

# Package ‘CellWindX’

May 27, 2026

**Title** Marker Gene Analysis and Visualization for Single-Cell Data

**Version** 1.0.0

**Description** Provides a 'Seurat'-compatible toolkit for marker gene identification, expression summarization, and visualization of annotated single-cell transcriptomic data. 'CellWindX' identifies top cell-type-enriched markers, calculates marker expression percentages and average expression values across cell groups, and generates publication-oriented dimensional reduction plots, marker heatmaps, and gene-level radar plots. The package includes built-in aesthetic palettes and supports both exploratory analysis and downstream figure preparation for single-cell atlas studies. The workflow is designed to complement single-cell analysis frameworks such as 'Seurat' described by Satija et al. (2015) <[doi:10.1038/nbt.3192](https://doi.org/10.1038/nbt.3192)> and Hao et al. (2021) <[doi:10.1016/j.cell.2021.04.048](https://doi.org/10.1016/j.cell.2021.04.048)>, as well as heatmap visualization methods implemented in 'ComplexHeatmap' described by Gu et al. (2016) <[doi:10.1093/bioinformatics/btw313](https://doi.org/10.1093/bioinformatics/btw313)>.

**License** GPL-3

**Encoding** UTF-8

**RoxygenNote** 7.3.3

**Imports** circlize, ComplexHeatmap, dplyr, ggplot2, grDevices, grid, Matrix, patchwork, Seurat, tidyr, stats

**Suggests** SeuratObject

**NeedsCompilation** no

**Author** Xiaofeng Yang [aut, cre] (affiliation: Chongqing Medical University),  
Shan Li [aut] (affiliation: Chongqing Medical University)

**Maintainer** Xiaofeng Yang <Youngxf02@163.com>

**Depends** R (>= 4.1.0)

**Repository** CRAN

**Date/Publication** 2026-05-27 19:50:27 UTC

## Contents

CellWindX_DimPlot . . . . .	2
-----------------------------	---

CellWindX_GeneRadar . . . . .	4
CellWindX_MarkerHeatmap . . . . .	7
CellWindX_TopMarkersStats . . . . .	9

<b>Index</b>	<b>13</b>
--------------	-----------

---

CellWindX_DimPlot	<i>Visualize Seurat embeddings with CellWindX palettes</i>
-------------------	--

---

## Description

CellWindX\_DimPlot() visualizes UMAP or t-SNE embeddings from a processed Seurat object using built-in CellWindX color palettes. The function is designed for annotated single-cell objects and supports three predefined aesthetic styles: Chinese landscape-inspired, Chongqing modern, and girlish palettes.

## Usage

```
CellWindX_DimPlot(
  object,
  group.by = "seurat_annotatations",
  reduction = c("umap", "tsne"),
  palette = c("shanshui", "chongqing_modern", "girlish"),
  label = TRUE,
  repel = TRUE,
  pt.size = 0.6,
  label.size = 4,
  alpha = 0.9,
  shuffle = TRUE,
  seed = 123,
  title = NULL,
  legend.position = "right"
)
```

## Arguments

object	A Seurat object containing dimensional reduction results.
group.by	Character string. Metadata column used to color cells. Default is "seurat_annotatations".
reduction	Character string. Dimensional reduction to visualize. One of "umap" or "tsne".
palette	Character string. Built-in CellWindX palette to use. One of "shanshui", "chongqing_modern", or "girlish".
label	Logical. Whether to show group labels on the plot. Default is TRUE.
repel	Logical. Whether to repel text labels using ggrepel. Default is TRUE.
pt.size	Numeric. Point size passed to <code>Seurat::DimPlot()</code> . Default is 0.6.
label.size	Numeric. Label font size. Reserved for future extension. Default is 4.

<code>alpha</code>	Numeric or NULL. Point transparency. Values should usually range from 0 to 1. Default is 0.9.
<code>shuffle</code>	Logical. Whether to randomly shuffle plotting order of cells. Default is TRUE.
<code>seed</code>	Integer. Random seed used when <code>shuffle = TRUE</code> . Default is 123.
<code>title</code>	Character string or NULL. Plot title. If NULL, a default title is automatically generated.
<code>legend.position</code>	Character string. Legend position passed to <code>ggplot2::theme()</code> . Common values include "right", "left", "bottom", "top", and "none". Default is "right".

