bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Combinatorial transcriptional codes specify macrophage polarization destinations

Sathyabaarathi Ravichandran[1], Bharat Bhatt[2], Kithiganahalli Narayanaswamy Balaji[2], Nagasuma Chandra[134*]

[1]IISc Mathematics Initiative, Indian Institute of Science, Bangalore 560012, India.

[2]Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 560012, India.

[3]Department of Biochemistry, Indian Institute of Science, Bangalore 560012, India.

[4]Centre for Biosystems Science and Engineering, Indian Institute of Science, Bangalore 560012, India.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Summary

Macrophages are driven to form distinct functional phenotypes in response to different immunological stimuli, in a process widely referred to as macrophage polarization. Transcriptional regulators that guide macrophage polarization in response to a given trigger remain largely unknown. In this study, we interrogate the programmable landscape in macrophages to find regulatory panels that determine the precise polarization state that a macrophage is driven to. Towards this, we configure an integrative network analysis pipeline that utilizes macrophage transcriptomes in response to 28 different stimuli and reconstructs contextualized human gene regulatory networks, and identifies epicentres of perturbations in each case. We find that these contextualized regulatory networks form a spectrum of thirteen distinct clusters with M1 and M2 at the two ends. Using our computational pipeline, we identify combinatorial panels of epicentric transcription factors (TFs) for each polarization state. We demonstrate that a set of three TFs i.e., *NFE2L2*, *BCL3* and *CEBPB*, is sufficient to change the polarization destination from M1 to M2. siRNA-mediated knockdown of the 3-TF set in THP1 derived M0 cells, despite exposure to an M1 stimulant, significantly attenuated the shift to M1 phenotype, and instead switched to an M2-like phenotype. Further, the siRNA-mediated knockdown of the 3-TF set rendered the macrophages hyper-susceptible to *S. aureus* infection, clearly showing the effect of the knockdown.

Keywords: macrophage polarization, gene regulatory networks, regulatory barcodes, switching factors, M1-M2 switching.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Background

Macrophages, the key components of the innate immune system, are responsible for processes ranging from the maintenance of healthy tissues, phagocytic clearance of microbes, to tissue repair and remodelling (Aderem, 2003; Greenberg and Grinstein, 2002; Mosser et al., 2021; Murray and Wynn, 2011). They do most of these in a bid to restore normal physiology and act as a bridge between innate and adaptive immunity. In response to a range of stimuli, including pathogen insults, macrophages are directed to distinct functional phenotypes in a process widely referred to as macrophage polarization (Benoit et al., 2008; Murray, 2017). The most well-described phenotypes are the M1 and M2 polarization states, where M1 macrophages are pro-inflammatory in nature with microbicidal activity, while M2 is associated with an anti-inflammatory response. Stimulation of macrophages with lipopolysaccharides (LPS) or interferon-gamma (IFNγ) typically leads to an M1 state, whereas treatment with interleukins (IL4, IL13) polarize macrophages into the M2 state (Mills et al., 2000; Murray, 2017). M2 macrophages are further sub-categorised into M2a, M2b, M2c and M2d depending on the activation triggers and the resulting transcriptional changes (Mantovani et al., 2004; Rőszer, 2015; Yao et al., 2019). However, this model is said to be oversimplified as it fails to capture the complexity exhibited by different macrophage states in an *in vivo* scenario where the interaction between different cytokines, interleukins, and other effector molecules define the final polarized state (Atri et al., 2018; Martinez and Gordon, 2014) . A recent study on the transcriptomes of human monocyte-derived macrophages treated with a variety of physiological stimulants identified nine different polarization states, which better fit the description of a spectrum of states rather than a classical binary M1/M2 axis (Xue et al., 2014). The different states in the polarization spectrum present complex transcriptional profiles as seen from altered gene expression patterns in several hundred genes. The importance of transcriptional regulation in defining macrophage polarization has been brought out by several studies (Gerrick et al., 2018; Hörhold et al., 2020; Palma et al., 2018; Xue et al., 2014; Zhao et al., 2021). *In vitro* and *ex-vivo* studies have shown complete reversal of macrophage polarization states depending upon the molecules (pathogen-associated molecular patterns (PAMPs), host-derived damage-associated molecular patterns (DAMPs) or cytokines, etc) present in the microenvironment (Sica and Mantovani, 2012; Stout et al., 2005; Zhang and Mosser, 2008).

Several studies have shown that activation of transcription factors (TFs) such as signal transducers and activators of transcription (STAT1, STAT3, STAT4 and STAT6), nuclear transcription factor-κB (NF-kB), NF-E2-related factors (NRF1 and NRF2), kruppel-like factors (KLF4 and KLF6),

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

CCAAT-enhancer-binding proteins (C/EBP-α and C/EBP-β) and peroxisome proliferator-activated receptors (PPAR-α, PPAR-δ and PPAR-γ) are associated with macrophage polarization (Lawrence, 2011; Li et al., 2018; Tugal et al., 2013). But, the available information is insufficient to predict the polarization state that will result in a given condition. A possible reason is that a set of TFs together defines the state rather than individual TFs by themselves. While the triggers that lead to different polarization states are known, the specific set of TFs that nucleate the change or serve as control points to guide macrophage polarization is not known.

The precise polarization state attained by the macrophages in response to pathogenic triggers have a direct bearing on the outcome of the disease (Benoit et al., 2008; Koziel et al., 2009; Pidwill et al., 2021). In response to bacterial pathogens such as *Staphylococcus aureus*, macrophages induce the transcriptional activity of pro-inflammatory genes, activating the M1 program. The M1 phenotype is known to have an activated microbicidal machinery that is utilized in clearing such acute infections. However, prolonged or excessive inflammatory response by M1 macrophages can be pathological as it results in tissue damage and destruction. M2 macrophages produce anti-inflammatory mediators and resolve the inflammation (Koziel et al., 2009; Pidwill et al., 2021). Not surprisingly, the skewing of macrophage polarization is observed in several cases of severe bacterial infections, sepsis, tumour progression in cancers, and a range of other inflammatory diseases (Ardura et al., 2019; Hamilton and Tak, 2009; Sica and Mantovani, 2012; Zheng et al., 2018). Although our understanding of macrophage biology is rapidly increasing, the precise molecular controls and their associated processes involved in macrophage switching to different functional states are not well characterized. Identifying combinations of molecules that guide macrophage polarization will greatly facilitate a deeper understanding of how infections are contained, provide a basis to predict the precise phenotype that will be attained in different situations and serve as a foundation for precision diagnostics and treatment strategies focused on the macrophages.

In this work, we seek to interrogate the programmable landscape in macrophages to identify regulatory panels that govern the precise macrophage polarization state. Towards this, we configured a computational pipeline that uses an integrative graph-theoretical approach where differential transcriptomes of macrophages were mapped onto the knowledge-based genome-scale gene regulatory network (hGRN), computed connected sets of top-ranked perturbations, and identified regulatory barcodes defining different polarization states ultimately. We then experimentally evaluated specific TFs in the M1 regulatory barcode in a THP-1 derived macrophage cell line. We also probed the role of these specific TF sets in the *S.aureus* infected

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

human monocyte-derived macrophages and cell-line infection model. Together, we validated genes in the M1 barcode and demonstrated that knocking down specific TF sets in the barcode switches M1 to an M2 state. We further showed that such a state switch leads to hyper-susceptibility to *S.aureus* infection.

Materials and Methods

Transcriptomes of macrophage polarization states: We retrieved transcriptomes of macrophages stimulated with a diverse set of stimulants (GSE46903) from NCBI-GEO (Barrett et al., 2013; Xue et al., 2014). Briefly, the authors of this study purified monocytes from peripheral blood mononuclear cells (PBMCs) of healthy controls and subjected them to granulocyte-macrophage colony-stimulating factor (GM-CSF) or macrophage-colony stimulating factor (M-CSF) for 72 hours (hrs) to generate baseline macrophages (Mo). Then, they treated the Mo with one of the twenty-eight different stimulants for different time points ranging from 30 mins to 72 hrs and generated transcriptomes from each stimulant and each time point. The stimulants spanned across the broad spectrum of pro and anti-inflammatory triggers to generate a wide range of polarization states. The triggers and their end time-points (n=145) considered for this analysis are shown in Table S1. Further, to study the relevance of macrophage polarization in acute bacterial infections, transcriptomes of peripheral macrophages exposed to *S. aureus* were retrieved from NCBI-GEO (GSE13670) (Koziel et al., 2009).

