Tutorial

Load the dataset   A subset of invasive breast carcinoma data from primary tumor tissue. See **?tcga** for more information on loading the full dataset or metadata.

```
library(tcgaViz)
library(ggplot2)
data(tcga)
head(tcga$genes)
#> # A tibble: 6 x 2
#>   sample         ICOS
#>   <chr>         <dbl>
#> 1 TCGA-3C-AAAU-01  1.25
#> 2 TCGA-3C-AALI-01  5.63
#> 3 TCGA-3C-AALJ-01  5.11
#> 4 TCGA-3C-AALK-01  3.79
#> 5 TCGA-4H-AAAK-01  4.24
#> 6 TCGA-5L-AAT0-01  5.73
head(tcga$cells$Cibersort_ABS)
#> # A tibble: 6 x 24
#>   sample         study B_cell_naive B_cell_memory B_cell_plasma T_cell_CD8.
#>   <chr>          <fct>        <dbl>         <dbl>         <dbl>      <dbl>
#> 1 TCGA-3C-AAAU-01 BRCA     0            0.0221        0.0192       0.0129
#> 2 TCGA-3C-AALI-01 BRCA     0.00754      0.00417       0            0.0645
#> 3 TCGA-3C-AALJ-01 BRCA     0.00520      0.00535       0            0.0358
#> 4 TCGA-3C-AALK-01 BRCA     0            0.00288       0.0516       0.0360
#> 5 TCGA-4H-AAAK-01 BRCA     0.00520      0             0            0.0373
#> 6 TCGA-5L-AAT0-01 BRCA     0.00674      0.0102        0            0.0845
#> # ... with 18 more variables: T_cell_CD4._naive <dbl>,
#> #   T_cell_CD4._memory_resting <dbl>, T_cell_CD4._memory_activated <dbl>,
#> #   T_cell_follicular_helper <dbl>, T_cell_regulatory_.Tregs. <dbl>,
#> #   T_cell_gamma_delta <dbl>, NK_cell_resting <dbl>, NK_cell_activated <dbl>,
#> #   Monocyte <dbl>, Macrophage_M0 <dbl>, Macrophage_M1 <dbl>,
#> #   Macrophage_M2 <dbl>, Myeloid_dendritic_cell_resting <dbl>,
#> #   Myeloid_dendritic_cell_activated <dbl>, Mast_cell_activated <dbl>,
#> #   Mast_cell_resting <dbl>, Eosinophil <dbl>, Neutrophil <dbl>
```
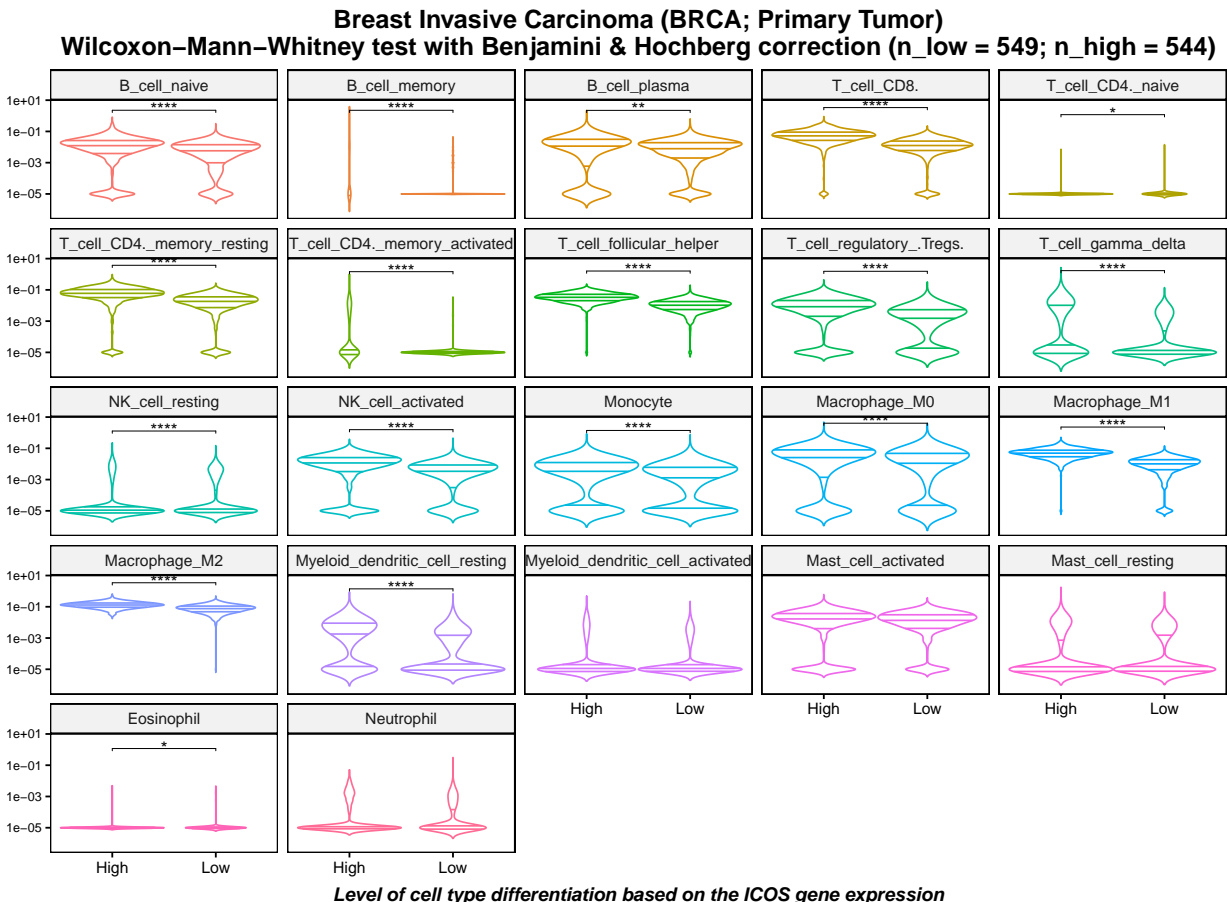
Violin plot of cell subtypes   And perform a significance of a Wilcoxon adjusted test according to the expression level (high or low) of a selected gene.

```
df <- convert2biodata(
  algorithm = "Cibersort_ABS",
  disease = "breast invasive carcinoma",
  tissue = "Primary Tumor",
  gene_x = "ICOS"
)
(stats <- calculate_pvalue(df))
#> Breast Invasive Carcinoma (BRCA; Primary Tumor)
