A lung Tgf-beta-signaling-mediated endothelial-interstitial macrophage axis prevents age-related abnormalities

2-Differential Expression between TGFbRII-KO and WT

Rendered 2024-12-17 13:42:49 +0100

Abstract

Lung interstitial macrophages (IMs) are monocyte-derived parenchymal macrophages whose homeostatic and tissue-supportive functions remain unclear. While recent progress has been made about the diversity and transcriptional regulation of lung IMs, the microenvironmental signals responsible for their development from monocytes and for their functional specification remain unidentified. Here we found, in mice, that lung endothelial cell-derived Tgf-beta1 specifically triggered a core Tgf-beta receptor-dependent IM signature in bone marrow-derived monocytes and macrophages (Macs). In vivo, myeloid-specific ablation of Tgf-beta receptor signaling severely impaired monocyte-to-IM development, resulting in the accumulation of perivascular monocytes, decreased IM numbers and a loss of IM-intrinsic identity. Of note, monocyte-to-IM development was similarly impaired in the absence of endothelial-specific Tgf-beta1. Functionally, lungs from mice selectively lacking Tgf-beta receptor in IMs exhibited spatial changes in monocyte and IM niche occupancies, a severe disruption in their immunoregulatory environment, and prematurely developed fibrosis, hyperinflation, increased compliance and decreased elastance, changes classically associated with aging. Our work identifies a novel endothelial-IM axis involving Tgf-beta1 - Tgf-beta receptor interactions that shapes IM development and identity and thereby sustains lung tissue integrity, thus providing foundations for IM-targeted interventions in the context of lung aging and other chronic inflammatory disorders.

Contents

1 Description

The differential expression (DE) analysis was made with `DESeq2` package *(1)*. The following packages were used to build plots: `ggplot2` *(2)*, `RColorBrewer` *(3)* and `EnhancedVolcano` *(4)*.

2 Prepare packages and data for DE expression analysis

```
suppressMessages(library(DESeq2))
suppressMessages(library(ggplot2))
suppressMessages(library(RColorBrewer))
suppressMessages(library(EnhancedVolcano))

cols <- brewer.pal(5, "Set2")
dds <- readRDS(file = "../1_Bulk_RNAseq_Data_preparation/dds.Rds")
```

Filter the genes which have at leat 10 reads in at least 4 samples:

```
smallestGroupSize <- 4
keep <- rowSums(counts(dds) >= 10) >= smallestGroupSize
dds <- dds[keep, ]
dds
```

```
## class: DESeqDataSet
## dim: 12824 8
## metadata(1): version
## assays(1): counts
## rownames(12824): Gnai3 Cdc45 ... Itprip Tmem179b
## rowData names(0):
## colnames(8): KO1 KO2 ... WT3 WT4
## colData names(2): genotype experiment
```

Set WT as reference group:

```
dds$genotype <- relevel(dds$genotype, ref = "WT")
```

3 Differentially expressed (DE) genes comparing TGFbRII-KO vs WT IMs

According to Michel Love (https://support.bioconductor.org/p/121408/#121419), just putting the batch in the design helps to control for changes in the counts due to batch. Putting batch in the design formula is our recommendation when batches are known and when they are not confounded with the condition. Thus the batch variation (**experiment** variant) has been already considered into design.

```
design(dds)
```

```
## ~experiment + genotype
```

