Single cell RNA sequencing based identification and profiling of Leishmania parasitized host cells in a murine model

Konstantinos **Karagiannis**[1] and Sreenivas **Gannavaram**[1], Luis **Santana-Quintero**[1], Hira L **Nakhasi**[1]

[1]Center for Biologics Evaluation and Research (CBER), Food and Drug Administration (FDA), Silver Spring, MD, USA

Introduction

Leishmaniasis is among the top ten neglected tropical diseases (NTD), with over 12 million infected people, and 2 million new cases annually. Visceral leishmaniasis (VL), a life-threatening form of leishmaniasis caused by Leishmania Donovani (L. Don), is the second most lethal parasitic disease after malaria. WHO identified US as an endemic country for this blood borne parasite for which there are no approved donor screening assays or vaccines. Post kala azar dermal leishmaniasis (PKDL) is a dermatological complication of L. Don infection which occurs in some VL patients after successful treatment.

Despite sequencing studies that compared L.Don strains and identified genomic differences, the cause of the different clinical outcome remains elusive. Little is known about Leishmania gene expression in different organs during infection and about the phagocytes that mediate parasite dissemination and/or serve as reservoirs for the parasites in visceral organs (spleen and bone marrow). Using a mouse model of L. Don infection, and an innovative analytical workflow we identify parasitized cells from spleen and bone marrow using a single cell RNA sequencing technologies.

Materials and Methods

To identify the diversity of the parasitized host cells and the heterogenous transcription responses that underlie pathogenesis and persistent infections, spleen samples from a Balb/C mouse infected with L. Don were collected and processed with 10X Genomics Chromium Single Cell 3' GEM Library and Gel Bead Kit v3.1. 3'seq libraries were sequenced using Illumina NovaSeq. Raw FASTQ files were processed 10x Genomics Cell Ranger 3.1.0, which performs filtering, barcode counting and UMI counting and alignment to the GRch38 and L. Don transcriptomes. The generated transcript counts per cell were processed using the Seurat[1] (v4.0). The total count of L. Don transcripts were used to infer presence or absence of the parasite per cell . Only mouse transcript were considered for downstream analysis. The cells were clustered by applying the K-nearest neighbors (KNN) graph approach, followed by Louvain's algorithm. The clustered cells were annotated with the SCSA[2] software and existing scRNA-seq specific markers available from the CellMarker[3] curated resource.

Results

Using a mouse model of L. Don infection, we have constructed an analytical workflow to identify parasitized cells using a concatenated host and parasite reference genome (Fig. 1). The workflow resulted to 7261 and 7356 cells from the bone marrow and mouse spleen derived sample respectively. In the spleen sample a total of 109 cell were detected to contain both mouse and L. Don transcripts and were identified as infected cells. Similarly, more than 230 cells were characterized as infected from the bone marrow derived sample (Fig.2). Visualization of the cells from each sample using Uniform Manifold Approximation and Projection (UMAP) dimensionality reduction technique, revealed that the parasitized cells appear to be clustering predominantly together in both bone marrow and spleen samples (Fig. 3 and Fig. 4), suggesting that L. Don infects specific cell populations.

Careful examination of the transcriptional profiles of the spleen derived cells revealed populations including, T cells, natural killer cells (NK cells), Medullary, E cells, Mesenchymal progenitor cells, mast, red blood cells, neutrophils, monocytes and macrophages (Fig. 5). The latter is the most numerous parasitized cell population with more than 200 infected cells, corresponding to >85% of total cells with detected L. Don transcripts (Fig. 6).

Transcription profiles of splenic macrophages, the most numerous parasitized population, were compared to uninfected macrophages. Consistent with the literature, several of the transcripts altered in the infected macrophages (Fig. 7) correspond to M1/M2 polarization (Xist, Mycbp2, klf13, Gpx1), cholesterol acquisition (Npc2), Inflammasome assembly (Neat 1), LPS-mediated IL-6 production (Sf3B1), and intracellular endosomal acidification (lfitm2).

Figure 3. Uniform Manifold Approximation and Projection (UMAP) embedding of 7261 cells from the bone marrow sample L.Don infected cells are colored in blue.

Figure 4. Uniform Manifold Approximation and Projection (UMAP) embedding of 7356 cells from the spleen sample, L.Don infected cells are colored in blue.

Figure 5. Single-cell transcriptomic UMAP embedding of spleen derived cells colored based on identified cell populations based on CellMarker database.