## Details

This function is a wrapper around `Seurat::DimPlot()` with customized color palettes and publication-oriented theme settings. It requires that the selected dimensional reduction has already been computed and stored in `object@reductions`, such as by running `Seurat::RunUMAP()` or `Seurat::RunTSNE()`.

The built-in palettes support up to 20 annotated groups. If the number of groups in `group.by` exceeds 20, the function will return an error.

## Value

A ggplot object generated from `Seurat::DimPlot()`.

## Author(s)

Xiaofeng Yang, Chongqing Medical University

## Examples

```
data("pbmc_small", package = "SeuratObject")

p1 <- CellWindX_DimPlot(
  object = pbmc_small,
  group.by = "groups",
  reduction = "tsne",
  palette = "shanshui"
)

p1
```

---

CellWindX\_GeneRadar *Draw gene-level radar plots across annotated cell groups*

---

## Description

CellWindX\_GeneRadar() visualizes the expression pattern of selected genes across annotated cell groups using straight-line spider radar plots. For each selected gene, the function generates two radar plots: one for the percentage of expressing cells and another for average expression.

## Usage

```
CellWindX_GeneRadar(
  marker_result,
  genes,
  source_cluster = NULL,
  aggregate_fun = c("mean", "max", "first"),
  palette = c("shanshui", "chongqing_modern", "girlish"),
  scale_avg = TRUE,
  scale_pct = FALSE,
  pct_max = 100,
  avg_scale_method = c("max", "zscore", "none"),
  grid_levels = c(25, 50, 75, 100),
  axis_label_size = 5.2,
  grid_label_size = 3.2,
  facet_title_size = 12,
  line_width = 1.2,
  point_size = 2.8,
  fill_alpha = 0.18,
  facet = TRUE,
  ncol = NULL,
  show_points = TRUE,
  show_fill = TRUE,
  show_grid_label = TRUE,
  output_file = NULL,
  output_width = 11,
  output_height = 5.5,
  dpi = 300
)
```

## Arguments

marker_result	A result object returned by CellWindX_TopMarkersStats(), or a data frame containing marker expression statistics. If a data frame is provided, it must contain the columns source_cluster, target_cluster, gene, pct_expr, and avg_expr.
genes	Character vector. Gene symbols to visualize.

source_cluster	Character vector or NULL. Optional source-cluster filter. If provided, only marker statistics from the selected source cluster(s) are used. Default is NULL.
aggregate_fun	Character string. Method used to aggregate duplicated gene-cell-group rows. One of "mean", "max", or "first". Default is "mean".
palette	Character string. Built-in CellWindX palette. One of "shanshui", "chongqing_modern", or "girlish".
scale_avg	Logical. Whether to scale average expression values before plotting. Default is TRUE.
scale_pct	Logical. Whether to scale expression percentage values by gene before plotting. If FALSE, the raw percentage values are used and capped by pct_max. Default is FALSE.
pct_max	Numeric. Maximum value used to cap percentage values when scale_pct = FALSE. Default is 100.
avg_scale_method	Character string. Scaling method for average expression. One of "max", "zscore", or "none". "max" rescales each gene to a 0-100 range by dividing by its maximum value. "zscore" performs gene-wise Z-score scaling and rescales the result to 0-100. "none" keeps the original average expression values. Default is "max".
grid_levels	Numeric vector. Radar grid levels. These values determine the radius of the straight polygon grid lines. Default is c(25, 50, 75, 100).
axis_label_size	Numeric. Font size of cell-group labels placed around the radar plot. Default is 5.2.
grid_label_size	Numeric. Font size of radar grid labels. Default is 3.2.
facet_title_size	Numeric. Font size of facet titles when multiple genes are plotted. Default is 12.
line_width	Numeric. Width of radar polygon lines. Default is 1.2.
point_size	Numeric. Size of points on radar axes. Default is 2.8.
fill_alpha	Numeric. Transparency of polygon fill. Values should usually range from 0 to 1. Default is 0.18.
facet	Logical. Whether to draw multiple genes as separate facets. Default is TRUE.
ncol	Integer or NULL. Number of columns used in facet layout when facet = TRUE. Default is NULL.
show_points	Logical. Whether to show points on radar axes. Default is TRUE.
show_fill	Logical. Whether to fill radar polygons. Default is TRUE.
show_grid_label	Logical. Whether to show numeric grid labels. Default is TRUE.
output_file	Character string or NULL. Optional path used to save the combined radar plot. Supported formats depend on <code>ggplot2::ggsave()</code> , such as .pdf, .png, .tiff, or .svg. Default is NULL.
output_width	Numeric. Width of the saved plot in inches. Default is 11.
output_height	Numeric. Height of the saved plot in inches. Default is 5.5.
dpi	Numeric. Resolution used when saving raster formats. Default is 300.