Make DE analysis using default design:

```
dds <- DESeq(dds)
res <- results(dds)

summary(res)
```

```
##
## out of 12824 with nonzero total read count
## adjusted p-value < 0.1
## LFC > 0 (up)       : 1524, 12%
## LFC < 0 (down)     : 1507, 12%
## outliers [1]       : 0, 0%
## low counts [2]     : 0, 0%
## (mean count < 7)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
res_Shrunk <- lfcShrink(dds, res = res, type = "apeglm", coef = "genotype_
    TGFbIIR_KO_vs_WT")
# refilled_vs_control <- merge(x=as.data.frame(res_refilled_vs_control),
# y=as.data.frame(res_refilled_vs_control_Shrunk), by=c(0,1))
DE.res.KO_vs_WT <- merge(x = as.data.frame(res[1:4]), y = as.data.frame(
    res_Shrunk),
     by = c(0, 1), )

colnames(DE.res.KO_vs_WT) <- c("Row.names", "baseMean", "log2FoldChange",
    "lfcSE",
     "stat", "log2FoldChange.shrunk", "lfcSE.shrunk", "pvalue", "padj")
rownames(DE.res.KO_vs_WT) <- DE.res.KO_vs_WT$Row.names
DE.res.KO_vs_WT <- DE.res.KO_vs_WT[, -1]
DE.res.KO_vs_WT <- DE.res.KO_vs_WT[order(DE.res.KO_vs_WT$log2FoldChange),
    ]
```

Add normalized counts of samples:

```
normalized.counts <- counts(dds, normalized = TRUE)
DE.res.KO_vs_WT.nCounts <- merge(x = DE.res.KO_vs_WT, y = normalized.
    counts, by = "row.names")
```

Significant DE genes:

```
threshold.padj <- 0.05
threshold.lfc <- 1
sigDE.res.KO_vs_WT <- DE.res.KO_vs_WT[abs(DE.res.KO_vs_WT$log2FoldChange.
    shrunk) >
     threshold.lfc & DE.res.KO_vs_WT$padj < threshold.padj, ]
```

Number of significant KO up-regulated genes (logFC > 1)

```
sum(sigDE.res.KO_vs_WT$log2FoldChange > 1)
```

```
## [1] 152
```

Number of significant KO down-regulated genes (logFC < -1)

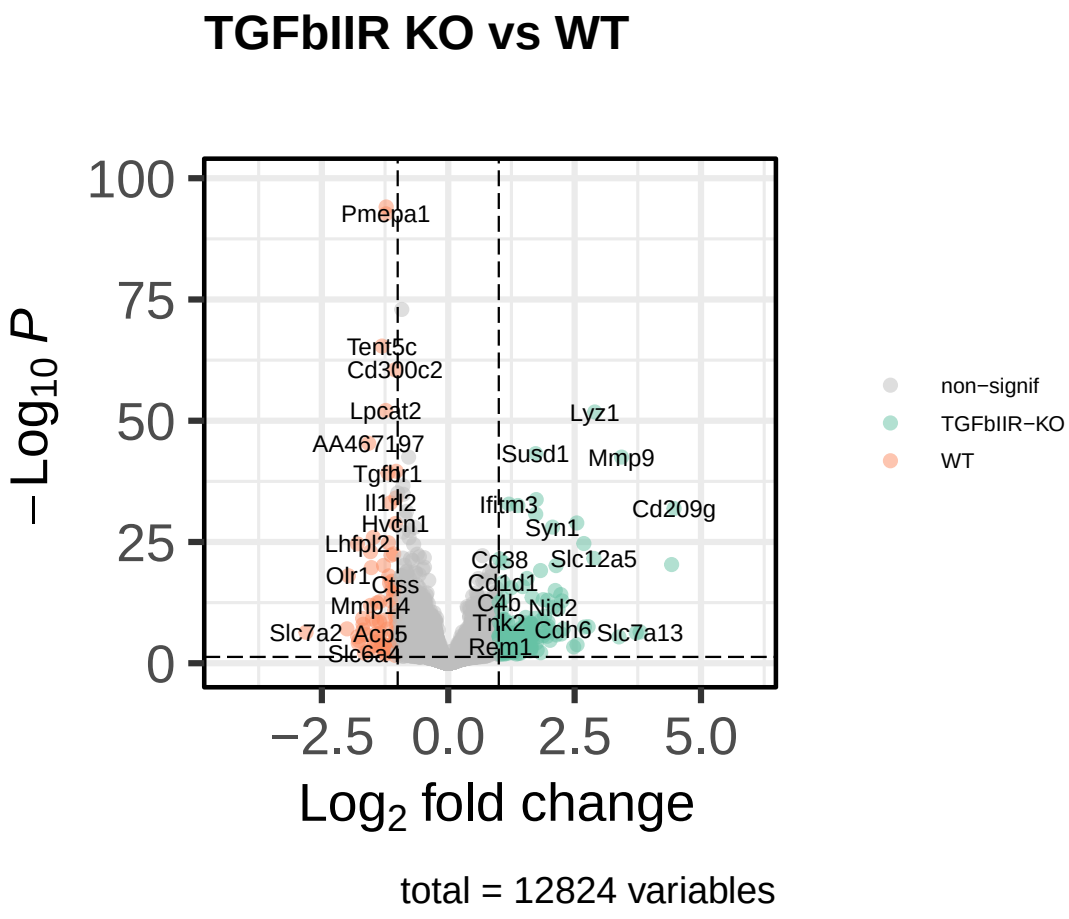
```
sum(sigDE.res.KO_vs_WT$log2FoldChange < (-1))
```

```
## [1] 79
```

