Annotation cells

Joan Abinet

2024-12-13 13:30:25 +0100

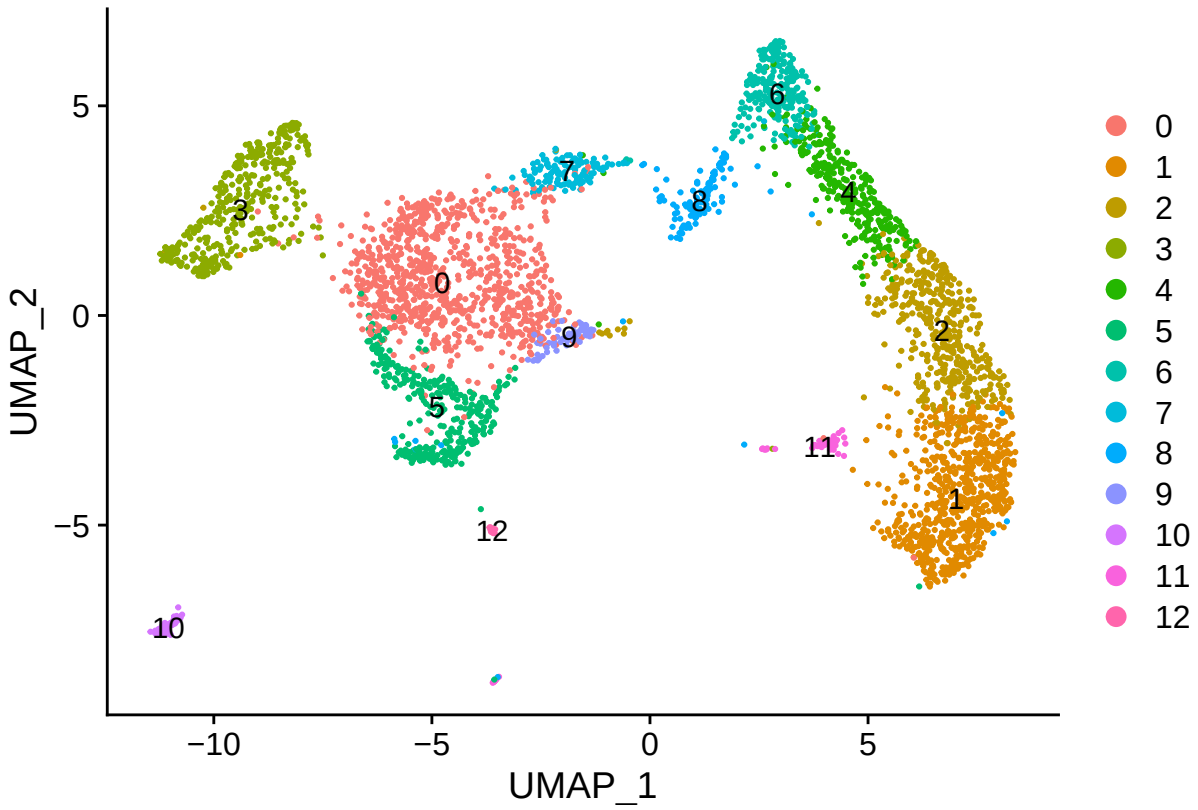
Contents

Load Packages

```
suppressMessages(library(dplyr))
suppressMessages(library(Seurat))
suppressMessages(library(patchwork))
suppressMessages(library(ggplot2))
suppressMessages(library(stringr))
suppressMessages(library(dittoSeq))
suppressMessages(library(ComplexHeatmap))
```