## Details

This function is designed to directly accept the output generated by `CellWindX_TopMarkersStats()`. It uses the `marker_expr_by_group` table from that result object, which should contain gene expression statistics across cell groups.

Each cell group is represented as one axis of the radar plot. Unlike polar coordinate radar plots, this function manually calculates polygon coordinates, resulting in straight polygon grid lines rather than curved circular grid lines. This style is generally clearer for comparing cell-group-specific expression patterns.

The function draws two panels:

- `pct_plot`: expression percentage across cell groups.
- `avg_plot`: average expression across cell groups.

The expression percentage and average expression panels use intentionally distinct color sets within the selected CellWindX palette. Three built-in palettes are available: "shanshui", "chongqing\_modern", and "girlish".

## Value

A list with class "CellWindX\_GeneRadar" containing:

**pct\_plot** A ggplot object showing expression percentage.

**avg\_plot** A ggplot object showing average expression.

**combined\_plot** A patchwork object combining percentage and average expression plots.

**plot\_data** A data frame used for plotting.

**pct\_coord** Radar coordinates for the expression percentage plot.

**avg\_coord** Radar coordinates for the average expression plot.

**palette** The selected CellWindX palette.

**genes** Genes used for visualization.

**parameters** A list of function parameters used in the plot.

## Author(s)

Xiaofeng Yang, Chongqing Medical University

## Examples

```
marker_df <- data.frame(
  source_cluster = rep(c("T cell", "B cell"), each = 4),
  target_cluster = rep(c("T cell", "B cell", "Myeloid", "Platelet"), times = 2),
  gene = rep(c("CD3D", "MS4A1"), each = 4),
  pct_expr = c(92, 8, 15, 4, 6, 88, 12, 3),
  avg_expr = c(2.8, 0.2, 0.5, 0.1, 0.1, 2.5, 0.4, 0.1),
  stringsAsFactors = FALSE
)

radar_res <- CellWindX_GeneRadar(
```

```

marker_result = marker_df,
genes = c("CD3D", "MS4A1"),
palette = "shanshui",
scale_avg = TRUE,
avg_scale_method = "max",
facet = TRUE,
ncol = 2
)

radar_res$combined_plot

```

---

CellWindX\_MarkerHeatmap

*Draw marker-gene heatmap across annotated cell groups*

---

### Description

CellWindX\_MarkerHeatmap() visualizes marker gene expression statistics across annotated cell groups using either ComplexHeatmap or ggplot2.

### Usage

```

CellWindX_MarkerHeatmap(
  marker_result,
  value_col = c("avg_expr", "pct_expr", "avg_expr_positive"),
  scale_method = c("zscore", "none"),
  plot_engine = c("complex", "ggplot"),
  palette = c("shanshui", "chongqing_modern", "girlish"),
  cluster_rows = FALSE,
  cluster_columns = FALSE,
  show_row_names = TRUE,
  show_column_names = TRUE,
  column_names_rot = 45,
  round_cell = TRUE,
  cell_width_mm = 5.5,
  cell_height_mm = 5.5,
  row_fontsize = 10,
  column_fontsize = 9,
  legend_position = "right",
  heatmap_title = NULL,
  zscore_clip = 2,
  output_file = NULL,
  output_width = 9,
  output_height = 5,
  dpi = 300,
  draw = TRUE
)

