

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$genes)
#> # A tibble: 6 x 2
#>   sample      ICOS
#>   <chr>      <dbl>
#> 1 TCGA-3C-AAAU-01 1.25
#> 2 TCGA-A2-A04Q-01 7.79
#> 3 TCGA-A2-A0T4-01 4.97
#> 4 TCGA-A8-A08S-01 3.69
#> 5 TCGA-A8-A09B-01 2.55
#> 6 TCGA-A8-A0AD-01 3.72
head(tcga$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-A2-A04Q-01 BRCA      0.0274      0.0249      0.0236      0.118
#> 3 TCGA-A2-A0T4-01 BRCA      0.0167      0          0.0159      0.0432
#> 4 TCGA-A8-A08S-01 BRCA      0          0.00425     0          0.0217
#> 5 TCGA-A8-A09B-01 BRCA      0.0146      0          0.00612     0.0256
#> 6 TCGA-A8-A0AD-01 BRCA      0.000919    0.000797    0.00290     0
#> # ... 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>
```

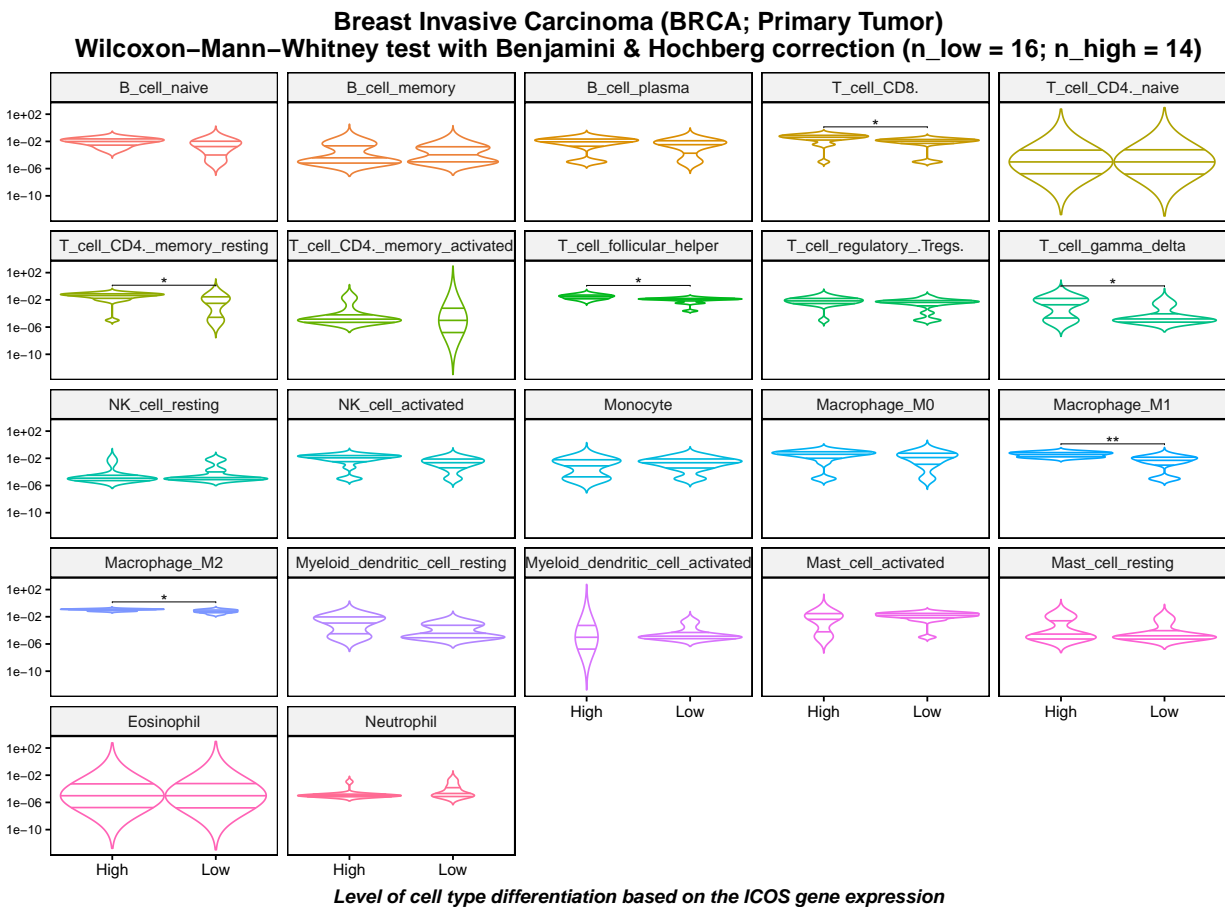
Violin plot of cell subtypes And perform a significance of a Wilcoxon adjusted test according to the expression level (high or low) of a selected gene.