4 Presentation of DE genes in volcano plots

4.1 TGFbIIR KO vs WT

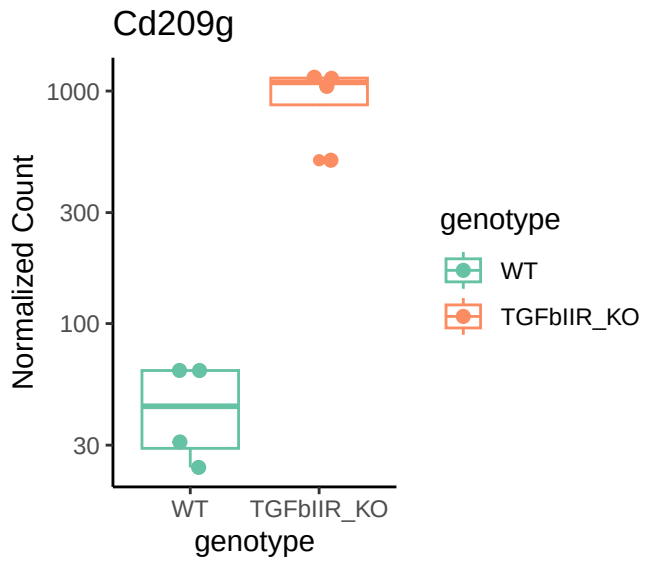
```
de.res <- DE.res.KO_vs_WT

de.res$Gene <- rownames(de.res)
keyvals <- rep("grey", nrow(de.res))
names(keyvals) <- rep("non-signif", nrow(de.res))

keyvals[which(de.res$log2FoldChange.shrunk > 1 )] <- cols[1]
names(keyvals)[which(de.res$log2FoldChange.shrunk > 1)] <- "TGFbIIR-KO"

keyvals[which(de.res$log2FoldChange.shrunk < -1)] <- cols[2]
names(keyvals)[which(de.res$log2FoldChange.shrunk < -1)] <- "WT"

plot.vol <- EnhancedVolcano(de.res, subtitle = "",
                lab = rownames(de.res),
                x = 'log2FoldChange.shrunk',
                y = 'padj',
                # xlim = c(-3, 3),
                # ylim=c(0, -log10(10e-75)),
                labSize = 3,
                pCutoff = 0.05,
                FCcutoff = 1,
                colAlpha = 0.5,
                colCustom = keyvals,
                legendLabSize = 8,
                legendIconSize = 2.0,
                border = "full",
                legendPosition = "right",
                axisLabSize = 20,
                title = "TGFbIIR␣KO␣vs␣WT")

plot.vol
```