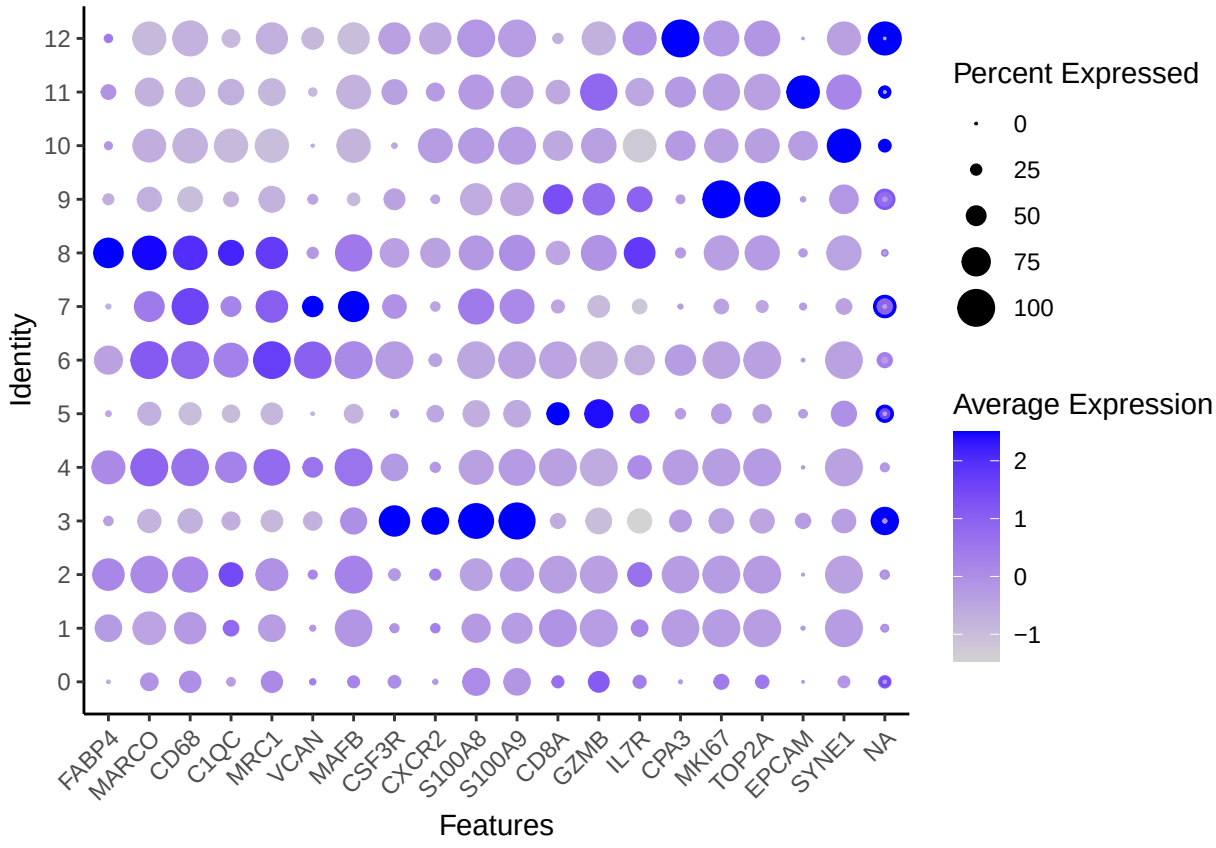
Load Seurat objects

```
BAL_Hive.integrated <- readRDS("../Hive-1-Pre_Processing/Hive_integrated_noDB.rds")

DimPlot(BAL_Hive.integrated, label = T)
```



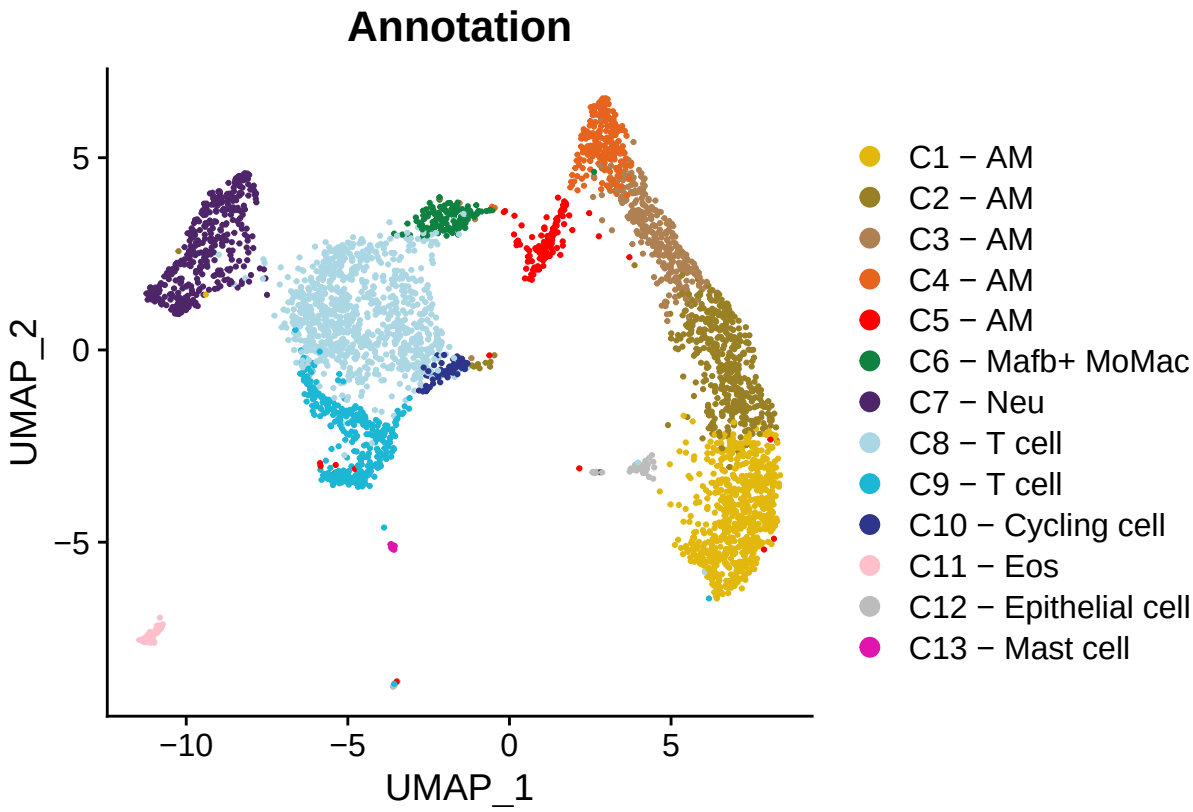
Annotating clusters

```
markers_neutro <- c("CSF3R", "FCGR3B", "CXCR2", "S100A8",
    "S100A9")
marker_epi <- c("EPCAM", "KRT8", "CDH1", "SIGLEC8")
Marker_Cycling_cell <- c("MKI67", "TOP2A")
markers_AM <- c("FABP4", "MARCO", "CD68", "C1QC", "MRC1")
markers_Mafb_macro <- c("CSF1R", "VCAN", "CD14", "MAFB")
markers_Mast <- c("KIT", "CPA3")
markers_LT <- c("CD3E", "CD8A", "GZMB", "KLGR1", "KLRK1",
    "CD4", "IL7R")
markers_eos <- c("CCR3", "SYNE1")

DotPlot(BAL_Hive.integrated, features = c(markers_AM,
    markers_Mafb_macro, markers_neutro, markers_LT,
    markers_Mast, Marker_Cycling_cell, marker_epi,
    markers_eos)) + theme_classic() + theme(axis.text.x = element_text(angle = 45,
    hjust = 1))
```

