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

1-Microarray data preparation

Rendered 2024-12-17 12:52:10 +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 Overall design

CD45.1/CD45.2 IMDTR mice were lethally irradiated with thorax protection and were fully reconstituted with BM cells either from CD45.2 Tgfbr2fl/fl mice or from CD45.2 Lyz2Cre Tgfbr2fl/fl mice. Four weeks later, chimeric IMDTR mice were treated with DT to specifically empty the IM niche and trigger IM niche refilling from either control or Tgfbr2-deficient monocytes. bulk RNA-seq was performed on reconstituted IMs 10 days after DT.

Bulk RNA-seq data have been deposited at the Gene Expression Omnibus (GEO) and are available under GEO accession GSE271467 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE271467). Counts were normalized and managed by `DESeq2` package *(1)* and the batch effects were corrected with `forcats` package *(2)*. The following packages were used to build plots: `ggplot2` *(3)*, `pheatmap` *(4)* and `RColorBrewer` *(5)*.

2 Counting from fastq data using nf-core/rnaseq pipeline

The following codes were used to do the mapping and counting.

```
nextflow run nf-core/rnaseq --input sample_list.csv --fasta GRCm38/fasta/
   genome.fa --gtf GRCm38/genes/genes.gtf --outdir counts/bulkRNAseq/ -
   profile docker
```

`sample_list.csv` is text file with 5 columns: group, replicate, fastq_1, fastq_2 and strandedness. Prepared following to the software's instructions.

3 Counts data processing

```
suppressMessages(library(DESeq2))
suppressMessages(library(ggplot2))
suppressMessages(library(pheatmap))
suppressMessages(library(RColorBrewer))
suppressMessages(library(forcats))

COUNTS <- read.table("../../data/salmon.merged.gene_counts.tsv", sep = "\t
   ", header = T,
    row.names = NULL)

Genes <- COUNTS$gene_id
rownames(COUNTS) = make.names(Genes, unique = TRUE)

COUNTS <- COUNTS[, c(-1, -2)]
COUNTS <- round(COUNTS, digits = 0)
head(COUNTS, 3)
```

```
## # A tibble: 3 x 8
##     KO1   KO2   KO3   KO4   WT1   WT2   WT3   WT4
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1  4196  3920  3538  3090  2205  3339  2926  5248
## 2     0     0     0     0     0     0     0     0
## 3   142   129   109   103    52   105    96   147
```

Annotation with org.Mm.eg.db package:

```
library(org.Mm.eg.db)
symbols <- mapIds(org.Mm.eg.db, keys = rownames(COUNTS), keytype = "
   ENSEMBL", column = "SYMBOL")
symbols.uniq <- na.omit(unique(symbols))

# remove adundant ensembl ids:
COUNTS <- COUNTS[match(symbols.uniq, symbols), ]

# use symbols as rownames:
rownames(COUNTS) <- symbols.uniq

head(COUNTS)
```

```
## # A tibble: 6 x 8
##     KO1   KO2   KO3   KO4   WT1   WT2   WT3   WT4
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1  4196  3920  3538  3090  2205  3339  2926  5248
## 2     0     0     0     0     0     0     0     0
## 3   142   129   109   103    52   105    96   147
## 4    33    28    26    23    10    18     9    19
## 5     1     0     0     0     0     0     0     0
## 6   405   311   249   249   274   281   204   374
```

4 Make metadata for bulkRNAseq samples

```
SampleSheet <- data.frame(genotype = rep(c("TGFbIIR_KO", "WT"), each = 4),
    experiment = c("exp1",
    rep("exp2", 3), rep("exp1", 2), rep("exp2", 2)))
rownames(SampleSheet) <- colnames(COUNTS)

SampleSheet
```

```
## # A tibble: 8 x 2
##   genotype   experiment
##   <chr>      <chr>
## 1 TGFbIIR_KO exp1
## 2 TGFbIIR_KO exp2
## 3 TGFbIIR_KO exp2
## 4 TGFbIIR_KO exp2
## 5 WT         exp1
## 6 WT         exp1
## 7 WT         exp2
## 8 WT         exp2
```

5 DESeq2 analysis

```
dds <- DESeqDataSetFromMatrix(countData = COUNTS, colData = SampleSheet,
   design = ~experiment +
    genotype)
```

```
# Since data were from two sequencing experiments, we applied the ComBat
# function to adjust for known batches

suppressMessages(library(sva))
batch = SampleSheet$experiment
modcombat = model.matrix(~1, data = SampleSheet)
group = SampleSheet$genotype

# using parametric empirical Bayesian adjustments.
combat_COUNTS = ComBat_seq(counts = as.matrix(COUNTS), batch = batch,
   group = group)
```

```
## Found 2 batches
## Using full model in ComBat-seq.
## Adjusting for 1 covariate(s) or covariate level(s)
## Estimating dispersions
## Fitting the GLM model
## Shrinkage off - using GLM estimates for parameters
## Adjusting the data
```