#> Wilcoxon-Mann-Whitney test with Benjamini & Hochberg correction (n_low = 549; n_high = 544).
```

```
#> # A tibble: 18 x 9
#>    `Cell type`            `Average(High)` `Average(Low)` `SD(High)` `SD(Low)`
#>    <fct>                            <dbl>          <dbl>      <dbl>     <dbl>
#>  1 B_cell_memory                  0.00823       0.000966    0.0264    0.00354
#>  2 B_cell_naive                   0.0204        0.00917     0.0293    0.0100
#>  3 B_cell_plasma                  0.0210        0.0134      0.0274    0.0171
#>  4 Eosinophil                     0.0000220     0.0000344   0.000180  0.000203
#>  5 Macrophage_M0                  0.0522        0.0297      0.0660    0.0378
#>  6 Macrophage_M1                  0.0584        0.0130      0.0410    0.0113
#>  7 Macrophage_M2                  0.140         0.0846      0.0642    0.0476
#>  8 Monocyte                       0.00832       0.00462     0.0114    0.00792
#>  9 Myeloid_dendritic_cell_r~      0.00707       0.00213     0.0121    0.00639
#> 10 NK_cell_activated              0.0179        0.00562     0.0200    0.00681
#> 11 NK_cell_resting                0.00104       0.00115     0.00391   0.00276
#> 12 T_cell_CD4._memory_activ~      0.00400       0.0000863   0.0107    0.000907
#> 13 T_cell_CD4._memory_resti~      0.0782        0.0249      0.0678    0.0231
#> 14 T_cell_CD4._naive              0.0000219     0.0000838   0.000237  0.000657
#> 15 T_cell_CD8.                    0.0683        0.0177      0.0594    0.0169
#> 16 T_cell_follicular_helper       0.0395        0.0133      0.0252    0.0110
#> 17 T_cell_gamma_delta             0.00947       0.00117     0.0189    0.00288
#> 18 T_cell_regulatory_.Tregs.      0.0152        0.00379     0.0190    0.00556
#> # ... with 4 more variables: Average(High - Low) <dbl>, P-value <dbl>,
#> #   P-value adjusted <dbl>, Significance <chr>
plot(df, stats = stats)
```


Breast Invasive Carcinoma (BRCA; Primary Tumor)
Wilcoxon–Mann–Whitney test with Benjamini & Hochberg correction (n_low = 549; n_high = 544)

Level of cell type differentiation based on the ICOS gene expression

Advanced parameters

With ggplot2::theme() expressions.