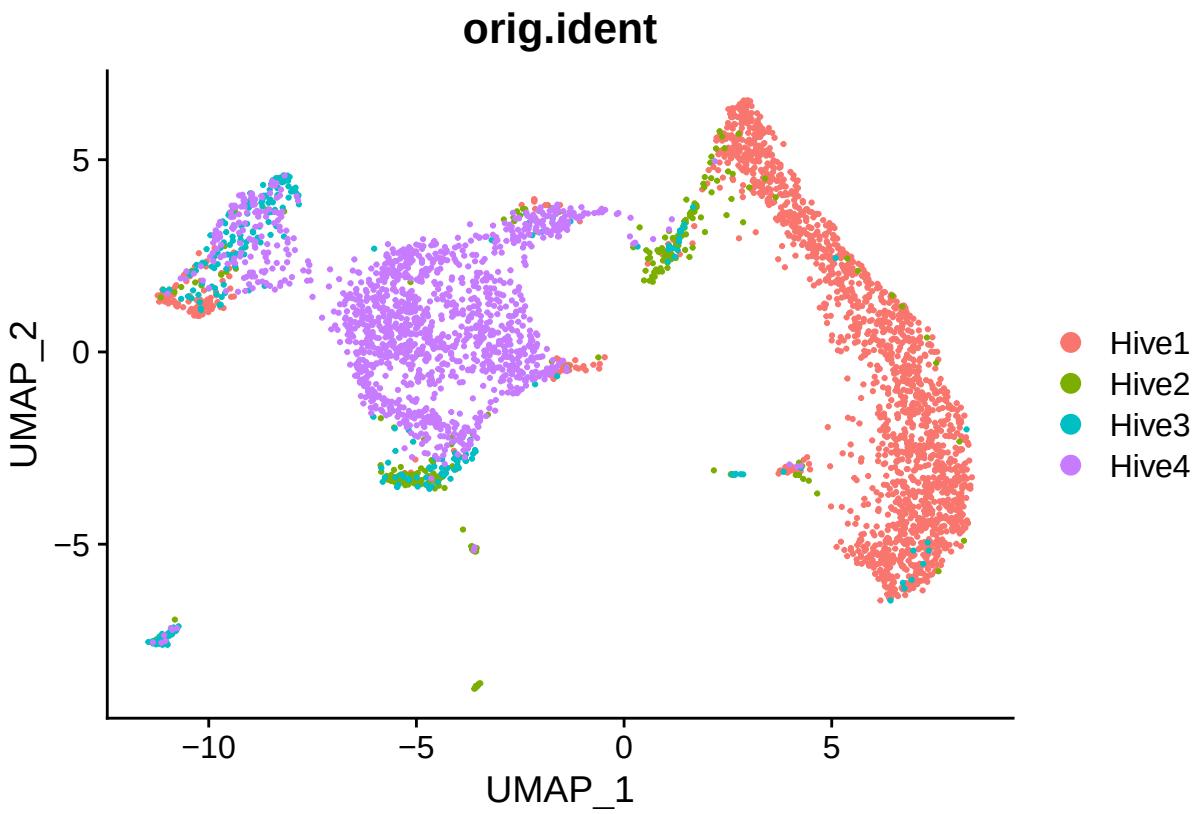


Annotation

```
BAL_Hive.integrated$Annotation <- BAL_Hive.integrated$seurat_clusters
BAL_Hive.integrated$Annotation <- as.factor(BAL_Hive.integrated$Annotation)
levels(BAL_Hive.integrated$Annotation) <- c("C8 - T cell",
    "C1 - AM", "C2 - AM", "C7 - Neu", "C3 - AM", "C9 - T cell",
    "C4 - AM", "C6 - Mafb+ MoMac", "C5 - AM", "C10 - Cycling cell",
    "C11 - Eos", "C12 - Epithelial cell", "C13 - Mast cell")

BAL_Hive.integrated$Annotation <- factor(BAL_Hive.integrated$Annotation,
    levels = c("C1 - AM", "C2 - AM", "C3 - AM", "C4 - AM",
        "C5 - AM", "C6 - Mafb+ MoMac", "C7 - Neu",