Raw data from the respective transcriptome experiments were retrieved and pre-processed in a platform-specific manner using appropriate packages in R (Carvalho and Irizarry, 2010; Gautier et al., 2004). Probes below the detection limit in ≥ 80% of the samples in each dataset were filtered out, and multiple probes associated with the same gene were summarized by taking an average. We then performed differential analysis for the macrophages treated with stimulants compared to their untreated counterparts, using the limma package in R(Ritchie et al., 2015). Genes with fold change ≥ ±1.3 and a False Discovery Rate corrected q-value ≤0.05 computed using the Benjamini-Hochberg procedure(Benjamini and Hochberg, 1995)were considered to be statistically significant differentially expressed genes (DEGs).

Reconstruction of human Gene Regulatory Network (hGRN): We reconstructed a comprehensive knowledge-based human Gene Regulatory Network (hGRN), consisting of Transcription Factor (TF) to Target Gene (TG) and TF to TF interactions. To achieve this, we curated the experimentally determined regulatory interactions (TF-TG, TF-TF) associated with human TFs (Wingender et al., 2013) in this study. These interactions were obtained from the

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

following resources: (a) literature curated resources such as the Human Transcriptional Regulation Interactions database (HTRIdb)(Bovolenta et al., 2012), Regulatory Network Repository (RegNetwork) (Liu et al., 2015), Transcriptional Regulatory Relationships Unravelled by Sentence-based Text-mining (TRRUST) (Han et al., 2015) , TRANSFAC resource from Harmonizome (Rouillard et al., 2016), (b) ChEA3 containing ChIP-seq determined interactions (Keenan et al., 2019) and (c) high confidence protein-protein binding interactions (TF-TF) were retrieved from human protein-protein interaction network -2 (hPPiN2) (Ravichandran et al., 2021). The resulting interactions from both categories were collated and made a non-redundant set. This resulted in a connected set of TF-TF and TF-TG interactions that comprise a 'master' human Gene Regulatory Network (hGRN) with 27, 702 nodes and 890, 991 interactions (Data file S1).

Model contexting and identification of epicentric TFset: Starting from the hGRN, stimulant-specific GRNs were derived by weighting the master hGRN network with differential transcriptomes of stimulants (ie., stimulant versus baseline) and generated weighted networks for individual stimulants. A sensitive network mining approach, previously established in the laboratory was then applied to identify top-ranked activated and repressed sub-networks using Eq. 1, Eq. 2, and Eq. 3. This method of computing response networks which involves a knowledge-based network and a sensitive interrogation algorithm has been shown to outperform data-driven network inference methods in capturing biologically relevant processes and genes (Ravichandran and Chandra, 2019; Sambaturu et al., 2021). Briefly, the network mining algorithm works by computing minimum weight shortest paths, in which each path begins from a source node and ends with a sink node, identifying connected sets of edges that make up the least-cost paths. The shortest paths between all pairs of genes were computed using Dijkstra's algorithm using Zen library implemented in Python2.7. For a path of length n, the path cost was calculated as a summation of the edge weights EW_{ij} of all edges forming the path, and normalized by the path length. Subsequently, all shortest paths were sorted based on their path costs from lowest to highest, and the top 1% were considered to constitute the top-perturbed network (TPN). Nodes forming the TPN were evaluated for enrichment of DEGs using Fisher's exact test. Further, the epicentric nodes in each TPN were identified using the EpiTracer algorithm developed previously in the laboratory (Sambaturu et al., 2016). Epitracer estimates the influential state of a node by computing ripple centrality, which is the product of closeness centrality and outward reachability.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The TPN has two components, a top-activated network and a top-repressed network, generated by using appropriate node-weighting schemes. To generate the activated network, the weight of node i in a condition treated with Stimulant A (S_A) was computed as:

$N_i(S_A) = FC_i(S_A/S_B)$ **Eq. 1**

Where FC was the fold change of gene i in a condition treated with stimulant A (S_A) with respect to the baseline S_B (antilog values were used to compute fold changes). To generate the repressed network, the node weight of node i in SA was computed as:

$N_i(S_A) = FC_i(S_B/S_A)$ **Eq. 2**

The edge weight EW_{ij} in a given condition S_A for an edge e comprised of nodes $N_i(S_A)$ and $N_j(SA)$ was calculated as

$EW_{ij}(S_A) = 1/ (N_i(S_A) * N_j(S_A))^{\wedge 1/2}$ **Eq. 3**

Where $N_i(S_A)$ and $N_j(S_A)$ are the node weights of nodes i and j, respectively. Lower the edge weight, stronger is the interaction between the nodes.

These subnetworks were combined to obtain a top perturbed network (TPN) for each stimulant. Nodes in the TPN were rank-ordered based on the ripple centrality score. Differentially expressed TFs with the ripple centrality score >0 were considered as influential transcription factors for a given TPN. The resultant TF-set defines the regulatory control for a given stimulant.

Clustering macrophage polarization states: We performed Monte Carlo reference-based consensus clustering using the M3C package implemented in R (John et al., 2020) to identify the optimal number of macrophage polarization states. Spectral clustering with Pearson's correlation coefficient as a correlation metric for 200 iterations with 250 replicates was used to estimate the cluster structure that best fits the differential transcriptome of (a) entire transcriptome (n=12,164), and (b) TF-set1 (n=265). Further, Relative Cluster Stability Index (RCSI) and Monte Carlo adjusted p-values were computed using M3C to test the number and structure of the clusters. For the cluster visualization and comparison, Dendextend package in R (Galili, 2015) was used.

Network visualization and Enrichment analysis: We have visualized networks using Cytoscape 3.2.0 in Allegro Spring-Electric layout (Shannon et al., 2003). We performed enrichment analysis for the genes of the regulatory cores of each cluster using Reactome pathway resource with default parameters (Fabregat et al., 2017) and the resultant hits with q-value ≤ 0.05 were considered to be significant (using a hypergeometric test).

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Reagents, Cell line and Bacteria: General laboratory chemicals were procured from Merck, Promega, Invitrogen, Himedia and Sigma-Aldrich. Tissue culture plastic wares were purchased from Biofil, BD Falcon and Tarson. Phorbol 12-myristate 13-acetate (PMA #P8139) and lipopolysaccharide (LPS #8630) were procured from Sigma-Aldrich. IL4 (#204-IL) were purchased from R&D Systems. THP-1 human monocytic cell line was obtained from National Centre for Cell Sciences, Pune, India and cultured in RPMI (GIBCO, Invitrogen Corporation, USA) supplemented with 10% heat inactivated FBS (Sigma-Aldrich, USA). THP-1 cells were treated with PMA (20ng/ml) for 24hrs for its differentiation into macrophages. Post PMA treatment, THP-1 derived macrophages were washed and kept in complete RPMI media followed by treatment/ infection as mentioned. *S. aureus Cowan 1* (MTCC - 902) was procured from Microbial Type Culture Collection, IMTECH Chandigarh. *S. aureus* was grown in Luria Broth at 37°C overnight. *S. aureus* was harvested and washed with phosphate-buffered saline (PBS) once the OD600 reached 0.6, and was then resuspended in the cell culture RPMI complete medium without antibiotics. *S. aureus* was used at 10 MOI for in vitro experiments.

RNA isolation and quantitative real-time PCR: Total RNA was isolated from PMA differentiated THP-1 cells using TRI Reagent (Sigma-Aldrich, USA) as per the manufacturer's protocol. For RT-PCR, 1µg of total RNA was converted into cDNA using First strand cDNA synthesis kit (BIO-RAD, USA) according to the manufacturer's protocol. Quantitative real time PCR was performed with SYBR Green PCR mix (KAPABIOSYSTEMS, USA) for quantification of target gene expression. Amplification of housekeeping gene GAPDH was used as an internal control. The list of primers used for the experiment is provided in Table S2.

Transient transfection: PMA differentiated THP-1 cells were transiently transfected with 100nM targeted siRNA using low molecular weight polyethyleneimine (Sigma-Aldrich, 40872-7). Further, 36 hrs post transfection, cells were treated or infected as indicated for required time and processed for analysis.