```
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 = 16; n_high = 14).
```

```

#> # A tibble: 6 x 9
#>   `Cell type`      `Average(High)` `Average(Low)` `SD(High)` `SD(Low)`
#>   <fct>          <dbl>         <dbl>         <dbl>      <dbl>
#> 1 Macrophage_M1      0.0454         0.00943       0.0328     0.0116
#> 2 Macrophage_M2      0.109          0.0697       0.0321     0.0368
#> 3 T_cell_CD4._memory_resting 0.0504         0.0122       0.0377     0.0124
#> 4 T_cell_CD8.        0.0498         0.0127       0.0387     0.00934
#> 5 T_cell_follicular_helper 0.0352         0.0119       0.0259     0.00691
#> 6 T_cell_gamma_delta 0.00823        0.000956     0.0101     0.00258
#> # ... with 4 more variables: Average(High - Low) <dbl>, P-value <dbl>,
#> # P-value adjusted <dbl>, Significance <chr>
plot(df, stats = stats)

```



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: 352 x 3
#>   high cell_type      value

```

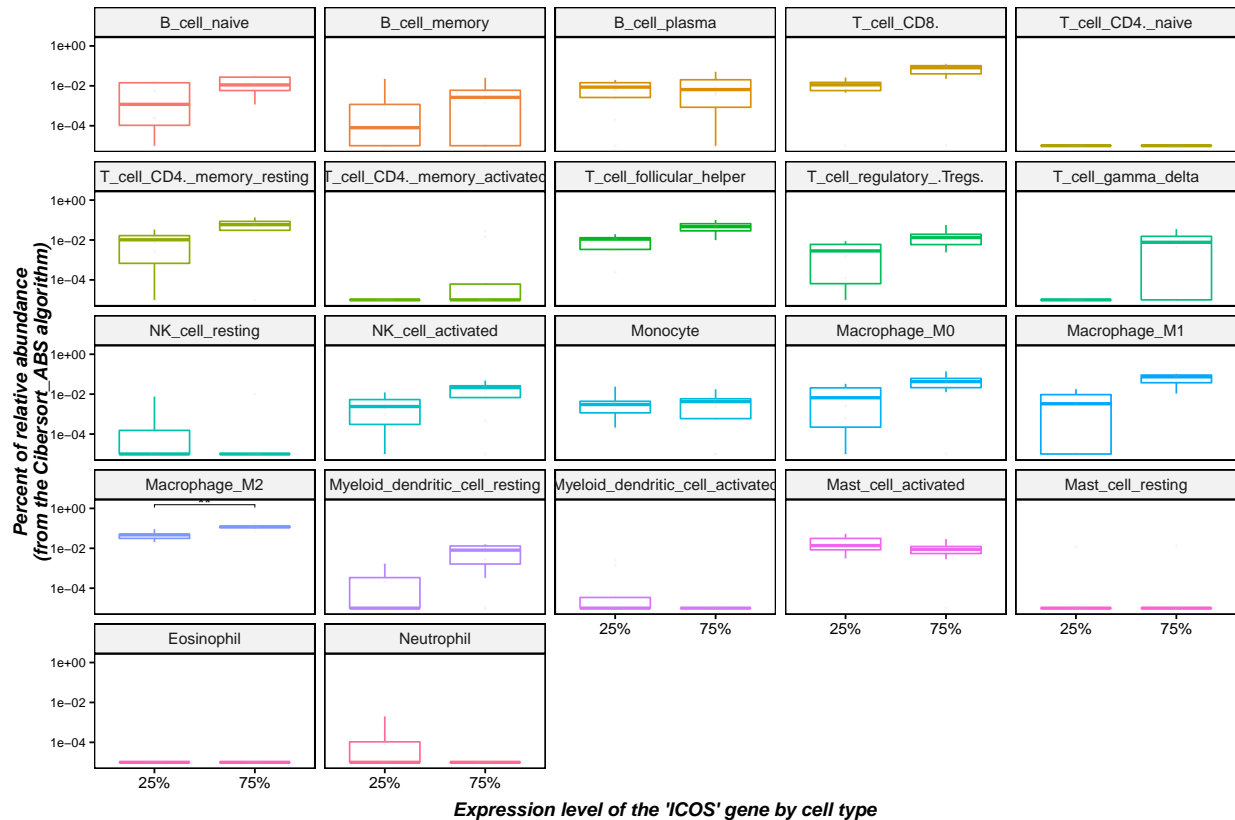
```

#> * <fct> <fct> <dbl>
#> 1 25% B_cell_naive 0.00001
#> 2 75% B_cell_naive 0.0274
#> 3 25% B_cell_naive 0.0146
#> 4 75% B_cell_naive 0.0112
#> 5 25% B_cell_naive 0.0141
#> 6 25% B_cell_naive 0.00546
#> 7 75% B_cell_naive 0.0289
#> 8 75% B_cell_naive 0.00376
#> 9 25% B_cell_naive 0.00001
#> 10 75% B_cell_naive 0.00118
#> # ... with 342 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 = 8; n_high = 8).
#> # A tibble: 1 x 9
#> `Cell type` `Average(75%)` `Average(25%)` `SD(75%)` `SD(25%)` `Average(75% - ~
#> <fct> <dbl> <dbl> <dbl> <dbl> <dbl>
#> 1 Macrophage~ 0.117 0.0456 0.0274 0.0216 0.0719
#> # ... with 3 more variables: 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 ($P < 0.01$)



Session information

```
#> 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/libopenblas-p0.3.8.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_1.0.0
#>
#> 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

```