        "C8 - T cell", "C9 - T cell", "C10 - Cycling cell",
        "C11 - Eos", "C12 - Epithelial cell", "C13 - Mast cell"))

col <- c("#E2B80C", "#998025", "#AE8052", "#e6641e",
    "red", "#0F8140", "#4F2569", "#ABD6E4", "#1CB7D5",
    "#2D368B", "#FFC0CB", "#BDBCBC", "#DF15AE")

DimPlot(BAL_Hive.integrated, group.by = "Annotation",
    cols = col)
```



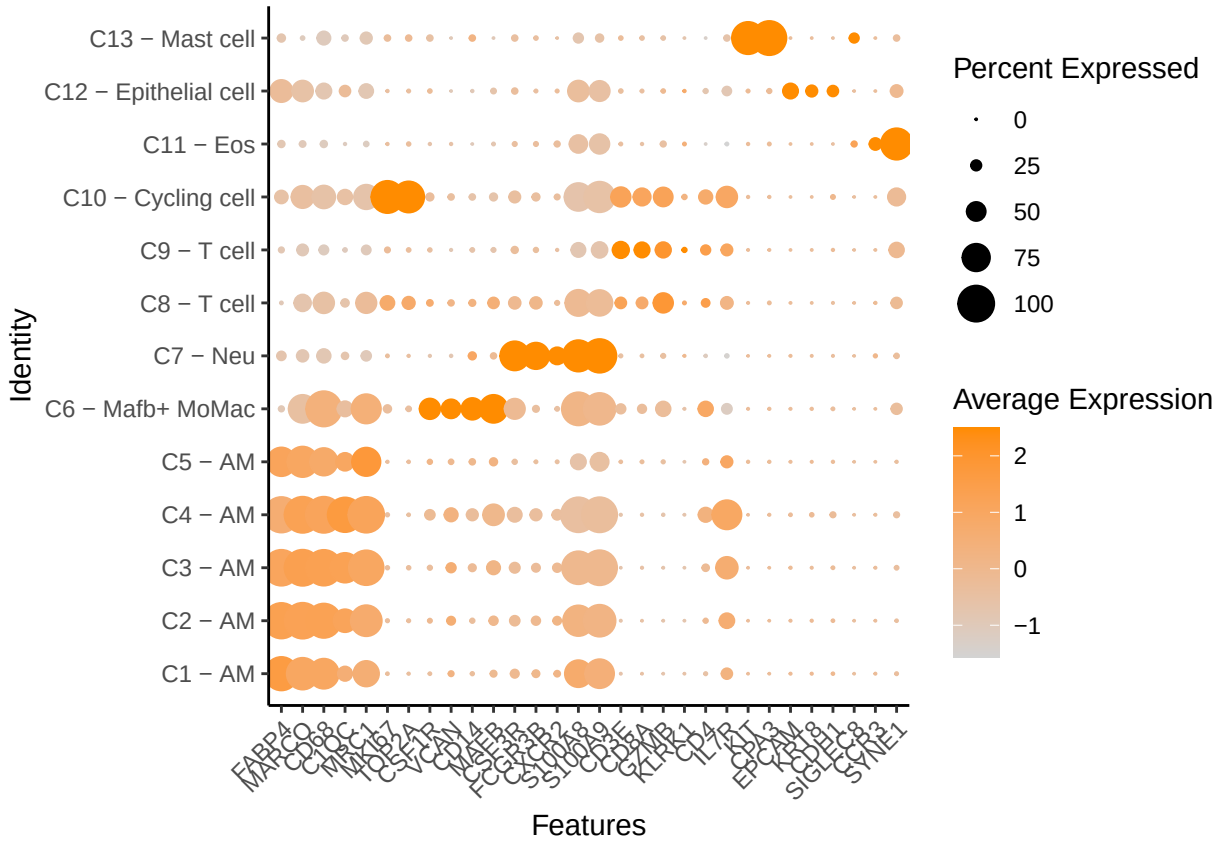

```
# ggsave('Umap_Annotation.pdf', width = 8, height
# = 4)
DimPlot(BAL_Hive.integrated, group.by = "orig.ident")
```




```
# ggsave('Umap_PerSample.pdf', width = 8, height
# = 4)
```