In vitro CFU analysis: PMA differentiated THP-1 cells were infected with *S. aureus Cowan I* at an MOI of 10 for 2 hrs to facilitate internalization. Following 2 hours, cells were washed with complete DMEM (without antibiotics) and Gentamicin (100µg/ml) treatment was given for 2 hours to remove extracellular bacilli. Cells were lysed in 0.1% TritonX-100 containing 1X PBS and CFU estimation was done at 0hr (after gentamicin treatment) and 24hrs by plating appropriate dilution on LB plates.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Statistical analysis: Levels of significance for comparison between samples were determined by the Student's t-test distribution and one-way ANOVA followed by Tukey's multiple-comparisons. The data in the graphs are expressed as the mean ± standard error for the values from at least 3 or more independent experiments and p values ≤ 0.05 were defined as significant. R version 3.6.3 was used for all statistical analyses.

Results

Subnetworks from contextualised hGRN identifies TF barcodes for macrophage polarization states: We first identified the 'programmable space' in macrophages in terms of the nature and range of transcriptional reprogramming that they achieve in response to different immunostimulants. The stimulants belong to broad classes of pro and anti-inflammatory cytokines such as Toll-like receptors (TLRs) ligands, immune complexes, interleukins, glucocorticoids or their selected combinations, comprehensively covering the macrophage polarization spectrum. We probed the programmable space by constructing and analyzing 28 unbiased 'response networks' of macrophages, corresponding to the treatment with 28 different immunostimulants. Towards this, we first constructed a knowledge-based human gene regulatory network (hGRN). Our hGRN covers regulatory elements from the whole genome and contains 27,702 genes and 890, 991 TF-TF and TF-TG directed interactions, which was rendered context-specific for each stimulant using their respective differential transcriptome data. We then subjected the 28 contextualized hGRNs to our network mining algorithm, which identified a connected set of genes that collectively vary the most in their expression levels in response to the given stimulant. This yielded 28 different subnetworks containing the highest ranked regulatory perturbations, with nodes ranging from 4,032 to 6,785 nodes (Data file S2). The resultant subnetworks containing connected sets of weighted and directed interactions capture cascades of regulatory events which are in essence the top-perturbed networks (TPNs) in each condition. Next, we configured a pipeline to identify the most influential transcription factors from the resultant TPNs. We used EpiTracer, an algorithm previously developed in the laboratory (Sambaturu et al., 2016) which identifies the highest ranked influential nodes in a given network using a ripple centrality measure. This step results in the identification of the genes that have the strongest 'causal' association with a given perturbation pattern. The word 'causal' here refers to the potential of the given gene for wielding an influence on the system, as estimated by the network analysis. Put together, we identified epicentric transcription factors (Ripple centrality score >0) which are differentially expressed in each of these TPNs. When pooled, these TFs from all stimulants, result in a set of 265 TFs (TF-set1 - Data file S3), defining the programmable space in relation to macrophage polarization.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Further, to probe the number of distinct polarization states, we carried out agnostic clustering based on the differential transcriptome of TF-set1, using a Monte Carlo reference-based consensus clustering (M3C) method. This resulted in thirteen different clusters. The clusters and the associated members are shown in (Table S3). The relative cluster stability index (RCSI) and Monte Carlo p-values further validated that the cluster stability (Fig. 1A, B). Put together, it is clear that the TPNs group into 13 broad clusters, each representing a distinct polarization state (Fig. 1C). Further, the cluster member associations inferred from TF-set1 were found to be concordant with the number of clusters (Fig. S1A, B) and clustering pattern (Fig. S2, cophenetic coefficient = 0.68; p-value = 1.25e−51) inferred by considering the entire differential transcriptome (n=12,164). This clearly indicates that the TFs in TF-Set-1 are sufficient to characterize a given polarization state and that the subnetworks from the contextualized hGRNs define the macrophage polarization spectrum.

Functional enrichment analysis of the genes encompassing the regulatory cores of each cluster established their biological relevance. For example, a statistically significant enrichment of innate immune response, apoptosis and signaling pathways associated with proinflammatory response such as NFK-B was seen in Cluster -11 (C11), while a significant enrichment of cell cycle regulation, basic cellular metabolism and anti-inflammatory responses by interleukin signaling was seen in Cluster-3 (C3). Next, we set out to identify combinations of key regulatory factors for each cluster. The subnetworks from which the TF-set1 and the subsequent clustering pattern was derived, serves as precise molecular definitions for each polarization state. Here, we seek to identify minimal gene sets from each cluster for identifying gene panels that are the most influential in guiding the differentiation into specific polarization destinations. Our ultimate goal was to probe if any of them could serve as switching factors between different functional states. The core networks of each cluster provide us connected sets of weighted and directed interactions which capture cascades of regulatory events. The complexity in these subnetworks arise due to the plurality of interactions between TFs since (i) a given TF can regulate multiple targets, (ii) multiple TFs can regulate a given target, and (iii) TFs themselves are regulated by other TFs. A computational pipeline was configured to discover combinations of transcription factors with specific expression patterns, from the subnetworks for each polarization state. The pipeline consists of computational steps to identify (a) common cores for each polarization state, and (b) differentially expressed epicentres of different members of the clusters. The first step was achieved by taking an edge-level intersection of the TPNs within each cluster, which resulted in pruning the subnetworks to retain only those nodes and edges that are common and exhibit the same pattern in all members within each cluster. The

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

pipeline yielded a 52 TF-set characteristic of C11 and a 42 -TF set characteristic of Cluster3 and similarly for all other states. These TF-sets serve as barcodes for each state where color gradient reflects the gene expression changes of each TF (as compared to an unstimulated state) (Fig. 1D). Put together, these barcodes amount to a union of 223 TFs (TF-set2) that can be considered to form a pool of central regulatory factors for the polarization states studied here (TF-set2 - Data file S3).

