

# Package ‘ggparallel’

August 21, 2015

**Version** 0.1.2

**Title** Variations of Parallel Coordinate Plots for Categorical Data

**Author** Heike Hofmann, Marie Vendettuoli

**Maintainer** Heike Hofmann <hofmann@iastate.edu>

**Description** Create hammock plots, parallel sets, and common angle plots with ggplot2.

**License** MIT + file LICENSE

**Depends** R (>= 2.10.0),

**Imports** ggplot2 (>= 0.9.2), reshape2, plyr

**Suggests** RColorBrewer

**URL** <http://github.com/heike/ggparallel>

**LazyData** true

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2015-08-21 07:38:47

## R topics documented:

genes . . . . .	2
ggparallel . . . . .	3
package-ggparallel . . . . .	4
<b>Index</b>	<b>5</b>

genes

*Data linking genes and pathways.***Description**

Table knownGene from track UCSC Genes was downloaded from the UCSC table browser for the human genome assembly (hg18, May 2006) and filtered for a selection of pathways associated with human metabolism was obtained from KEGG PATHWAY database. Bioconductor package KEGG.db was used to provide mappings between gene and pathway identifiers.

**Usage**

genes

**Format**

```
'data.frame': 2768 obs. of 14 variables:
 $ value      : int  9 9 9 9 9 9 9 10 10 ...
 $ name       : Factor w/ 59950 levels "uc001aab.2","uc001aac.2",...: 43054 43058 43053 43055 43056 43057 ...
 $ chrom      : Factor w/ 24 levels "chr1","chr2",...: 8 8 8 8 8 8 8 8 8 ...
 $ strand     : Factor w/ 2 levels "-","+": 2 2 2 2 2 2 2 2 2 ...
 $ txStart    : int  18111894 18113073 18111894 18111894 18113073 18113073 18113073 18123836 18301793 18301793 ...
 $ txEnd      : int  18125100 18125100 18125100 18125100 18125100 18125100 18125100 18124709 18302666 18302666 ...
 $ cdsStart   : int  18123836 18123836 18123836 18118682 18118682 18123836 18123836 18123836 18301793 18301793 ...
 $ cdsEnd     : int  18124709 18124709 18124709 18124709 18124709 18124709 18124709 18124709 18302666 18302666 ...
 $ exonCount  : int  4 4 3 4 4 4 3 1 1 2 ...
 $ exonStarts: Factor w/ 64524 levels "0,1137,5766,12786",...: 19845 19848 19847 19846 19849 19848 19848 19848 ...
 $ exonEnds   : Factor w/ 64568 levels "100000370,100002888,100004625,100007887,100009499,100010117",...: 13218 13218 13218 23139 23139 13218 13218 ...
 $ proteinID  : Factor w/ 36856 levels "", "A0A183", "A0A4V9",...: 48418 48422 48417 48419 48420 48420 48420 ...
 $ alignID    : Factor w/ 66803 levels "uc001aaa.2","uc001aab.2",...: 48418 48422 48417 48419 48420 48420 48420 ...
 $ path       : chr  "hsa00232" "hsa00232" "hsa00232" "hsa00232" ...
```

**References**

Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, Cline MS, Goldman M, Barber GP, Clawson H, Coelho A, Diekhans M, Dreszer TR, Giardine BM, Harte RA, Hillman-Jackson J, Hsu F, Kirkup V, Kuhn RM, Learned K, Li CH, Meyer LR, Pohl A, Raney BJ, Rosenbloom KR, Smith KE, Haussler D, Kent WJ. The UCSC Genome Browser database: update 2011. *Nucleic Acids Res.* 2010 Oct 18. <http://genome.ucsc.edu/index.html?org=Human&db=hg19&hgid=289810087>

Marc Carlson, Seth Falcon, Herve Pages and Nianhua Li (). KEGG.db: A set of annotation maps for KEGG. R package version 2.6.1.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M.; KEGG for integration and interpretation of large-scale molecular datasets. *Nucleic Acids Res.* 40, D109-D114 (2012)

Kanehisa, M. and Goto, S.; KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27-30 (2000).