Dotplot Annotated

```
DefaultAssay(BAL_Hive.integrated) <- "RNA"
Idents(BAL_Hive.integrated) <- "Annotation"
Marker_gene <- c("FABP4", "MARCO", "CD68", "C1QC",
    "MRC1", "MKI67", "TOP2A", "CSF1R", "VCAN", "CD14",
    "MAFB", "CSF3R", "FCGR3B", "CXCR2", "S100A8", "S100A9",
    "CD3E", "CD8A", "GZMB", "KLGR1", "KLRK1", "CD4",
    "IL7R", "KIT", "CPA3", "EPCAM", "KRT8", "CDH1",
    "SIGLEC8", "CCR3", "SYNE1")

DotPlot(BAL_Hive.integrated, features = Marker_gene,
    cols = c("lightgray", "darkorange"), ) + theme_classic() +
    theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



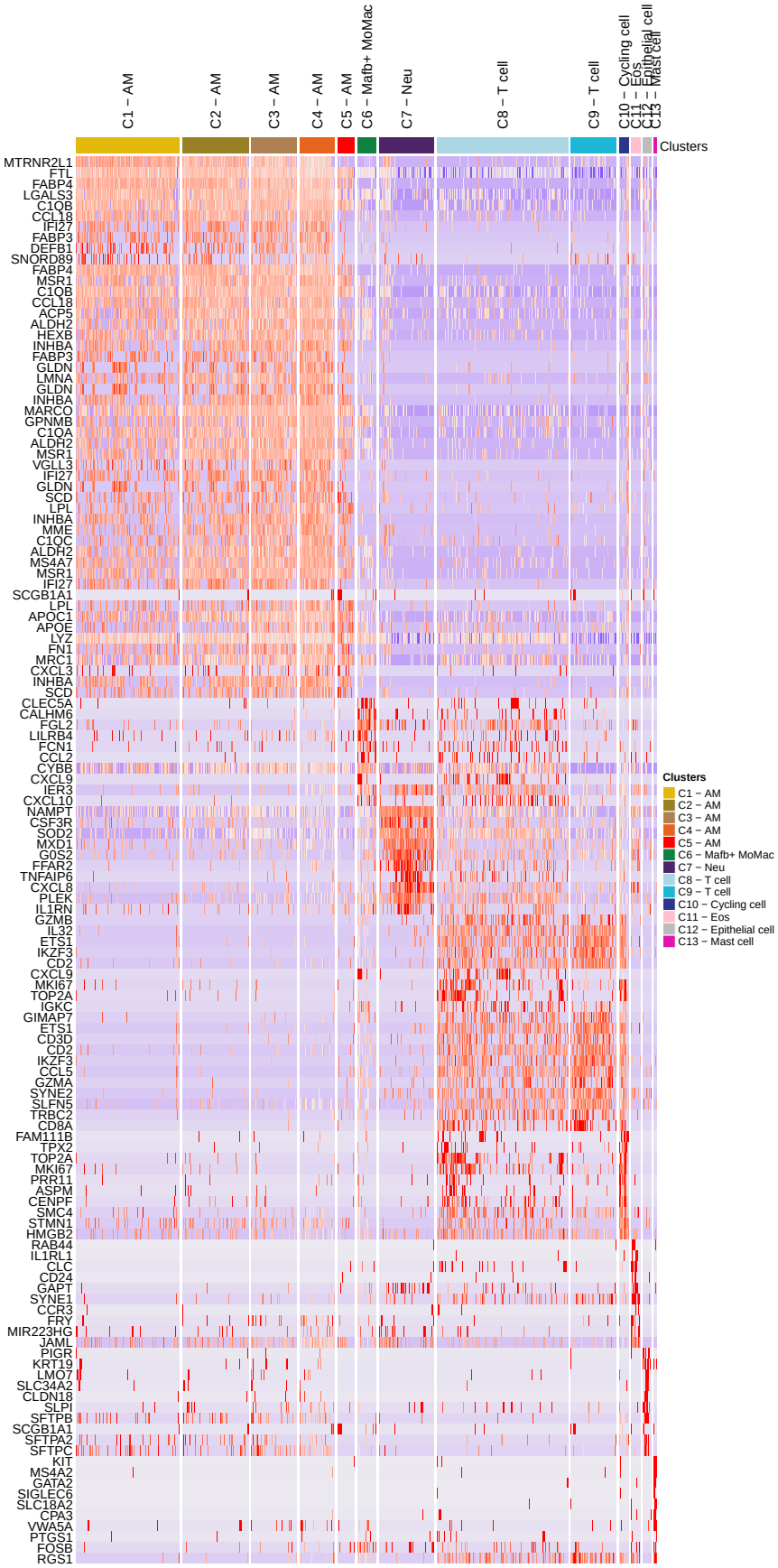