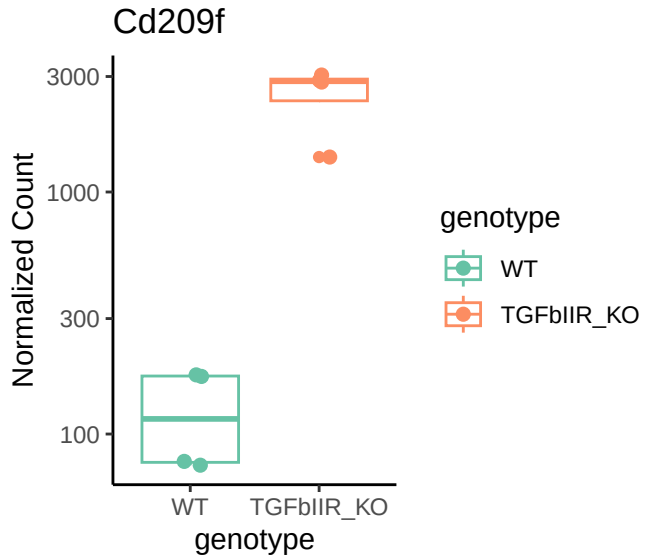
4.2 Plot the top DE genes

`plotCounts` uses normalized counts in dds.

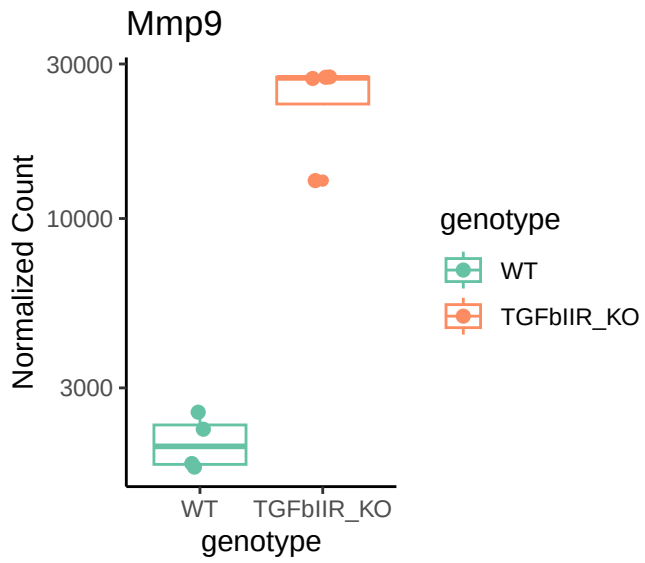
```
genesymbol <- "Cd209g"

data <- plotCounts(dds = dds, gene = genesymbol, intgroup = c("experiment"
    , "genotype"),
    returnData = TRUE)

ggplot(data, aes(x = genotype, y = count, colour = genotype)) + scale_y_
    log10() +
    geom_boxplot() + geom_point(position = position_jitter(width = 0.1,
        height = 0),
    size = 2) + ggtitle(genesymbol) + theme(plot.title = element_text(
        hjust = 0.5,
    size = 20)) + scale_color_manual(values = c(cols[1], cols[2])) + ylab(
        "Normalized Count") +
    theme_linedraw() + theme_classic()
```



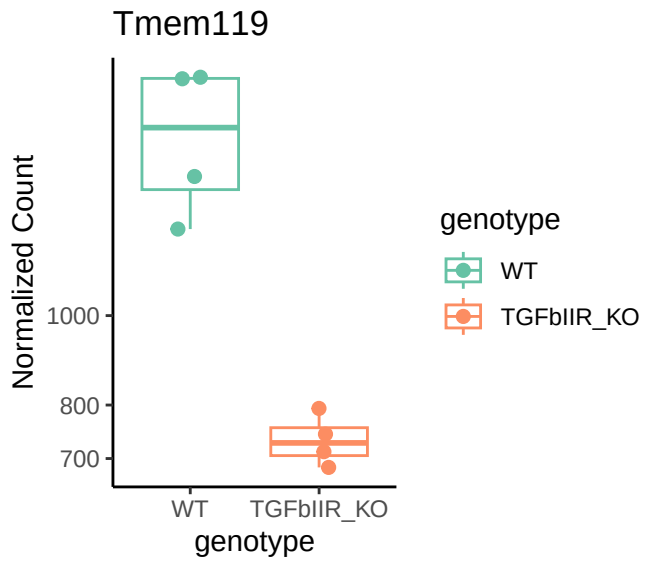

```
genesymbol <- "Cd209f"

data <- plotCounts(dds = dds, gene = genesymbol, intgroup = c("experiment"
    , "genotype"),
     returnData = TRUE)

ggplot(data, aes(x = genotype, y = count, colour = genotype)) + scale_y_
    log10() +
     geom_boxplot() + geom_point(position = position_jitter(width = 0.1,
         height = 0),
     size = 2) + ggtitle(genesymbol) + theme(plot.title = element_text(
         hjust = 0.5,
     size = 20)) + scale_color_manual(values = c(cols[1], cols[2])) + ylab(
         "Normalized Count") +
     theme_linedraw() + theme_classic()
```