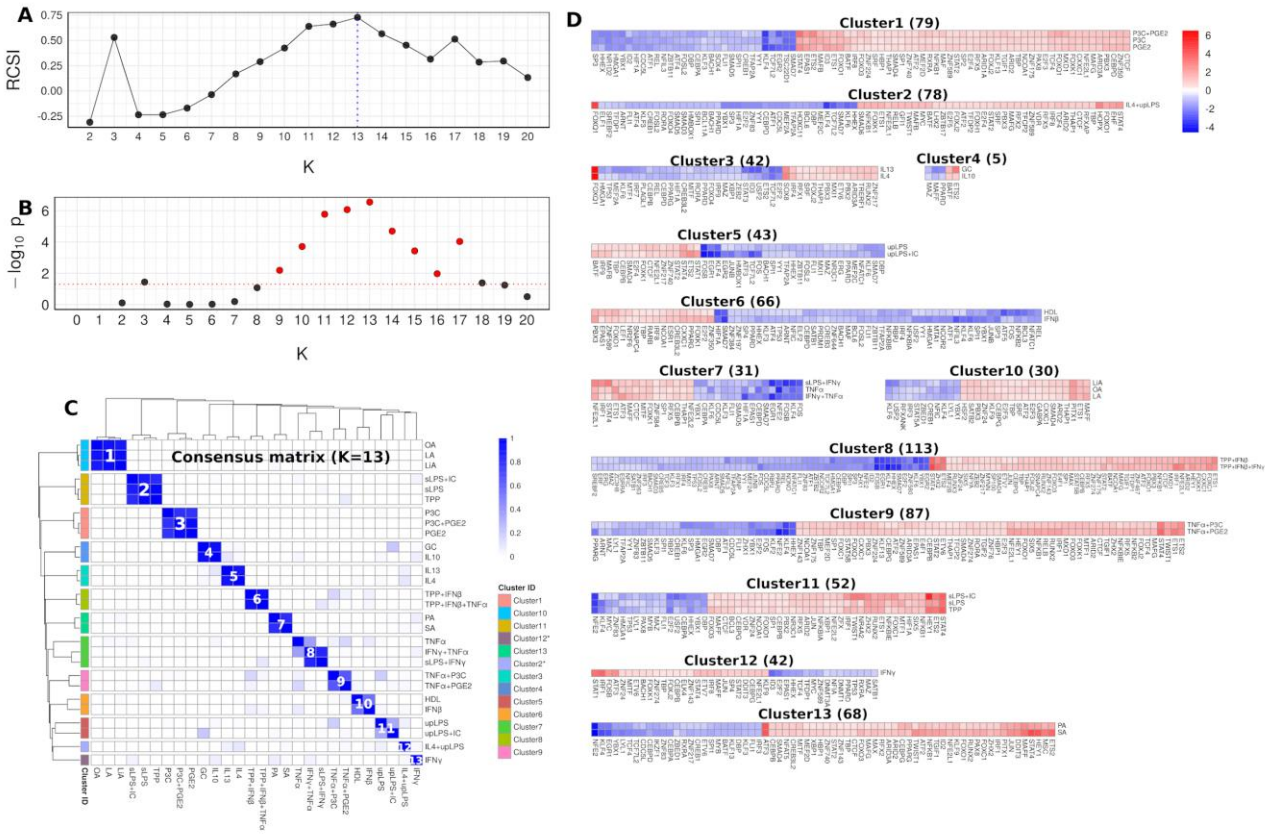


Fig.1. A clustergram showing K=13 for samples spanning the macrophage polarization spectrum and TF-barcodes characteristic of each polarization state. The clustering pattern was obtained by using M3C, a widely used agnostic clustering method based on Monte Carlo referencing that uses stability selection to estimate the number of clusters (John et al., 2020). (A) The robustness metrics, Relative Cluster Stability Index (RCSI) spiked with a significant p-value (B) at K =13. (C) Cluster-member associations in the consensus matrix for K=13. Clusters with a single member are starred. Clusters 2 and 12 are singletons. (D) TF barcodes which are characteristic of the thirteen different clusters. Each barcode is a heatmap of the differential expression profile of the identified transcription factors (as compared to the unstimulated Mo state) in each cluster. The length of the barcode (which is the number of TFs in each cluster) is indicated

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

within parentheses. Barcodes of members within a cluster resemble each other extensively and are shown together.

Identification of a 3-TF set as M1 to M2 switching factors: We investigated if the barcodes could serve as switching factors to tune the polarization destination of macrophages. For this, we selected a stimulant that is known to drive unstimulated macrophages (Mo) to M1 and asked if we could switch it to an M2 destination instead. We start with the barcodes for both states (C11 - representing M1 and C3 representing M2) and pruned them to retain (a) only those genes that occur in the barcodes for both the states and (b) exhibit reversal in the gene expression variation trend between the two states. This resulted in identifying a panel consisting of sixteen TFs (*STAT4*, *ETS2*, *NFKB1*, *BCL3*, *NFE2L2*, *MTF1*, *HIF1A*, *NFKBIA*, *HEY1*, *FOXO3*, *ZFX*, *ZNF24*, *XBP1*, *CEBPB*, *MYC* and *NFE2)* where each gene exhibits a significant quantitative difference in gene expression variation between M1 and M2 states, both with reference to the Mo state (Fig. 2A). Of these, we selected the top-ranked genes based on the extent of gene expression variation in the two states and the number of targets they influenced. We identified two well known TFs: *STAT4*, *NFKB1* and three additional TFs *NFE2L2*, *CEBPB*, *BCL3*, that qualified our criteria to be potential switching factors. *STAT4* and *NFKB1* are well studied and clearly linked to the proinflammatory M1 state and have many targets. We verified that in our study system, they are indeed upregulated upon LPS induction and downregulated upon IL4 treatment (both with reference to Mo, Fig. 2A). We focused our attention on the three newly identified TFs *NFE2L2, CEBPB, BCL3* (referred to as NCB hereafter), which show the same expression trend and are positioned to have a high degree of influence in the network. Put together, the 3-TF combination has 1293 targets, covering about 53% of the top-perturbed network shown in Fig.2B.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

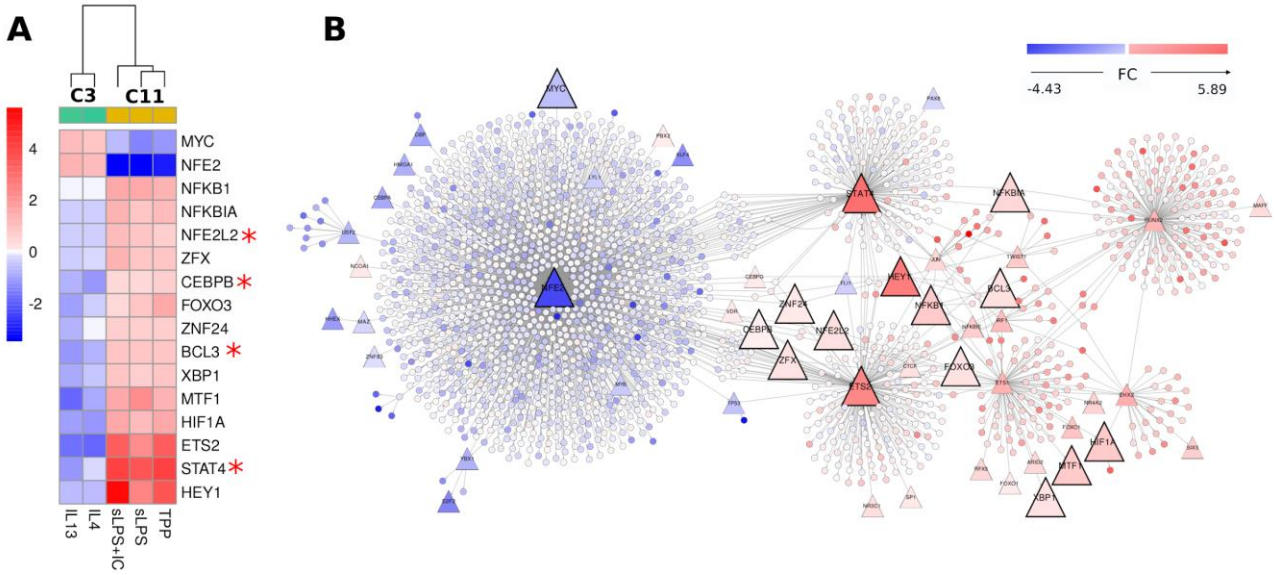
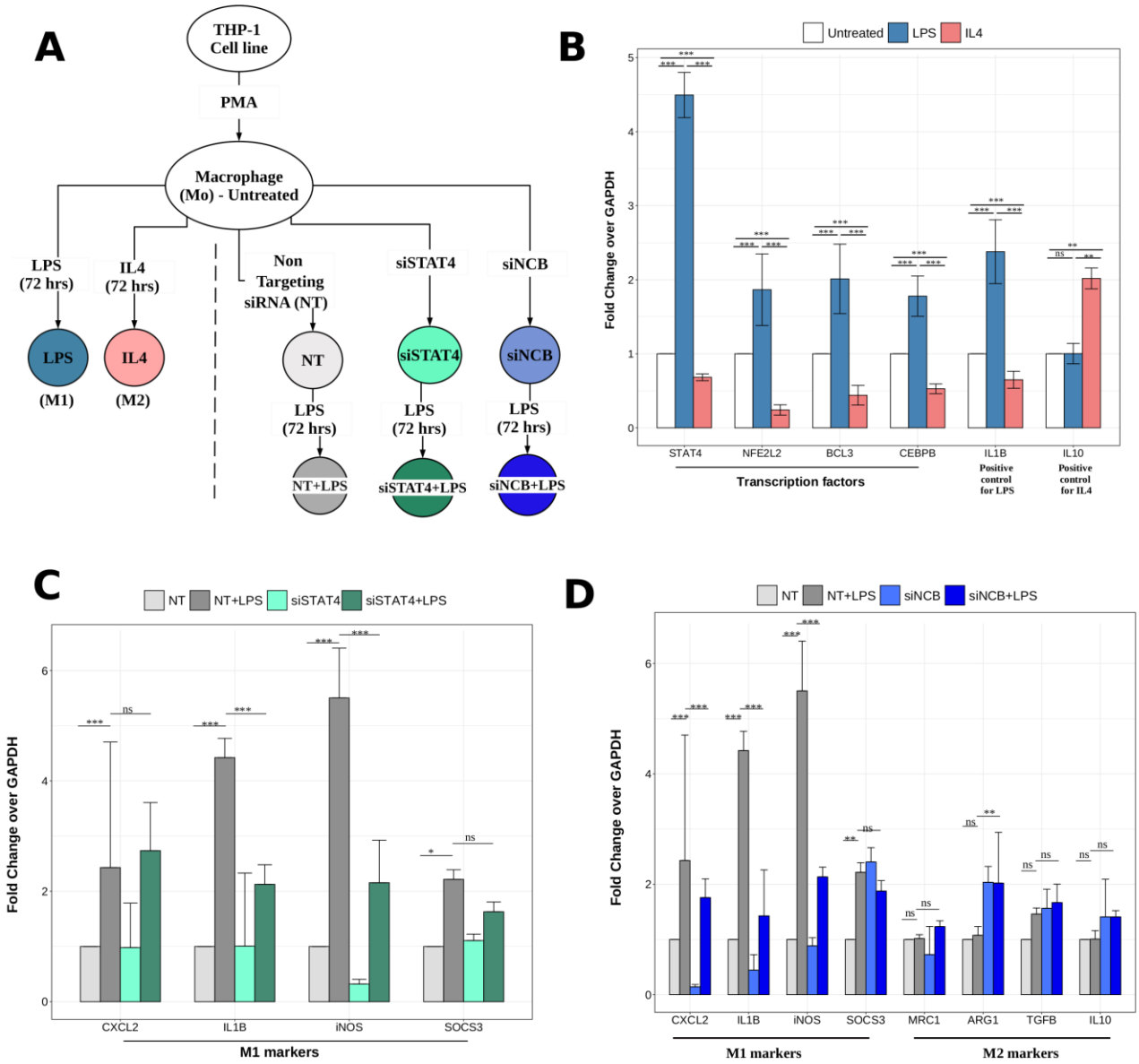