```
# ggsave('Dotplot_Non-annotated.pdf', width = 35,
# height = 10)
```

```
lavage.markers <- FindAllMarkers(BAL_Hive.integrated,
    only.pos = TRUE, min.pct = 0.25)

lavage.markers %>%
    group_by(cluster) %>%
    top_n(n = 10, wt = avg_log2FC) -> top10

mat <- as.matrix(GetAssayData(object = BAL_Hive.integrated,
    slot = "data")[as.character(top10$gene), ])

df <- as.data.frame(BAL_Hive.integrated$Annotation)
colnames(df) <- "Clusters"
color_df <- list(Clusters = c(`C1 - AM` = "#E2B80C",
    `C2 - AM` = "#998025", `C3 - AM` = "#AE8052", `C4 - AM` = "#e6641e",
    `C5 - AM` = "red", `C6 - Mafb+ MoMac` = "#0F8140",
    `C7 - Neu` = "#4F2569", `C8 - T cell` = "#ABD6E4",
    `C9 - T cell` = "#1CB7D5", `C10 - Cycling cell` = "#2D368B",
    `C11 - Eos` = "#FFC0CB", `C12 - Epithelial cell` = "#BDBCBC",
    `C13 - Mast cell` = "#DF15AE"))

Heatmap <- Heatmap(t(scale(t(mat))), show_column_names = F,
    column_split = BAL_Hive.integrated$Annotation,
    cluster_column_slices = F, cluster_rows = F, top_annotation = HeatmapAnnotation(df = df,
        col = color_df), use_raster = F, show_heatmap_legend = F,
```

```
    show_column_dend = F, column_title_rot = 90, row_names_side = "left")

# tidyHeatmap::save_pdf(Heatmap,
# 'Heatmap_Hive.pdf',width = 30, height = 45,
# units ='cm')
```

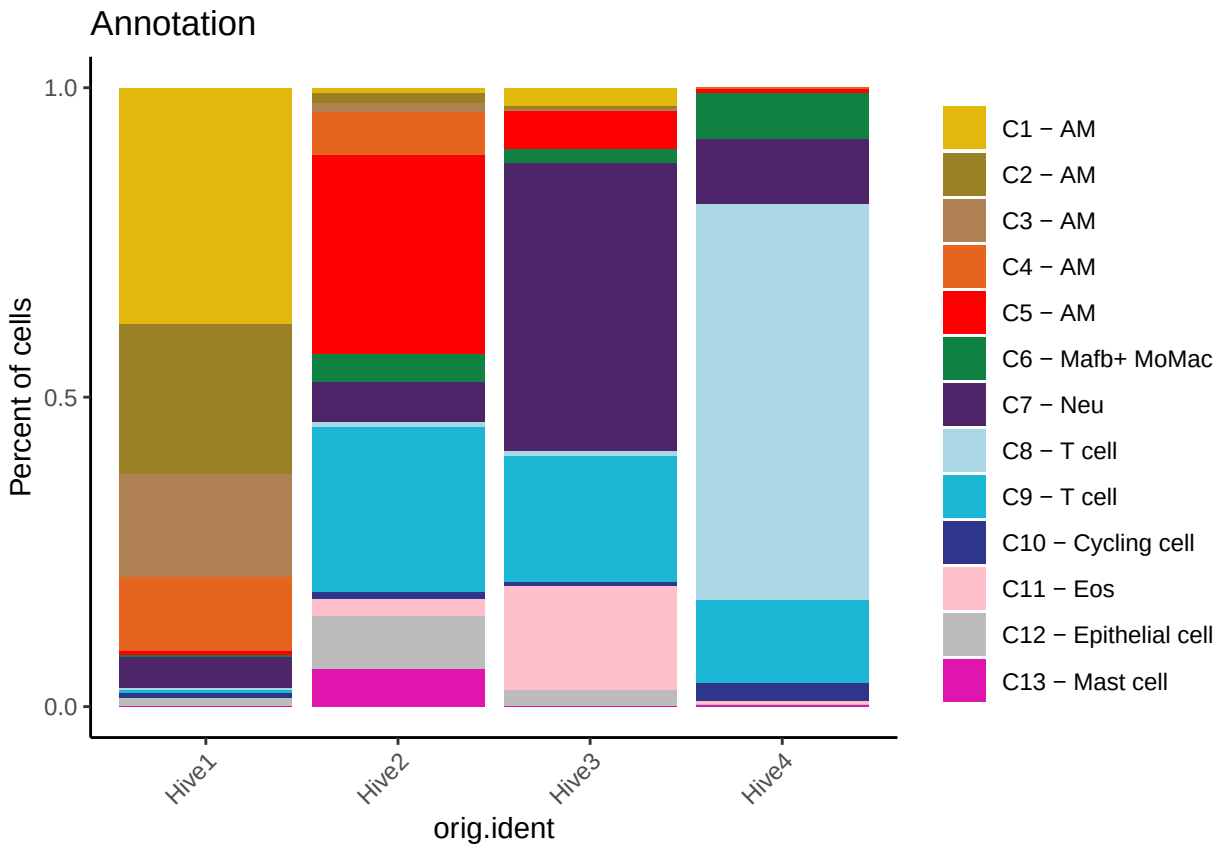
Heatmap top genes per clusters