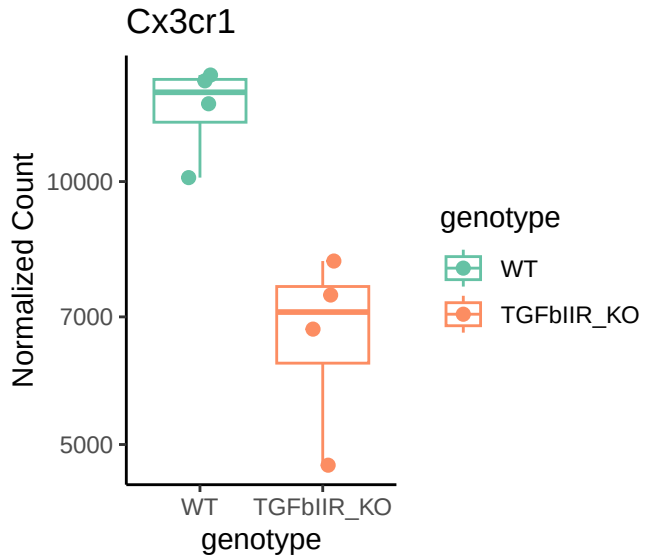
```
genesymbol <- "Mmp9"

data <- plotCounts(dds = dds, gene = genesymbol, intgroup = c("experiment"
    , "genotype"),
     returnData = TRUE)

ggplot(data, aes(x = genotype, y = count, colour = genotype)) + scale_y_
    log10() +
     geom_boxplot() + geom_point(position = position_jitter(width = 0.1,
         height = 0),
     size = 2) + ggtitle(genesymbol) + theme(plot.title = element_text(
         hjust = 0.5,
     size = 20)) + scale_color_manual(values = c(cols[1], cols[2])) + ylab(
         "Normalized Count") +
     theme_linedraw() + theme_classic()
```




```
genesymbol <- "Tmem119"

data <- plotCounts(dds = dds, gene = genesymbol, intgroup = c("experiment"
    , "genotype"),
     returnData = TRUE)

ggplot(data, aes(x = genotype, y = count, colour = genotype)) + scale_y_
    log10() +
     geom_boxplot() + geom_point(position = position_jitter(width = 0.1,
         height = 0),
     size = 2) + ggtitle(genesymbol) + theme(plot.title = element_text(
         hjust = 0.5,
     size = 20)) + scale_color_manual(values = c(cols[1], cols[2])) + ylab(
         "Normalized Count") +
     theme_linedraw() + theme_classic()
```




```
genesymbol <- "Cx3cr1"

data <- plotCounts(dds = dds, gene = genesymbol, intgroup = c("experiment"
    , "genotype"),
     returnData = TRUE)

ggplot(data, aes(x = genotype, y = count, colour = genotype)) + scale_y_
    log10() +
     geom_boxplot() + geom_point(position = position_jitter(width = 0.1,
         height = 0),
     size = 2) + ggtitle(genesymbol) + theme(plot.title = element_text(
         hjust = 0.5,
     size = 20)) + scale_color_manual(values = c(cols[1], cols[2])) + ylab(
         "Normalized Count") +
     theme_linedraw() + theme_classic()
```