```

**Arguments**

marker_result	A result object returned by CellWindX_TopMarkersStats(), or a data frame containing marker expression statistics.
value_col	Character string. Value column used for heatmap visualization. One of "avg_expr", "pct_expr", or "avg_expr_positive".
scale_method	Character string. Scaling method. One of "zscore" or "none".
plot_engine	Character string. Plotting engine. One of "complex" or "ggplot".
palette	Character string. Built-in CellWindX palette.
cluster_rows	Logical. Whether to cluster rows.
cluster_columns	Logical. Whether to cluster columns.
show_row_names	Logical. Whether to show row names.
show_column_names	Logical. Whether to show column names.
column_names_rot	Numeric. Rotation angle of column names.
round_cell	Logical. Whether to draw rounded cells when using ComplexHeatmap.
cell_width_mm	Numeric. Cell width in millimeters.
cell_height_mm	Numeric. Cell height in millimeters.
row_fontsize	Numeric. Row name font size.
column_fontsize	Numeric. Column name font size.
legend_position	Character string. Legend position.
heatmap_title	Character string or NULL. Heatmap title.
zscore_clip	Numeric. Z-score clipping threshold.
output_file	Character string or NULL. Optional output file path.
output_width	Numeric. Output width in inches.
output_height	Numeric. Output height in inches.
dpi	Numeric. Output resolution.
draw	Logical. Whether to draw the heatmap immediately when using ComplexHeatmap.

**Value**

A list with class "CellWindX\_MarkerHeatmap".

---

CellWindX\_TopMarkersStats

*Identify top marker genes and summarize expression statistics*

---

## Description

CellWindX\_TopMarkersStats() identifies the top marker genes for each annotated cell group in a Seurat object and summarizes their expression percentage and average expression. The output is designed to be directly used by downstream CellWindX visualization functions, including CellWindX\_MarkerHeatmap() and CellWindX\_GeneRadar().

## Usage

```
CellWindX_TopMarkersStats(  
  object,  
  group.by = "seurat_annotatations",  
  assay = NULL,  
  slot = "data",  
  top_n = 5,  
  only.pos = TRUE,  
  min.pct = 0.25,  
  logfc.threshold = 0.25,  
  test.use = "wilcox",  
  min.diff.pct = -Inf,  
  expression.threshold = 0,  
  exclude.mt = TRUE,  
  exclude.ribo = FALSE,  
  output_dir = NULL,  
  file_prefix = "CellWindX_top_markers",  
  seed = 123,  
  verbose = TRUE  
)
```

## Arguments

object	A Seurat object.
group.by	Character string. Metadata column used as cell-group annotation for marker detection. Default is "seurat_annotatations".
assay	Character string or NULL. Assay used for marker detection and expression summarization. If NULL, <a href="#">Seurat::DefaultAssay()</a> is used. Default is NULL.
slot	Character string. Assay slot or layer used to calculate expression statistics. Common choices are "data" for normalized expression and "counts" for raw counts. Default is "data".
top_n	Integer. Number of top marker genes selected for each cell group. Default is 5.
only.pos	Logical. Whether to return only positive markers in <a href="#">Seurat::FindAllMarkers()</a> . Default is TRUE.

<code>min.pct</code>	Numeric. Minimum fraction of cells expressing a gene in either of the two groups tested by <code>Seurat::FindAllMarkers()</code> . Default is 0.25.
<code>logfc.threshold</code>	Numeric. Log-fold-change threshold used by <code>Seurat::FindAllMarkers()</code> . Default is 0.25.
<code>test.use</code>	Character string. Statistical test used by <code>Seurat::FindAllMarkers()</code> . Default is "wilcox".
<code>min.diff.pct</code>	Numeric. Minimum difference in detection percentage between the two groups compared by <code>Seurat::FindAllMarkers()</code> . Default is -Inf.
<code>expression.threshold</code>	Numeric. Threshold used to define whether a gene is expressed in a cell when calculating <code>n_expr_cells</code> and <code>pct_expr</code> . Default is 0.
<code>exclude.mt</code>	Logical. Whether to remove mitochondrial genes from selected top markers. Human mitochondrial genes beginning with "MT-" and mouse mitochondrial genes beginning with "mt-" are removed. Default is TRUE.
<code>exclude.ribo</code>	Logical. Whether to remove ribosomal protein genes beginning with "RPS", "RPL", "Rps", or "Rpl". Default is FALSE.
<code>output_dir</code>	Character string or NULL. Optional directory for saving result tables as CSV files. If NULL, no files are written. Default is NULL.
<code>file_prefix</code>	Character string. Prefix used for output CSV files when <code>output_dir</code> is not NULL. Default is "CellWindX_top_markers".
<code>seed</code>	Integer. Random seed set before marker detection. Default is 123.
<code>verbose</code>	Logical. Whether to print progress messages. Default is TRUE.