```
Heatmap
```


C1 – AM
C2 – AM
C3 – AM
C4 – AM
C5 – AM
C6 – Mafb+ MoMac
C7 – Neu
C8 – T cell
C9 – T cell
C10 – Cycling cell
C11 – Eos
C12 – Epithelial cell
C13 – Mast cell
Clusters
MTRNR2L1
FTL
FABP4
LGALS3
C1QB
CCL18
IFI27
FABP3
DEFB1
SNORD89
FABP4
MSR1
C1QB
CCL18
ACP5
ALDH2
HEXB
INHBA
FABP3
GLDN
LMNA
GLDN
INHBA
MARCO
GPNMB
C1QA
ALDH2
MSR1
VGLL3
IFI27
GLDN
SCD
LPL
INHBA
MME
C1QC
ALDH2
MS4A7
MSR1
IFI27
SCGB1A1
LPL
APOC1
APOE
LYZ
FN1
MRC1
CXCL3
INHBA
SCD
CLEC5A
CALHM6
FGL2
LILRB4
FCN1
CCL2
CYBB
CXCL9
IER3
CXCL10
NAMPT
CSF3R
SOD2
MXD1
G0S2
FFAR2
TNFAIP6
CXCL8
PLEK
IL1RN
GZMB
IL32
ETS1
IKZF3
CD2
CXCL9
MKI67
TOP2A
IGKC
GIMAP7
ETS1
CD3D
CD2
IKZF3
CCL5
GZMA
SYNE2
SLFN5
TRBC2
CD8A
FAM111B
TPX2
TOP2A
MKI67
PRR11
ASPM
CENPF
SMC4
STMN1
HMGB2
RAB44
IL1RL1
CLC
CD24
GAPT
SYNE1
CCR3
FRY
MIR223HG
JAML
PIGR
KRT19
LMO7
SLC34A2
CLDN18
SLPI
SFTPB
SCGB1A1
SFTPA2
SFTPC
KIT
MS4A2
GATA2
SIGLEC6
SLC18A2
CPA3
VWA5A
PTGS1
FOSB
RGS1
Clusters
C1 – AM
C2 – AM
C3 – AM
C4 – AM
C5 – AM
C6 – Mafb+ MoMac
C7 – Neu
C8 – T cell
C9 – T cell
C10 – Cycling cell
C11 – Eos
C12 – Epithelial cell
C13 – Mast cell

Cluster frequency per sample

```
var_order <- c(1, 6, 7, 8, 9, 10, 11, 12, 13, 2, 3,
    4, 5)
dittoBarPlot(BAL_Hive.integrated, "Annotation", group.by = "orig.ident",
    var.labels.reorder = var_order, color.panel = col)
```




```
# ggsave('bar_freq_Hive.pdf')
```

Saving results for later

```
saveRDS(BAL_Hive.integrated, "BAL_hive.annotated_noDB.rds")
```

```
sessionInfo()
```

```
## R version 4.3.3 (2024-02-29)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 22.04.4 LTS
##
## Matrix products: default
## BLAS:   /usr/lib/x86_64-linux-gnu/openblas-pthread/libblas.so.3
## LAPACK: /usr/lib/x86_64-linux-gnu/openblas-pthread/libopenblasp-r0.3.20.so;  LAPACK version 3.10.0
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=fr_BE.UTF-8        LC_COLLATE=en_US.UTF-8
```