Fig.2. Switching factors that guide M1 to M2 destination. (A) A heatmap of the genes in the barcode that exhibits a reversal in the fold change pattern between M1 (C11) and M2 (C3). TFs that are selected for experimental validation are starred. (B) A regulatory core for the M1 cluster (C11). The core network (nodes: 2028; interactions: 2145) that contains the top perturbed regulatory interactions common across C11 members (sLPS, sLPS+IC and TPP) is shown. Larger triangle nodes represent TFs identified to show switching behaviour; smaller triangles represent TFs present in the C11 barcodes, and circles represent genes common to all C11 members.The nodes are colored based on the fold change (color scale is shown as an inset).

We hypothesized that knock-down of *NFE2L2*, *CEBPB*, *BCL3* in an M1 state should switch it to M2. To test this hypothesis, we designed an experiment (experimental design shown in Fig. 3A) where we take THP-1 derived macrophages (taken as Mo) and stimulate them with sLPS to obtain an M1 state. We first verified that sLPS treatment is leading to an M1 state by measuring known M1 markers (*IL1B* in Fig. 3B), which confirmed that they were indeed in an M1 state. We also measured the levels of M2 markers (*IL10* in Fig.3B) and verified that there was no significant change. *STAT4* is known to have an effect on the polarization state as its knockdown in M1 was shown to significantly attenuate the M1 state (Ishii et al., 2009) and we therefore used *STAT4* as a positive control. We checked whether STAT4 was exhibiting the expected behaviour in our experimental setup and we found that was indeed the case (Fig. S3, Fig. 3C). We then performed siRNA mediated combined knock-down of NCB, the 3-TF set, to evaluate the M1 to M2 switch.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

We verified that the siRNA mediated knockdown (referred to as siNCB) had worked by testing the gene expression values of the 3 NCB genes and finding them to be significantly reduced (Fig. S4, Fig. 3D). Upon stimulation with sLPS in the siRNA treated cells, we observed that the known M1 markers were significantly reduced while the known M2 markers were significantly pronounced, clearly showing that the siRNA knockouts lead the cells to shift from M0 to M2 (Fig. 3D). The change was statistically significant when compared with the parent THP-1 cells without knockdown as well as with non-targeting siRNA controls. In summary, we observe that the knockdown of the 3-gene set leads to a clear shift of the polarization state away from M1 towards M2, despite the sLPS trigger.



bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig. 3. siRNA knockdown of *NFE2L2, CEBPB* and *BCL3* in LPS infected THP-1 derived macrophage cell line attenuates macrophages polarizing to C11 state (M1). (A) Schema describing the experimental design for testing the role of the 3-gene set NCB, predicted to be a switching factor combination between M1 and M2 destinations. The LPS and IL4 treatment conditions are included to verify the experimental setup. The non-targeting siRNA is used as a negative control while siSTAT4 is used as a positive control. siNCB is a knock-down of the 3-gene set (NCB) which is evaluated in this study. (B) Gene expression changes of selected TFs and known markers of M1 (IL1B) and M2 (IL10) in THP1 cell lines upon LPS and IL4 treatment (for 72 hr) respectively. Color coding scheme is similar to that in (A). THP1 monocytes were treated with PMA (20ng/ml) for 16h. Post 24h, LPS (100ng/ml) and IL-4(1000IU) treatment was given for 72 hr. Two-way ANOVA ($P > 0.05$:ns, $P < 0.05$:*, $P < 0.01$:**, $P < 0.001$:*** ; N=3) is used for estimating significance in gene expression fold change. (C) Gene expression changes in the M1 markers upon siRNA knockdown of STAT4 (positive control). THP1 monocytes were treated with PMA (20ng/ml) for 16h. STAT4 specific siRNA was transfected. Post 8h, LPS (100ng/ml) treatment was given for 72h. Gene expression changes in M1 markers were quantified (Two-way ANOVA; $P > 0.05$:ns, $P < 0.05$:*, $P < 0.01$:**, $P < 0.001$:*** ; N=3; NT: non-targeting siRNA) (D) Gene expression changes in the M1 and M2 markers upon siRNA knockdown of the NCB set. THP1 monocytes were treated with PMA (20ng/ml) for 16h. NCB (*NFE2L2*, *CEBPB* , *BCL3*) specific siRNA was transfected. Post 8h, LPS (100ng/ml) treatment was given for 72h. Gene expression changes of M1 and M2 markers were quantified (Two-way ANOVA; $P > 0.05$:ns, $P < 0.05$:*, $P < 0.01$:**, $P < 0.001$:*** ; N=3; NT: non-targeting siRNA)

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

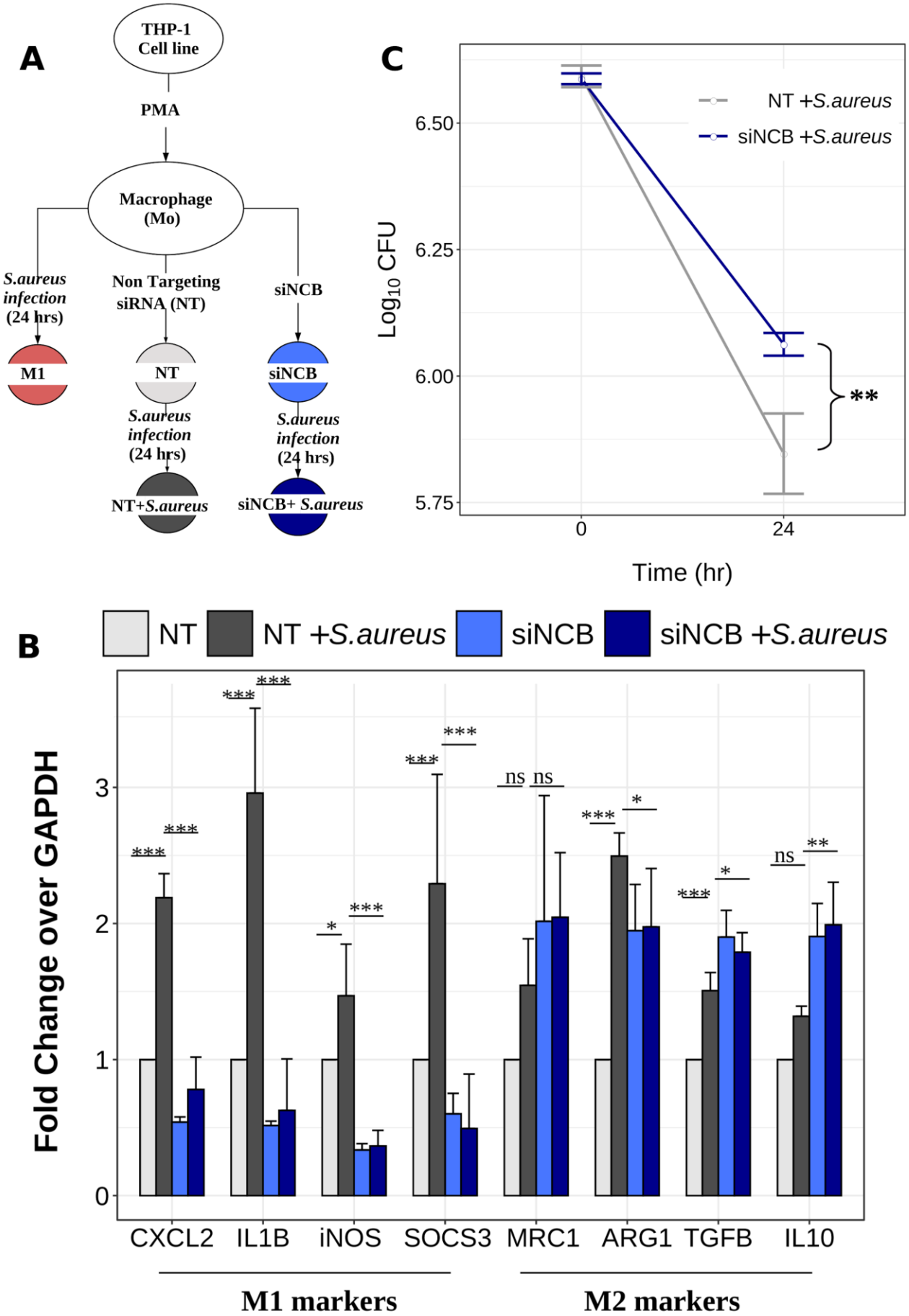
Knockdown of NCB restricts the growth of *S. aureus* within macrophages