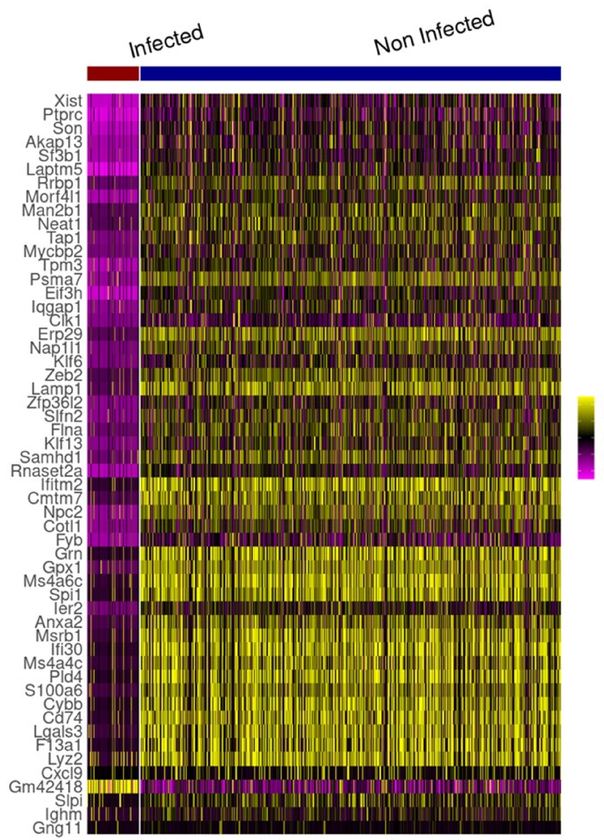

Figure 7. Heatmap of differentially expressed genes between the infected and non infected macrophage cells from the spleen sample.

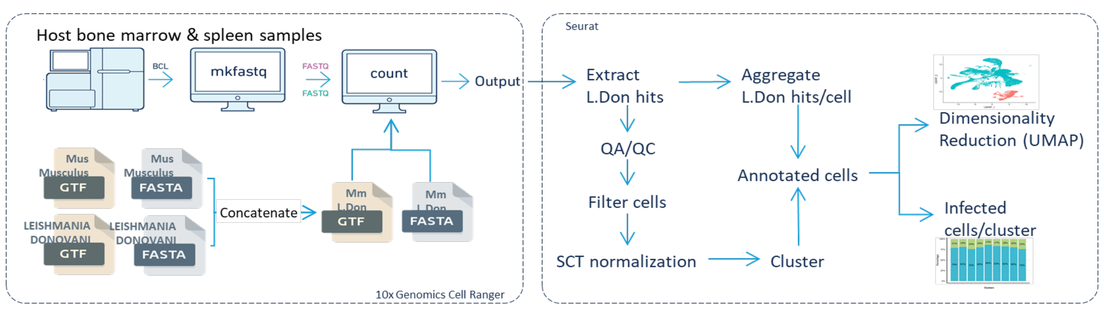

Figure 1. Schematic representation of the analytical workflow. Base calling, read alignment and hit calling performed by CellRange. Mus musculus and Leishmania Donovani sequences and annotation files were concatenated, and the reads were aligned against a single reference. The hits were processed using Seurat package.

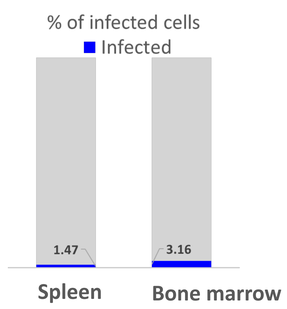

Figure 2. Bar plot showing the percentage of infected cells detected in each sample.

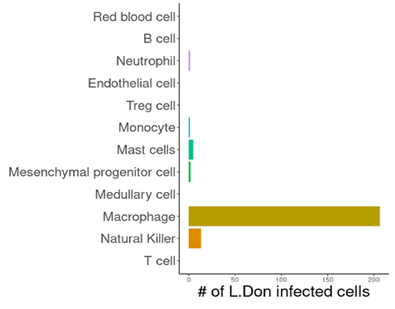

Figure 6. Distribution of infected cells across different populations following Fig. 5 color coding

Conclusion

This first of its kind scRNA-Seq application successfully quantified parasitized cells, identified cell types harboring parasites and examined their transcriptional profile. Hence, this approach can reveal new types of parasitized cells and simultaneously provide insight into their transcriptional landscape, moving towards better diagnostic assays and clinical interventions.

References

1. Stuart T. et al. *Comprehensive integration of Single-cell data*. Cell, 2019
2. Yinghao C. et al. *SCSA: A cell type annotation tools for single-cell RNA-seq data.* Front Genet, 2020
3. Zhang X. et al. *CellMarker: a manually curated resource of cell markers in human and mouse*. NAR, 2019