```
##  [5] LC_MONETARY=fr_BE.UTF-8    LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=fr_BE.UTF-8       LC_NAME=C
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
## [11] LC_MEASUREMENT=fr_BE.UTF-8 LC_IDENTIFICATION=C
##
## time zone: Europe/Brussels
## tzcode source: system (glibc)
##
## attached base packages:
## [1] grid      stats     graphics  grDevices utils     datasets  methods
## [8] base
##
## other attached packages:
## [1] ComplexHeatmap_2.16.0 dittoSeq_1.12.0       stringr_1.5.0
## [4] ggplot2_3.4.2         patchwork_1.1.2       SeuratObject_4.1.3
## [7] Seurat_4.3.0          dplyr_1.1.2
##
## loaded via a namespace (and not attached):
##   [1] RColorBrewer_1.1-3          shape_1.4.6
##   [3] rstudioapi_0.14             jsonlite_1.8.7
##   [5] magrittr_2.0.3              magick_2.7.5
##   [7] spatstat.utils_3.0-3        farver_2.1.1
##   [9] rmarkdown_2.23              GlobalOptions_0.1.2
##  [11] zlibbioc_1.46.0             vctrs_0.6.3
##  [13] ROCR_1.0-11                 Cairo_1.6-2
##  [15] spatstat.explore_3.2-1      RCurl_1.98-1.12
##  [17] S4Arrays_1.2.1              htmltools_0.5.5
##  [19] sctransform_0.3.5           parallelly_1.36.0
##  [21] KernSmooth_2.23-22          htmlwidgets_1.6.2
##  [23] ica_1.0-3                   plyr_1.8.8
##  [25] plotly_4.10.2               zoo_1.8-12
##  [27] igraph_1.5.0.1              iterators_1.0.14
##  [29] mime_0.12                   lifecycle_1.0.3
##  [31] pkgconfig_2.0.3             Matrix_1.6-1
##  [33] R6_2.5.1                    fastmap_1.1.1
##  [35] clue_0.3-64                 GenomeInfoDbData_1.2.10
##  [37] MatrixGenerics_1.12.2       fitdistrplus_1.1-11
##  [39] future_1.33.0               shiny_1.7.4.1
##  [41] digest_0.6.33               colorspace_2.1-0
##  [43] S4Vectors_0.38.1            tensor_1.5
##  [45] irlba_2.3.5.1               GenomicRanges_1.52.0
##  [47] labeling_0.4.2              progressr_0.13.0
##  [49] fansi_1.0.4                 spatstat.sparse_3.0-2
##  [51] httr_1.4.6                  polyclip_1.10-4
##  [53] abind_1.4-5                 compiler_4.3.3
##  [55] doParallel_1.0.17           withr_2.5.0
##  [57] highr_0.10                  MASS_7.3-60.0.1
##  [59] DelayedArray_0.26.3         rjson_0.2.21
##  [61] tools_4.3.3                 lmtest_0.9-40
##  [63] httpuv_1.6.11               future.apply_1.11.0
##  [65] goftest_1.2-3               glue_1.6.2
##  [67] nlme_3.1-164                promises_1.2.0.1
##  [69] Rtsne_0.16                  cluster_2.1.6
##  [71] reshape2_1.4.4              generics_0.1.3
```

```
##  [73] gtable_0.3.3               spatstat.data_3.0-1
##  [75] tidyr_1.3.0                data.table_1.14.8
##  [77] XVector_0.40.0             sp_2.0-0
##  [79] utf8_1.2.3                 BiocGenerics_0.46.0
##  [81] spatstat.geom_3.2-4        RcppAnnoy_0.0.21
##  [83] foreach_1.5.2              ggrepel_0.9.3
##  [85] RANN_2.6.1                 pillar_1.9.0
##  [87] limma_3.56.2               spam_2.9-1
##  [89] later_1.3.1                circlize_0.4.15
##  [91] splines_4.3.3              lattice_0.22-5
##  [93] survival_3.5-8             deldir_1.0-9
##  [95] tidyselect_1.2.0           SingleCellExperiment_1.22.0
##  [97] miniUI_0.1.1.1             pbapply_1.7-2
##  [99] knitr_1.43                 gridExtra_2.3
## [101] IRanges_2.34.0             SummarizedExperiment_1.30.2
## [103] scattermore_1.2            stats4_4.3.3
## [105] xfun_0.39                  Biobase_2.60.0
## [107] matrixStats_1.0.0          pheatmap_1.0.12
## [109] stringi_1.7.12             lazyeval_0.2.2
## [111] yaml_2.3.7                 evaluate_0.21
## [113] codetools_0.2-19           tibble_3.2.1
## [115] cli_3.6.1                  uwot_0.1.16
## [117] xtable_1.8-4               reticulate_1.30
## [119] munsell_0.5.0              Rcpp_1.0.11
## [121] GenomeInfoDb_1.36.0        globals_0.16.2
## [123] spatstat.random_3.1-5      png_0.1-8
## [125] parallel_4.3.3             ellipsis_0.3.2
## [127] dotCall64_1.0-2            bitops_1.0-7
## [129] listenv_0.9.0              viridisLite_0.4.2
## [131] scales_1.2.1               ggridges_0.5.4
## [133] crayon_1.5.2               leiden_0.4.3
## [135] purrr_1.0.1                GetoptLong_1.0.5
## [137] rlang_1.1.1                cowplot_1.1.1
## [139] formatR_1.14
```