S. aureus is a facultative intracellular pathogen causing a broad spectrum of diseases such as minor skin infections to serious bloodstream infections and endocarditis. This pathogen is known to survive inside the host cell and escapes from professional phagocytic cells such as macrophages through various immune evasion strategies, especially through cytoprotective mechanisms by altering the host gene expression (Koziel et al., 2009; Pidwill et al., 2021).

We performed differential analysis of peripheral macrophages infected with S. aureus using the publicly available data (Koziel et al., 2009), and profiled the expression levels of TFs in M1 (C11) barcode and M1 markers. This analysis showed an upregulation of NCB and other TFs in the M1 barcode at the 8hr, 24 hr and 48hr time point (Fig. S5). We then tested the same in an *in vitro* S.aureus infection model using THP-1 macrophages and made a very similar observation, where the mRNA expression levels of the STAT1, NCB, M1, and M2 markers were quantified at 12hr, 24hr, 48hr and 72hr time points (Fig S6, S7). Specifically, the expression levels of NFE2L2, CEBPB, and BCL3 were significantly upregulated in the THP-1 macrophages upon *S. aureus* infection with respect to the uninfected control at 24 hrs (Fig. S7). This suggests that macrophages infected with S.aureus polarize to the M1 state within 24 hrs. Hence, this time point was used for further experiments.

Next, to evaluate the potential of NCB in regulating polarization of macrophages during *S. aureus* infection, we used the same in vitro infection setup but using the NCB knockdown cells. We observed a significant shift from M1 to M2 with the downregulation in the expression of M1 markers (*CXCL2, IL1B, iNOS, SOCS3*) (Fig. 4B, Fig. S8) and upregulation in the expression of M2 markers (*TGFB* and *IL10*) (Fig. 4B). Interestingly, we observed 1.6 fold significant increase in the burden of S.aureus within these macrophages (Fig. 4C). This clearly indicates that knockdown of NCB combination renders the macrophage more conducive for S. aureus to proliferate.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig. 4. siRNA knockdown of NCB (NFE2L2, CEBPB and BCL3) in S.aureus infected macrophage increases infectivity. (A) Schema describing the experimental design. (B) Effect of the NCB set knockdown in the S. aureus infection model. siRNA knockdown of NFE2L2, CEBPB and BCL3 in THP-1 cells infected with S. aureus leads to a significant decrease in the gene expression of the M1 markers (*CXCL2, IL1B, iNOS and SOCS3*) and an increase in the expression of the M2 markers (TGFβ and IL10). Color coding is based on the schema shown in (A). THP1 monocytes were treated with PMA (20ng/ml) for 16hrs. NFE2L2, CEBPB and BCL3 specific siRNAs were transfected. Post 24hrs, S.aureus infection at 1:10 MOI was given for 24hrs. Gene expression levels of M1 and M2 markers were assessed. (Two-way ANOVA; $P > 0.05$:ns, $P < 0.05$:*, $P < 0.01$:**, $P < 0.001$:*** ; N=3; NT: non-targeting siRNA). (C) In vitro S.aureus CFU assay. THP-1 cells were differentiated to macrophages with PMA for 16 hrs. Post 24 hrs, siRNA specific for NFE2L2, BCL3 and CEBPB were transfected. S.aureus infection (MOI 10) was given for 2 hrs prior to 2 hrs gentamicin treatment. CFU assay was performed at 0 hr and 24 hrs. (Two-way ANOVA; $P > 0.05$:ns, $P < 0.05$:*, $P < 0.01$:**, $P < 0.001$:*** ; N=3; NCB : *NFE2L2, CEBPB and BCL3*). CFU-Colony Forming Units.

Discussion

Macrophages display a wide array of functions, from being killers of pathogens to healers of damaged tissue (Murray, 2017; Murray and Wynn, 2011; Wynn et al., 2013). Their versatility is attributed to their ability to respond to specific physiological or pathological cues, leading to their polarization into different states, each with a distinct functional phenotype. A range of molecular triggers that cause differentiation in macrophages is known (Hu et al., 2021; Jaguin et al., 2013; Xue et al., 2014). It is also well known that the differentiation spans a full spectrum of phenotypes with the classically activated M1 and the alternatively activated M2 states at two ends of the spectrum. The molecular profiles that define or cause these states, however, have not been sufficiently characterized. In particular, very little is known about the regulatory genes that control and guide the polarization in a specific direction. It is of great interest to identify TFs that are most influential in guiding the polarization. Given the importance of macrophages, several studies have investigated the variations in the transcriptomes in different diseases (Benoit et al., 2008; Mosser and Edwards, 2008; Ross et al., 2021; Schultze, 2016; Zheng et al., 2018) as well as upon

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

exposure of macrophages to different exogenous and endogenous stimuli (Hu et al., 2021; Jaguin et al., 2013; Xue et al., 2014), which have resulted in a wealth of data that are publicly accessible. These large datasets provide opportunities for addressing a variety of questions.

In this study, we algorithmically identify distinct polarization states, minimal regulatory signatures for each state as transcription factor barcodes and demonstrate that manipulation of the barcode leads to loss of switching, in an example case. Additionally, our work has indicated that arresting the switch to M1 leads to hypersusceptibility to *S.aureus* infection in cell lines, suggesting that M1 has a distinct role in limiting the extent of infection. Two aspects in our computational workflow are distinct from the previous methods, which facilitated the identification of barcodes characteristic of different macrophage polarization states. The first is the use of an unbiased knowledge-based gene regulatory network as compared to the more common correlation-based interaction networks. In the latter, the interactions are statistical associations, with no guarantee that the interaction is 'real' in the given biological system (Petri et al., 2015; Ravichandran and Chandra, 2019). We overcome this limitation by consolidating a comprehensively curated knowledge-based network where each interaction is based on experimental data reported in databases and in primary literature. We have previously carried out a systematic comparison and have shown that knowledge-based networks are superior to correlation-based networks in capturing biologically relevant genes and pathways (Ravichandran and Chandra, 2019). Our macrophage specific GRN (subset of hGRN) network contains 326, 281 interactions among 742 TFs and 11, 423 target genes, which is more comprehensive than the previously reported knowledge based macrophage GRNs in literature (Hörhold et al., 2020; Palma et al., 2018). The second is the network interrogation method. Our method of contextualizing the knowledge-based network for diverse stimulants using the corresponding transcriptomes allowed a quantitative comparison of variations over a common background network, yielding a top-perturbed regulatory subnetwork in each case. Our network interrogation method identifies the most influential 'epicentric' transcription factors, which lead to the derivation of quantitative barcodes of TF panels. To the best of our knowledge, such barcodes have not been reported earlier even for the well studied M1 and M2 states. We are hopeful our results of epicentric TFs and their molecular interaction trajectories will serve as key inputs for several investigations on the TFs individually as well as for specific TF panels. The response networks for each stimulant also serves as a rich resource of the top-ranked regulatory pathways in each case.