## Examples

```
library(ggplot2)
library(RColorBrewer)
genes$chrom <- factor(genes$chrom, levels=c(paste("chr", 1:22, sep=""),
      "chrX", "chrY"))
ggparallel(
  list("path", "chrom"),
  text.offset = c(0.03, 0, -0.03),
  data = genes,
  width = 0.1,
  order = c(1, 0),
  angle = 0,
  color = "white",
  factorlevels = c(sapply(unique(genes$chrom), as.character), unique(genes$path))
) +
scale_fill_manual(
  values = c(brewer.pal("YlOrRd", n = 9), rep("grey80", 24)),
  guide = "none"
) +
scale_colour_manual(
  values = c(brewer.pal("YlOrRd", n = 9), rep("grey80", 24)),
  guide = "none"
) +
coord_flip()
```

---

ggparallel

*Variations of parallel coordinate plots*

---

## Description

[ggparallel](#) implements and combines different types of parallel coordinate plots for categorical data: hammock plots, parallel sets plots, common angle plots, and common angle plots with a hammock-like adjustment for line widths.

## Usage

```
ggparallel(vars = list(), data, weight = NULL, method = "angle",
  alpha = 0.5, width = 0.25, order = 1, ratio = NULL, asp = NULL,
  label = TRUE, label.size = 4, text.angle = 90, text.offset = NULL,
  ...)
```

## Arguments

vars	list of variable names to be included in the plotting. Order of the variables is preserved in the display
data	data frame
weight	weighting variable - use character string

method	plotting method to use - one of <code>angle</code> , <code>adj.angle</code> , <code>parset</code> , or <code>hammock</code> , for a hammock plot the aspect ratio needs to be fixed.
alpha	level of alpha blending for the fill color in ribbons, value has to be between 0 and 1, defaults to 0.5.
width	width of variables
order	flag variable with three levels -1, 0, 1 for levels in decreasing order, levels in increasing order and levels unchanged. This variable can be either a scalar or a vector
ratio	used for methods with angle adjustments ( <code>method = 'hammock'</code> , <code>'adj.angle'</code> ): specifies the height (width for horizontal displays) of the widest line as ratio of the overall display height (width for horizontal displays).
asp	aspect ratio of the plot - it will be set to a default of 1 in the case of hammock plots.
label	binary variable (vector), whether labels should be shown.
label.size	numeric value to determine the size in which labels are shown, defaults to 4
text.angle	numeric value in degrees, by which text for labelling is rotated. Ignored if <code>label = FALSE</code>
text.offset	(vector) of values for offset the labels
...	passed on directly to all of the <code>ggplot2</code> commands

### Details

Parallel sets have been suggested by *kosara:2006* as a visualization technique to incorporate categorical variables into a parallel coordinate plot introduced by *wegman:1990* and *inselberg:1985*. The parallel sets implemented here are reduced to representations of neighboring two-dimensional relationships only rather than the hierarchical version originally suggested.

Both versions, however, show perceptual problems with interpreting line widths, leading to potentially wrong conclusions about the data. The hammock display, introduced by *schonlau:2003*, and the common angle plots are two approaches at fixing this problem: in Hammock plots the linewidth is adjusted by a factor countering the strength of the illusion, in the common angle plot all lines are adjusted to show the same angle - making line widths again comparable across ribbons.

Additionally, we can also adjust ribbons in the common angle display for the angle, to make them appear having the same width (or height) across the display. We refer to this method as `adj.angle`.

### Value

returns a `ggplot2` object that can be plotted directly or used as base layer for additional modifications.

---

package-ggparallel      *A package for creating parallel coordinates for categorical data*

---

### Description

The main function `ggparallel` implements three types of parallel coordinate plots for categorical data: hammock plots, parallel sets plots, and common angle plots.

# Index

\*Topic **datasets**

genes, [2](#)

genes, [2](#)

ggparallel, [3](#), [3](#), [4](#)

package-ggparallel, [4](#)

package-ggparallel-package

(package-ggparallel), [4](#)