## Details

This function first runs `Seurat::FindAllMarkers()` using the cell-group annotation specified by `group.by`. For each group, the top N marker genes are selected according to adjusted P value, log-fold change, and detection percentage.

For each selected marker gene, the function then calculates:

- `n_cells`: number of cells in the target group.
- `n_expr_cells`: number of cells with expression greater than `expression.threshold`.
- `pct_expr`: percentage of expressing cells.
- `avg_expr`: average expression across all cells in the group.
- `avg_expr_positive`: average expression among expressing cells only.

The function returns both source-cluster-specific statistics and expression summaries across all annotated groups. This makes it suitable for marker inspection, heatmap visualization, dot-plot-style summaries, and radar plots.

The function is compatible with Seurat v4 and v5 by attempting to retrieve assay data through `layer` first and falling back to `slot` when needed.

**Value**

A list with class "CellWindX\_TopMarkersStats" containing:

**markers\_all** Complete marker table returned by `Seurat::FindAllMarkers()`.

**top\_markers** Top marker genes selected for each source cell group.

**top\_marker\_stats** Expression statistics of each selected marker gene within its source cell group.

**marker\_expr\_by\_group** Expression statistics of selected marker genes across all target cell groups.  
This table is used by downstream CellWindX heatmap and radar functions.

**top\_gene\_table** A compact table listing top marker genes for each cell group.

**parameters** A list of parameters used in the analysis.

**Author(s)**

Xiaofeng Yang, Chongqing Medical University

**Examples**

```
counts <- matrix(
  c(
    25, 26, 24, 27, 25, 26, 24, 28, 25, 27,
    1, 2, 1, 1, 2, 1, 1, 2, 1, 1,
    1, 1, 2, 1, 1, 2, 1, 1, 2, 1,

    1, 1, 2, 1, 1, 2, 1, 1, 2, 1,
    25, 27, 26, 24, 28, 25, 27, 26, 24, 28,
    1, 2, 1, 1, 2, 1, 1, 2, 1, 1,

    1, 2, 1, 1, 2, 1, 1, 2, 1, 1,
    1, 1, 2, 1, 1, 2, 1, 1, 2, 1,
    25, 26, 28, 24, 27, 25, 26, 28, 24, 27,

    5, 6, 5, 6, 5, 6, 5, 6, 5, 6,
    5, 6, 5, 6, 5, 6, 5, 6, 5, 6,
    5, 6, 5, 6, 5, 6, 5, 6, 5, 6
  ),
  nrow = 4,
  byrow = TRUE
)

rownames(counts) <- c("CD3D", "MS4A1", "LYZ", "ACTB")
colnames(counts) <- paste0("Cell", seq_len(ncol(counts)))

counts <- Matrix::Matrix(counts, sparse = TRUE)

object <- Seurat::CreateSeuratObject(counts = counts)
object$cell_type <- rep(c("T cell", "B cell", "Myeloid"), each = 10)
object <- Seurat::NormalizeData(object, verbose = FALSE)

res_marker <- CellWindX_TopMarkersStats(
  object = object,
```

```
group.by = "cell_type",
assay = "RNA",
slot = "data",
top_n = 1,
only.pos = TRUE,
min.pct = 0.10,
logfc.threshold = 0,
test.use = "t",
verbose = FALSE
)

head(res_marker$top_markers)
head(res_marker$top_marker_stats)
head(res_marker$marker_expr_by_group)
res_marker$top_gene_table
```

# Index

CellWindX\_DimPlot, [2](#)  
CellWindX\_GeneRadar, [4](#)  
CellWindX\_MarkerHeatmap, [7](#)  
CellWindX\_TopMarkersStats, [9](#)

ggplot2::ggsave(), [5](#)  
ggplot2::theme(), [3](#)

Seurat::DefaultAssay(), [9](#)  
Seurat::DimPlot(), [2, 3](#)  
Seurat::FindAllMarkers(), [9–11](#)  
Seurat::RunTSNE(), [3](#)  
Seurat::RunUMAP(), [3](#)