As a first step into the identification of barcodes, we used a pooled set of differentially expressed epicentric TFs (n=265) in different conditions that we identify from the networks and perform an

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

unbiased clustering of the macrophages. We used a rigorous clustering method, which produced a robust clustering pattern which clearly indicated that the 28 stimulants group into thirteen polarization states. A clustering pattern was described earlier for the same dataset based on the top 1000 most varying DEGs from each condition, which resulted in nine distinct states (Xue et al., 2014). The clustering pattern that we get here from the top-ranked epicentres in each, yielded a similar pattern, indicating that a subset of DEGs that occur in our topnets are sufficient to achieve the differentiation. In fact, our gene-regulatory models implied that the saturated (SA and PA) and unsaturated (LA, LiA, and, OA) fatty acids mediate different modes of resolution that were grouped together in the previous clustering pattern, and are now resolved into two sub-branches. Similarly, the effects of IFNγ and sLPS which were clubbed in the previous pattern are now resolved, consistent with the known regulatory differences (Hoeksema et al., 2015; Kang et al., 2019). The exact number of clusters is not the main message from this analysis. Instead, it is deriving a molecular basis for the differentiation of the functional states, of which M1 and M2 are two ends of the spectrum, while several other states are located in between the two in the polarization state spectrum, which can be described as a punctuated continuum.

Our results suggest that the barcodes could serve as the regulatory handles to attain different polarization states with different functional phenotypes. We test that hypothesis for the M1/M2 axis and indeed demonstrate that siRNA knockdown of the TFs in the barcode achieves polarization state switching from M1 to M2 polarization states. In the M1 barcode, we selected the TFs upregulated as compared to the baseline (M_0): (NBC- Nuclear Factor, Erythroid 2 Like 2 (*NFE2L2*), B-Cell Lymphoma-3 (*BCL3*), and CCAAT/enhancer-binding protein beta (*CEBPB*)) and by knocking down the expression of these transcription factors in LPS treated THP-1 derived macrophage cell line, the cells were clearly seen not only to have not attained an M1 state but to have shifted towards the M2 state. This establishes their role as a switch between the states- where an up-regulation moves them from an M_0 to M1 while a knock-down shifts them from M_0 to M2, even after exposure to a M1-stimulant, thus indicating them to be essential for establishing the M1 state. After having established an arrest of polarization to M1, we were curious to see if it resulted in a functional phenotype. Our experiments with *S.aureus* infection, clearly showed that the NBC knockdown led to increased infectivity of *S.aureus*. Expression of these TFs and their target genes are involved in cellular response to stress, Tumor Necrosis Factor (TNF) signaling pathway and senescence associated secretory phenotype, suggesting that inhibition of inflammatory associated processes halts M1 phenotype and promotes the M2-like Phenotype, which renders them hyper-susceptible to *S.aureus* infection.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

In conclusion, we report a systems level analysis of the gene regulatory networks to identify regulatory factors which define different macrophage activation states. Perturbation of these factors for reprogramming the macrophage subpopulations to desired polarization states can be explored for applications in controlling the inflammation gradient in different diseases such as sepsis, autoimmune conditions and chronic infections.

Author contributions: SR and NC conceptualized the study. KNB planned the validation experiments. SR carried out the data curation, computational methodology, analysis, and interpretation. BB carried out the experiments. NC and KNB acquired funding for the research. NC supervised the whole project. SR and NC wrote the manuscript with inputs from KNB and BB. All authors read and approved the final manuscript.

Declaration of interests: NC is a co-founder of qBiome Pvt. Ltd and HealthSeq Precision Medicine Pvt Ltd, which had no role in this manuscript. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements: This work was supported by funds from the Department of Science and Technology- Fund for Improvement of S&T Infrastructure in Higher Educational Institutions (DST-FIST), the University Grants Commission and the Department of Biotechnology (DBT), Govt of India; (DBT)-Indian Institute of Science partnership program (Phase-II at IISc, BT/PR27952/INF/22/212/2018); (DBT, No. BT/PR13522/COE/34/27/2015 and BT/PR27352/BRB/10/1639/2017 to KNB)) and the Department of Science and Technology (DST, EMR/2014/000875 to KNB). KNB thanks Science and Engineering Research Board (SERB), DST for the award of J. C. Bose National Fellowship (No. SB/S2/JCB-025/2016 dt 25/7/15) and core research grant (CRG/2019/002062); and for the funding (SP/DSTO-19-0176 dt 06/02/2020).

Supplementary materials

Fig S1. M3C based identification of the optimal number of clusters for the entire transcriptome.

Fig S2. Clustering based on TF-set1 Vs. Entire transcriptome.

Fig S3. Validation of STAT4 siRNA.

Fig S4. Validation of NCB siRNA.

Fig S5. S. auerus infected macrophages polarize to M1.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig S6. Gene expression changes in M1 and M2 markers upon S.aureus infection at different time points.

Fig S7. Gene expression changes in STAT4 and NCB-set upon S.aureus infection at different time points.

Fig S8. NCB (NFE2L2, CEBPB, BCL3) siRNA validation in S.aureus data.

Supplementary tables

Table S1. In vitro stimulant data.

Table S2. The list of primers used for the qRT-PCT

Table S3. The list of clusters, associated members, network details of regulatory cores and the number of TFs in the barcodes.

Supplementary data file

Data file S1: The human Gene Regulatory Network.

Data file S2: Twenty eight different subnetworks capturing highest ranked regulatory perturbations.

Data file S3: TF-set1, TF-set2

References

Aderem, A. (2003). Phagocytosis and the Inflammatory Response. The Journal of Infectious Diseases *187*, S340-5.

Ardura, J.A., Rackov, G., Izquierdo, E., Alonso, V., Gortazar, A.R., and Escribese, M.M. (2019). Targeting Macrophages: Friends or Foes in Disease? Frontiers in Pharmacology *10*, 1255.

Atri, C., Guerfali, F.Z., and Laouini, D. (2018). Role of Human Macrophage Polarization in Inflammation during Infectious Diseases. Int J Mol Sci *19*, E1801.

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., et al. (2013). NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res *41*, D991–D995.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological) *57*, 289–300.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Benoit, M., Desnues, B., and Mege, J.-L. (2008). Macrophage polarization in bacterial infections. J Immunol *181*, 3733–3739.

Bovolenta, L.A., Acencio, M.L., and Lemke, N. (2012). HTRIdb: an open-access database for experimentally verified human transcriptional regulation interactions. BMC Genomics *13*, 405.

Carvalho, B.S., and Irizarry, R.A. (2010). A framework for oligonucleotide microarray preprocessing. Bioinformatics *26*, 2363–2367.

Fabregat, A., Sidiropoulos, K., Viteri, G., Forner, O., Marin-Garcia, P., Arnau, V., D'Eustachio, P., Stein, L., and Hermjakob, H. (2017). Reactome pathway analysis: a high-performance in-memory approach. BMC Bioinformatics *18*, 142.

Galili, T. (2015). dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. Bioinformatics *31*, 3718–3720.

Gautier, L., Cope, L., Bolstad, B.M., and Irizarry, R.A. (2004). affy--analysis of Affymetrix GeneChip data at the probe level. Bioinformatics *20*, 307–315.

Gerrick, K.Y., Gerrick, E.R., Gupta, A., Wheelan, S.J., Yegnasubramanian, S., and Jaffee, E.M. (2018). Transcriptional profiling identifies novel regulators of macrophage polarization. PLoS One *13*.

Greenberg, S., and Grinstein, S. (2002). Phagocytosis and innate immunity. Curr Opin Immunol *14*, 136–145.

Hamilton, J.A., and Tak, P.P. (2009). The dynamics of macrophage lineage populations in inflammatory and autoimmune diseases. Arthritis Rheum *60*, 1210–1221.

Han, H., Shim, H., Shin, D., Shim, J.E., Ko, Y., Shin, J., Kim, H., Cho, A., Kim, E., Lee, T., et al. (2015). TRRUST: a reference database of human transcriptional regulatory interactions. Sci Rep *5*, 11432.

Hoeksema, M.A., Scicluna, B.P., Boshuizen, M.C.S., Velden, S. van der, Neele, A.E., Bossche, J.V. den, Matlung, H.L., Berg, T.K. van den, Goossens, P., and Winther, M.P.J. de (2015). IFN-γ Priming of Macrophages Represses a Part of the Inflammatory Program and Attenuates Neutrophil Recruitment. The Journal of Immunology *194*, 3909–3916.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Hörhold, F., Eisel, D., Oswald, M., Kolte, A., Röll, D., Osen, W., Eichmüller, S.B., and König, R. (2020). Reprogramming of macrophages employing gene regulatory and metabolic network models. PLOS Computational Biology *16*, e1007657.