5.1 Perform rlog / vst transformation for distances and PCA

```
dds.rmBatchEff <- DESeqDataSetFromMatrix(countData = combat_COUNTS,
   colData = SampleSheet,
    design = ~experiment + genotype)

# keep only genes with more than a single read
dds.rmBatchEff <- dds.rmBatchEff[rowSums(counts(dds.rmBatchEff)) > 1, ]

# perform vst transformation for distances (for clustering) and PCA
vst.rmBatchEff <- vst(dds.rmBatchEff)
rld.rmBatchEff <- rlog(dds.rmBatchEff)
```

Calculate sample-to-sammple distances

```
sampleDists <- dist(t(assay(rld.rmBatchEff)))
sampleDistMatrix <- as.matrix(sampleDists)
```

5.2 Heatmap

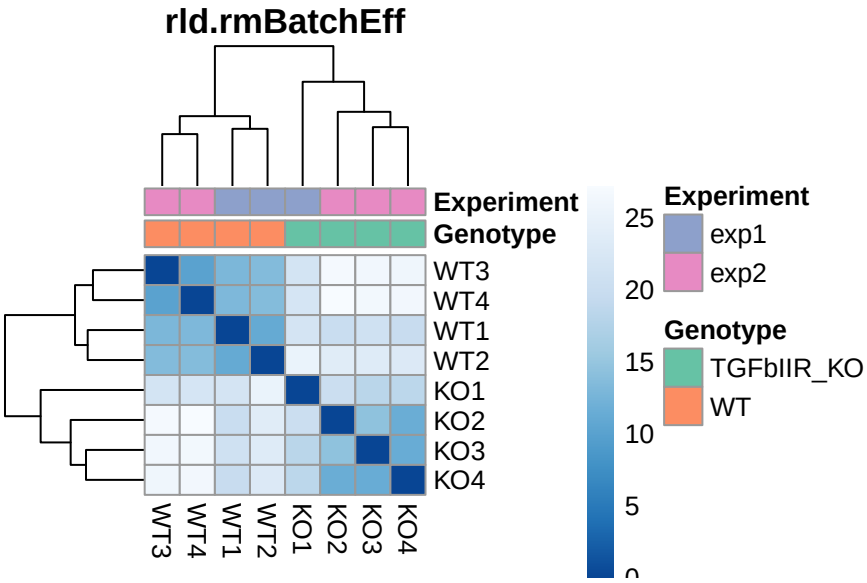
```
colors <- colorRampPalette(rev(brewer.pal(ncol(COUNTS), "Blues")))(255)
cols <- brewer.pal(5, "Set2")

annotation_col = data.frame(Genotype = factor(SampleSheet$genotype),
   Experiment = factor(SampleSheet$experiment),
```

```
    row.names = rownames(SampleSheet))

ann_colors = list(Genotype = c(TGFbIIR_KO = cols[1], WT = cols[2]),
   Experiment = c(exp1 = cols[3],
    exp2 = cols[4]))

heatmap <- pheatmap(sampleDistMatrix, clustering_distance_rows =
   sampleDists, clustering_distance_cols = sampleDists,
    annotation_col = annotation_col, annotation_colors = ann_colors, col =
        colors,
    main = "rld.rmBatchEff")
```



5.3 PCA analysis

Calculate PCs:

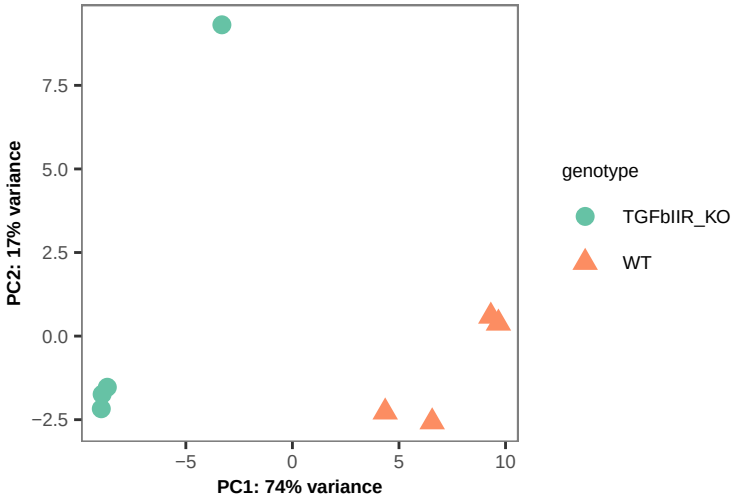
```
PlotData <- plotPCA(rld.rmBatchEff, intgroup = c("genotype", "experiment")
   , returnData = TRUE)
```