```
(df <- convert2biodata(
  algorithm = "Cibersort_ABS",
  disease = "breast invasive carcinoma",
  tissue = "Primary Tumor",
  gene_x = "ICOS",
  stat = "quantile"
))
#> # A tibble: 11,990 x 3
#>    high  cell_type     value
#>  * <fct> <fct>         <dbl>
#>  1 25%   B_cell_naive 0.00001
#>  2 75%   B_cell_naive 0.00853
#>  3 25%   B_cell_naive 0.00001
#>  4 25%   B_cell_naive 0.00001
#>  5 25%   B_cell_naive 0.0196
#>  6 25%   B_cell_naive 0.0443
#>  7 25%   B_cell_naive 0.0118
#>  8 75%   B_cell_naive 0.00529
#>  9 25%   B_cell_naive 0.0236
#> 10 25%   B_cell_naive 0.00001
#> # ... with 11,980 more rows
(stats <- calculate_pvalue(
  df,
  method_test = "t_test",
  method_adjust = "bonferroni",
  p_threshold = 0.01
))
#> Breast Invasive Carcinoma (BRCA; Primary Tumor)
#> Student's t-test with bonferroni correction (n_low = 273; n_high = 272).
#> # A tibble: 16 x 9
#>    `Cell type`              `Average(75%)` `Average(25%)` `SD(75%)` `SD(25%)`
#>    <fct>                          <dbl>          <dbl>     <dbl>     <dbl>
#>  1 B_cell_memory                 0.0135        0.000924    0.0355   0.00339
#>  2 B_cell_naive                  0.0260        0.00710     0.0376   0.00792
#>  3 B_cell_plasma                 0.0236        0.0109      0.0294   0.0138
#>  4 Macrophage_M0                 0.0590        0.0273      0.0742   0.0348
#>  5 Macrophage_M1                 0.0811        0.00689     0.0443   0.00629
#>  6 Macrophage_M2                 0.153         0.0664      0.0667   0.0384
#>  7 Mast_cell_activated           0.0268        0.0175      0.0285   0.0192
#>  8 Monocyte                      0.00924       0.00360     0.0124   0.00566
#>  9 Myeloid_dendritic_cell_res~   0.00867       0.00145     0.0134   0.00514
#> 10 NK_cell_activated             0.0247        0.00415     0.0240   0.00469
#> 11 T_cell_CD4._memory_activat~   0.00724       0.0000449   0.0141   0.000300
#> 12 T_cell_CD4._memory_resting    0.103         0.0163      0.0802   0.0143
#> 13 T_cell_CD8.                   0.0945        0.0111      0.0699   0.0107