Hu, G., Su, Y., Kang, B.H., Fan, Z., Dong, T., Brown, D.R., Cheah, J., Wittrup, K.D., and Chen, J. (2021). High-throughput phenotypic screen and transcriptional analysis identify new compounds and targets for macrophage reprogramming. Nat Commun *12*, 773.

Jaguin, M., Houlbert, N., Fardel, O., and Lecureur, V. (2013). Polarization profiles of human M-CSF-generated macrophages and comparison of M1-markers in classically activated macrophages from GM-CSF and M-CSF origin. Cell Immunol *281*, 51–61.

John, C.R., Watson, D., Russ, D., Goldmann, K., Ehrenstein, M., Pitzalis, C., Lewis, M., and Barnes, M. (2020). M3C: Monte Carlo reference-based consensus clustering. Sci Rep *10*, 1816.

Kang, K., Bachu, M., Park, S.H., Kang, K., Bae, S., Park-Min, K.-H., and Ivashkiv, L.B. (2019). IFN-γ selectively suppresses a subset of TLR4-activated genes and enhancers to potentiate macrophage activation. Nat Commun *10*, 3320.

Keenan, A.B., Torre, D., Lachmann, A., Leong, A.K., Wojciechowicz, M.L., Utti, V., Jagodnik, K.M., Kropiwnicki, E., Wang, Z., and Ma'ayan, A. (2019). ChEA3: transcription factor enrichment analysis by orthogonal omics integration. Nucleic Acids Res *47*, W212–W224.

Koziel, J., Maciag-Gudowska, A., Mikolajczyk, T., Bzowska, M., Sturdevant, D.E., Whitney, A.R., Shaw, L.N., DeLeo, F.R., and Potempa, J. (2009). Phagocytosis of Staphylococcus aureus by macrophages exerts cytoprotective effects manifested by the upregulation of antiapoptotic factors. PloS One *4*, e5210.

Lawrence, T. (2011). Transcriptional regulation of macrophage polarization: enabling diversity with identity. Nat. Rev. Immunol. 750–761.

Li, H., Jiang, T., Li, M.-Q., Zheng, X.-L., and Zhao, G.-J. (2018). Transcriptional Regulation of Macrophages Polarization by MicroRNAs. Frontiers in Immunology *9*, 1175.

Liu, Z.-P., Wu, C., Miao, H., and Wu, H. (2015). RegNetwork: an integrated database of transcriptional and post-transcriptional regulatory networks in human and mouse. Database (Oxford) *2015*.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Mantovani, A., Sica, A., Sozzani, S., Allavena, P., Vecchi, A., and Locati, M. (2004). The chemokine system in diverse forms of macrophage activation and polarization. Trends in Immunology *25*, 677–686.

Martinez, F.O., and Gordon, S. (2014). The M1 and M2 paradigm of macrophage activation: time for reassessment. F1000Prime Reports *6*.

Mills, C.D., Kincaid, K., Alt, J.M., Heilman, M.J., and Hill, A.M. (2000). M-1/M-2 macrophages and the Th1/Th2 paradigm. Journal of Immunology (Baltimore, Md.: 1950) *164*, 6166–6173.

Mosser, D.M., and Edwards, J.P. (2008). Exploring the full spectrum of macrophage activation. Nat Rev Immunol *8*, 958–969.

Mosser, D.M., Hamidzadeh, K., and Goncalves, R. (2021). Macrophages and the maintenance of homeostasis. Cellular & Molecular Immunology *18*, 579–587.

Murray, P.J. (2017). Macrophage Polarization. Annual Review of Physiology *79*, 541–566.

Murray, P.J., and Wynn, T.A. (2011). Protective and pathogenic functions of macrophage subsets. Nature Reviews. Immunology *11*, 723–737.

Palma, A., Jarrah, A.S., Tieri, P., Cesareni, G., and Castiglione, F. (2018). Gene Regulatory Network Modeling of Macrophage Differentiation Corroborates the Continuum Hypothesis of Polarization States. Front Physiol *9*, 1659.

Petri, T., Altmann, S., Geistlinger, L., Zimmer, R., and Küffner, R. (2015). Addressing false discoveries in network inference. Bioinformatics *31*, 2836–2843.

Pidwill, G.R., Gibson, J.F., Cole, J., Renshaw, S.A., and Foster, S.J. (2021). The Role of Macrophages in Staphylococcus aureus Infection. Frontiers in Immunology *11*, 3506.

Ravichandran, S., and Chandra, N. (2019). Interrogation of genome-wide networks in biology: comparison of knowledge-based and statistical methods. Int J Adv Eng Sci Appl Math *11*, 119–137.

Ravichandran, S., Banerjee, U., DR, G.D., Kandukuru, R., Thakur, C., Chakravortty, D., Balaji, K.N., Singh, A., and Chandra, N. (2021). VB10, a new blood biomarker for differential diagnosis and recovery monitoring of acute viral and bacterial infections. EBioMedicine *67*, 103352.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47.

Ross, E.A., Devitt, A., and Johnson, J.R. (2021). Macrophages: The Good, the Bad, and the Gluttony. Frontiers in Immunology *12*, 3234.

Rőszer, T. (2015). Understanding the Mysterious M2 Macrophage through Activation Markers and Effector Mechanisms. Mediators Inflamm *2015*.

Rouillard, A.D., Gundersen, G.W., Fernandez, N.F., Wang, Z., Monteiro, C.D., McDermott, M.G., and Ma'ayan, A. (2016). The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. Database (Oxford) *2016*.

Sambaturu, N., Mishra, M., and Chandra, N. (2016). EpiTracer - an algorithm for identifying epicenters in condition-specific biological networks. BMC Genomics *17*, 543.

Sambaturu, N., Pusadkar, V., Hannenhalli, S., and Chandra, N. (2021). PathExt: a general framework for path-based mining of omics-integrated biological networks. Bioinformatics *37*, 1254–1262.

Schultze, J.L. (2016). Reprogramming of macrophages—new opportunities for therapeutic targeting. Current Opinion in Pharmacology *26*, 10–15.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

Sica, A., and Mantovani, A. (2012). Macrophage plasticity and polarization: in vivo veritas. J Clin Invest *122*, 787–795.

Stout, R.D., Jiang, C., Matta, B., Tietzel, I., Watkins, S.K., and Suttles, J. (2005). Macrophages sequentially change their functional phenotype in response to changes in microenvironmental influences. J Immunol *175*, 342–349.

Tugal, D., Liao, X., and Jain, M.K. (2013). Transcriptional Control of Macrophage Polarization. Arteriosclerosis, Thrombosis, and Vascular Biology *33*, 1135–1144.

Wingender, E., Schoeps, T., and Dönitz, J. (2013). TFClass: an expandable hierarchical

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.19.464946; this version posted October 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

classification of human transcription factors. Nucleic Acids Res *41*, D165–D170.

Wynn, T.A., Chawla, A., and Pollard, J.W. (2013). Macrophage biology in development, homeostasis and disease. Nature *496*, 445–455.

Xue, J., Schmidt, S.V., Sander, J., Draffehn, A., Krebs, W., Quester, I., De Nardo, D., Gohel, T.D., Emde, M., Schmidleithner, L., et al. (2014). Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. Immunity *40*, 274–288.

Yao, Y., Xu, X.-H., and Jin, L. (2019). Macrophage Polarization in Physiological and Pathological Pregnancy. Front Immunol *10*.

Zhang, X., and Mosser, D. (2008). Macrophage activation by endogenous danger signals. J Pathol *214*, 161–178.

Zhao, C., Medeiros, T.X., Sové, R.J., Annex, B.H., and Popel, A.S. (2021). A data-driven computational model enables integrative and mechanistic characterization of dynamic macrophage polarization. IScience *24*.

Zheng, P., Luo, Q., Wang, W., Li, J., Wang, T., Wang, P., Chen, L., Zhang, P., Chen, H., Liu, Y., et al. (2018). Tumor-associated macrophages-derived exosomes promote the migration of gastric cancer cells by transfer of functional Apolipoprotein E. Cell Death Dis *9*, 1–14.