Give adjusted p-values for the following genes: - Tmem119 - Cx3cr1 - Cd209f - Cd209g

```
DE.res.KO_vs_WT[c("Tmem119", "Cx3cr1", "Cd209f", "Cd209g"), "padj"]
```

```
## [1] 5.169681e-23 2.121482e-14 4.513817e-21 1.266367e-32
```

5 Session information

R sesssion:

```
sessionInfo()
```

```
## R version 4.3.3 (2024-02-29)
## Platform: aarch64-apple-darwin20 (64-bit)
## Running under: macOS 15.1.1
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/
   lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/
   lib/libRlapack.dylib;  LAPACK version 3.11.0
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## time zone: Europe/Paris
## tzcode source: internal
##
## attached base packages:
## [1] stats4    stats     graphics  grDevices utils     datasets  methods
## [8] base
##
## other attached packages:
##  [1] EnhancedVolcano_1.20.0      ggrepel_0.9.6
##  [3] RColorBrewer_1.1-3          ggplot2_3.5.1
##  [5] DESeq2_1.42.1               SummarizedExperiment_1.32.0
##  [7] Biobase_2.62.0              MatrixGenerics_1.14.0
##  [9] matrixStats_1.4.1           GenomicRanges_1.54.1
## [11] GenomeInfoDb_1.38.8         IRanges_2.36.0
## [13] S4Vectors_0.40.2            BiocGenerics_0.48.1
##
## loaded via a namespace (and not attached):
##  [1] gtable_0.3.6            xfun_0.49               lattice_0.22-6
##  [4] numDeriv_2016.8-1.1     vctrs_0.6.5             tools_4.3.3
##  [7] bitops_1.0-9            generics_0.1.3          parallel_4.3.3
## [10] tibble_3.2.1            fansi_1.0.6             pkgconfig_2.0.3
## [13] Matrix_1.6-5            lifecycle_1.0.4         GenomeInfoDbData_1
   .2.11
## [16] farver_2.1.2            compiler_4.3.3          munsell_0.5.1
## [19] codetools_0.2-20        htmltools_0.5.8.1       RCurl_1.98-1.16
## [22] yaml_2.3.10             pillar_1.9.0            crayon_1.5.3
```

```
## [25] MASS_7.3-60.0.1          BiocParallel_1.36.0     DelayedArray_0
    .28.0
## [28] emdbook_1.3.13           abind_1.4-8             bdsmatrix_1.3-7
## [31] tidyselect_1.2.1         locfit_1.5-9.10         digest_0.6.37
## [34] mvtnorm_1.3-2            dplyr_1.1.4             labeling_0.4.3
## [37] fastmap_1.2.0            grid_4.3.3              colorspace_2.1-1
## [40] cli_3.6.3                SparseArray_1.2.4       magrittr_2.0.3
## [43] S4Arrays_1.2.1           utf8_1.2.4              withr_3.0.2
## [46] scales_1.3.0             rmarkdown_2.29          XVector_0.42.0
## [49] coda_0.19-4.1            evaluate_1.0.1          knitr_1.49
## [52] bbmle_1.0.25.1           rlang_1.1.4             Rcpp_1.0.13-1
## [55] glue_1.8.0               formatR_1.14            apeglm_1.24.0
## [58] rstudioapi_0.17.1        plyr_1.8.9              R6_2.5.1
## [61] zlibbioc_1.48.2
```

References

1. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).

2. H. Wickham, *ggplot2: Elegant graphics for data analysis* (Springer-Verlag New York, 2016; https://ggplot2.tidyverse.org).

3. E. Neuwirth, *RColorBrewer: ColorBrewer palettes* (2022; https://CRAN.R-project.org/package=RColorBrewer).

4. K. Blighe, S. Rana, M. Lewis, *EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling* (2023; https://bioconductor.org/packages/EnhancedVolcano).