#> 14 T_cell_follicular_helper      0.0520        0.00963     0.0266   0.00781
#> 15 T_cell_gamma_delta            0.0142        0.000706    0.0244   0.00178
#> 16 T_cell_regulatory_.Tregs.     0.0211        0.00258     0.0233   0.00330
#> # ... with 4 more variables: Average(75% - 25%) <dbl>, P-value <dbl>,
#> #   P-value adjusted <dbl>, Significance <chr>
plot(
```

```
  df,
  stats = stats,
  type = "boxplot",
  dots = TRUE,
  xlab = "Expression level of the 'ICOS' gene by cell type",
  ylab = "Percent of relative abundance\n(from the Cibersort_ABS algorithm)",
  title = toupper("Differential analysis of immune cell type abundance
    based on RNASeq gene-level expression from The Cancer Genome Atlas"),
  axis.text.y = element_text(size = 8, hjust = 0.5),
  plot.title =  element_text(face = "bold", hjust = 0.5),
  plot.subtitle =  element_text(size = , face = "italic", hjust = 0.5),
  draw = FALSE
) + labs(
  subtitle = paste("Breast Invasive Carcinoma (BRCA; Primary Tumor):",
                   "Student's t-test with Bonferroni (P < 0.01)")
)
```


DIFFERENTIAL ANALYSIS OF IMMUNE CELL TYPE ABUNDANCE
BASED ON RNASEQ GENE–LEVEL EXPRESSION FROM THE CANCER GENOME ATLAS
Breast Invasive Carcinoma (BRCA; Primary Tumor): Student's t–test with Bonferroni (P < 0.01)

Expression level of the 'ICOS' gene by cell type
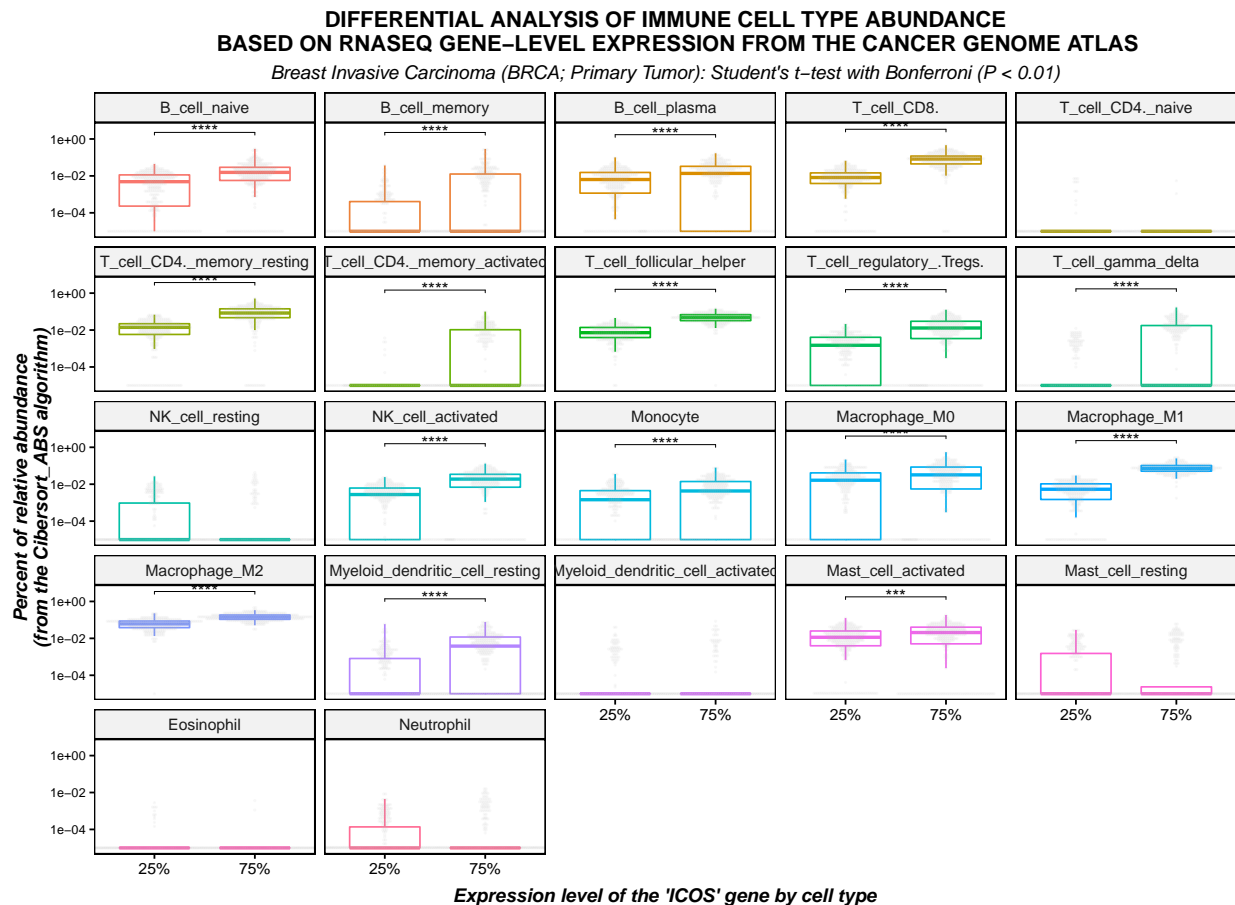
```
sessionInfo()
#> R version 4.0.5 (2021-03-31)
#> Platform: x86_64-pc-linux-gnu (64-bit)
#> Running under: Ubuntu 20.04.4 LTS
#>
#> Matrix products: default
#> BLAS/LAPACK: /usr/lib/x86_64-linux-gnu/openblas-pthread/libopenblasp-r0.3.8.so
#>
```

```
#> locale:
#>  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
#>  [3] LC_TIME=en_US.UTF-8        LC_COLLATE=en_US.UTF-8
#>  [5] LC_MONETARY=en_US.UTF-8    LC_MESSAGES=en_US.UTF-8
#>  [7] LC_PAPER=en_US.UTF-8       LC_NAME=C
#>  [9] LC_ADDRESS=C               LC_TELEPHONE=C
#> [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
#>
#> attached base packages:
#> [1] stats     graphics  grDevices utils     datasets  methods   base
#>
#> other attached packages:
#> [1] ggplot2_3.3.3 tcgaViz_0.6.0
