukey filtering
`palette_name`	RColorBrewer palette name
`species_to_encircle`	Species to encircle in multivariate plots
`seeds`	A named list of integer seeds for reproducibility, with elements: `betadisper` for beta-dispersion permutation tests and `permanova` for PERMANOVA permutation tests. Defaults to `list(betadisper = 123, permanova = 456)`.
`point_aes`	List of point aesthetics
`mean_aes`	List of mean-point aesthetics
`violin_aes`	List of violin aesthetics
`box_aes`	List of boxplot aesthetics
`label_aes`	List of label/text aesthetics
`label_templates`	Optional plot label templates
`verbose`	Logical; if TRUE, print progress messages. Defaults to FALSE.

Value

A list containing results from all analyses

Examples

```
# Create a tiny example dataset in a temporary file
tmp <- tempfile(fileext = ".csv")
toy_data <- data.frame(
  species = c("A", "A", "B", "B", "C", "C"),
  trait1  = c(1, 2, 5, 6, 9, 10),
  trait2  = c(3, 4, 7, 8, 11, 12),
  trait3  = c(2, 3, 6, 7, 10, 11)
)
write.csv(toy_data, tmp, row.names = FALSE)

# Create a temporary output directory
out_dir <- tempdir()

# Set a named list of seeds for reproducibility
seeds <- list(betadisper = 123, permanova = 456)

# Run Orangutan on the toy dataset
res <- run_orangutan(
  data_path = tmp,
  output_dir = out_dir,
  seeds = seeds,
  verbose = FALSE
)

# Inspect returned object
str(res)

# Clean up temporary dataset file
unlink(tmp)
```

Index