Construct plot:

```
percentVar<-round(100 * attr(PlotData, "percentVar"))

ggplot(PlotData,aes(PC1,PC2)) +
    geom_point(size=3, aes(color = genotype, shape = genotype)) +
    xlab(paste0("PC1:␣", percentVar[1],"%␣variance")) +
    ylab(paste0("PC2:␣", percentVar[2],"%␣variance"))  + theme(
    aspect.ratio=1,
    panel.background = element_rect(fill = "white", colour = "grey50"),
    axis.text=element_text(size=7),
        axis.title=element_text(size=7,face="bold"),
    legend.key = element_blank(),
    legend.text = element_text(size=7),
    legend.title = element_text(size = 7),
```

```
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank()
  ) + scale_color_manual(
    breaks = c( "TGFbIIR_KO",
                "WT"),
    values = c(
                      cols[1], # TGFbIIR_KO
                      cols[2] # WT
                    ))
```



6 Session information

Nextflow:

```
Nextflow version: version 21.03.0.edge, build 5518 (05-03-2021 10:52 UTC)
Workflow profile: docker
Workflow repository: https://github.com/nf-core/rnaseq, revision master (
   commit hash 3643a94411b65f42bce5357c5015603099556ad9)
```

Software version used by Workflow:

```
bedtools    2.29.2
bioconductor-summarizedexperiment   1.20.0
bioconductor-tximeta    1.8.0
deseq2  1.28.0
dupradar    1.18.0
fastqc  0.11.9
nextflow    21.03.0.edge
nf-core/rnaseq  3.0
picard  2.23.9
preseq  2.0.3
qualimap    2.2.2-dev
rseqc   3.0.1
```

```
salmon  1.4.0
samtools    1.10
star    2.6.1d
stringtie   2.1.4
subread 2.0.1
trimgalore  0.6.6
ucsc    377
```

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] sva_3.50.0                  BiocParallel_1.36.0
##  [3] genefilter_1.84.0           mgcv_1.9-1
##  [5] nlme_3.1-166                org.Mm.eg.db_3.18.0
##  [7] AnnotationDbi_1.64.1        forcats_1.0.0
##  [9] RColorBrewer_1.1-3          pheatmap_1.0.12
## [11] ggplot2_3.5.1               DESeq2_1.42.1
## [13] SummarizedExperiment_1.32.0 Biobase_2.62.0
## [15] MatrixGenerics_1.14.0       matrixStats_1.4.1
## [17] GenomicRanges_1.54.1        GenomeInfoDb_1.38.8
## [19] IRanges_2.36.0              S4Vectors_0.40.2
## [21] BiocGenerics_0.48.1
##
## loaded via a namespace (and not attached):
##  [1] tidyselect_1.2.1        farver_2.1.2            dplyr_1.1.4
##  [4] blob_1.2.4              Biostrings_2.70.3       bitops_1.0-9
##  [7] fastmap_1.2.0           RCurl_1.98-1.16         XML_3.99-0.17
## [10] digest_0.6.37           lifecycle_1.0.4         statmod_1.5.0
## [13] survival_3.7-0          KEGGREST_1.42.0         RSQLite_2.3.9
## [16] magrittr_2.0.3          compiler_4.3.3          rlang_1.1.4
## [19] tools_4.3.3             utf8_1.2.4              yaml_2.3.10
## [22] knitr_1.49              labeling_0.4.3          S4Arrays_1.2.1
```

```
## [25] bit_4.5.0.1              DelayedArray_0.28.0     abind_1.4-8          41
## [28] withr_3.0.2              grid_4.3.3              fansi_1.0.6          42
## [31] xtable_1.8-4             colorspace_2.1-1        edgeR_4.0.16         43
## [34] scales_1.3.0             cli_3.6.3               rmarkdown_2.29       44
## [37] crayon_1.5.3             generics_0.1.3          rstudioapi_0.17.1    45
## [40] httr_1.4.7               DBI_1.2.3               cachem_1.1.0         46
## [43] zlibbioc_1.48.2          splines_4.3.3           parallel_4.3.3       47
## [46] formatR_1.14             XVector_0.42.0          vctrs_0.6.5          48
## [49] Matrix_1.6-5             bit64_4.5.2             locfit_1.5-9.10      49
## [52] limma_3.58.1             annotate_1.80.0         glue_1.8.0           50
## [55] codetools_0.2-20         gtable_0.3.6            munsell_0.5.1        51
## [58] tibble_3.2.1             pillar_1.9.0            htmltools_0.5.8.1    52
## [61] GenomeInfoDbData_1.2.11 R6_2.5.1                 evaluate_1.0.1       53
## [64] lattice_0.22-6           png_0.1-8               memoise_2.0.1        54
## [67] Rcpp_1.0.13-1            SparseArray_1.2.4       xfun_0.49            55
## [70] pkgconfig_2.0.3                                                       56
```

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, *Forcats: Tools for working with categorical variables (factors)* (2023; https://CRAN.R-project.org/package=forcats).

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

4. R. Kolde, *Pheatmap: Pretty heatmaps* (2019; https://CRAN.R-project.org/package=pheatmap).

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