#>
#> loaded via a namespace (and not attached):
#>   [1] fs_1.5.0           usethis_2.0.1      devtools_2.4.1
#>   [4] httr_1.4.2         rprojroot_2.0.2    tools_4.0.5
#>   [7] backports_1.2.1    utf8_1.2.1         R6_2.5.0
#>  [10] DT_0.18            lazyeval_0.2.2     colorspace_2.0-1
#>  [13] withr_2.4.2        tidyselect_1.1.1   prettyunits_1.1.1
#>  [16] processx_3.5.2     curl_4.3.1         compiler_4.0.5
#>  [19] cli_2.5.0          shinyjs_2.0.0      xml2_1.3.2
#>  [22] desc_1.3.0         plotly_4.9.3       scales_1.1.1
#>  [25] readr_1.4.0        callr_3.7.0        shinyFeedback_0.3.0
#>  [28] stringr_1.4.0      digest_0.6.27      foreign_0.8-81
#>  [31] rmarkdown_2.8      rio_0.5.26         pkgconfig_2.0.3
#>  [34] htmltools_0.5.1.1  sessioninfo_1.1.1  attempt_0.3.1
#>  [37] highr_0.9          fastmap_1.1.0      htmlwidgets_1.5.3
#>  [40] rlang_0.4.11       readxl_1.3.1       rstudioapi_0.13
#>  [43] shiny_1.6.0        farver_2.1.0       generics_0.1.0
#>  [46] jsonlite_1.7.2     dplyr_1.0.6        zip_2.1.1
#>  [49] car_3.0-10         config_0.3.1       magrittr_2.0.1
#>  [52] Rcpp_1.0.6         munsell_0.5.0      fansi_0.4.2
#>  [55] abind_1.4-5        lifecycle_1.0.0    stringi_1.6.1
#>  [58] yaml_2.2.1         carData_3.0-4      plyr_1.8.6
#>  [61] pkgbuild_1.2.0     grid_4.0.5         promises_1.2.0.1
#>  [64] forcats_0.5.1      crayon_1.4.1       haven_2.4.1
#>  [67] hms_1.0.0          knitr_1.33         ps_1.6.0
#>  [70] pillar_1.6.1       ggpubr_0.4.0       ggsignif_0.6.1
#>  [73] reshape2_1.4.4     pkgload_1.2.1      glue_1.4.2
#>  [76] evaluate_0.14      golem_0.3.1        data.table_1.14.0
#>  [79] remotes_2.3.0      vctrs_0.3.8        httpuv_1.6.1
#>  [82] testthat_3.0.2     cellranger_1.1.0   gtable_0.3.0
#>  [85] purrr_0.3.4        tidyr_1.1.3        cachem_1.0.5
#>  [88] xfun_0.23          openxlsx_4.2.3     mime_0.10
#>  [91] xtable_1.8-4       broom_0.7.6        roxygen2_7.1.1
#>  [94] rstatix_0.7.0      later_1.2.0        viridisLite_0.4.0
#>  [97] dockerfiler_0.1.3  tibble_3.1.2       memoise_2.0.0
#> [100] ellipsis_0.3.2
```