## **Cell Systems**

## **Quantifying stimulus-response specificity to probe** the functional state of macrophages

### **Graphical abstract**



### **Authors**

Katherine M. Sheu, Aditya A. Guru, Alexander Hoffmann

### Correspondence

ahoffmann@ucla.edu

### In brief

Emerging evidence suggests that macrophages are capable of immune threat-appropriate responses. To quantify response specificity, we introduce experimental and analytical workflows. Our studies reveal alterations by microenvironmental cytokines *in vitro* and within unhealthy mice, suggesting a means to assess the functional state of macrophages.

### **Highlights**

- Macrophages mount pathogen/stimulus-specific responses, but how specific are they?
- An experimental and computational workflow to quantify response specificity
- Response specificity is distinctively altered by microenvironmental cytokines
- The response specificity index quantifies macrophage functional states in old and obese mice



## **Cell Systems**



## Article

# Quantifying stimulus-response specificity to probe the functional state of macrophages

Katherine M. Sheu,<sup>1</sup> Aditya A. Guru,<sup>1</sup> and Alexander Hoffmann<sup>1,2,\*</sup>

<sup>1</sup>Department of Microbiology, Immunology, and Molecular Genetics, and Institute for Quantitative and Computational Biosciences, University of California, Los Angeles, 611 Charles E. Young Dr S, Los Angeles, CA 90093, USA

<sup>2</sup>Lead contact

\*Correspondence: ahoffmann@ucla.edu https://doi.org/10.1016/j.cels.2022.12.012

SUMMARY

Immune sentinel macrophages initiate responses to pathogens via hundreds of immune response genes. Each immune threat demands a tailored response, suggesting that the capacity for stimulus-specific gene expression is a key functional hallmark of healthy macrophages. To quantify this property, termed "stimulus-response specificity" (SRS), we developed a single-cell experimental workflow and analytical approaches based on information theory and machine learning. We found that the response specificity of macrophages is driven by combinations of specific immune genes that show low cell-to-cell heterogeneity and are targets of separate signaling pathways. The "response specificity profile," a systematic comparison of multiple stimulus-response distributions, was distinctly altered by polarizing cytokines, and it enabled an assessment of the functional state of macrophages. Indeed, the response specificity profile of peritoneal macrophages from old and obese mice showed characteristic differences, suggesting that SRS may be a basis for measuring the functional state of innate immune cells.

A record of this paper's transparent peer review process is included in the supplemental information.

#### INTRODUCTION

Macrophages reside in almost all tissues of the body, where they are sensors for injury, infection, or disease.<sup>1</sup> The many functions they perform require them to respond appropriately to pathogens (pathogen-associated molecular patterns [PAMPs]), injury and danger (danger-associated molecular patterns [DAMPs]), and to cytokines. Immune response genes code for potent bioactivities that are not constitutively expressed because they may be detrimental to the host. Thus, immune response gene programs should only be deployed as needed and to the extent necessary. In fact, the precise deployment of immune response genes is critical for preventing abnormal immune sequelae. Both weak or overactive immune responses may arise from the failure of immune sentinel cells to respond with appropriate specificity to immune threats, resulting in poor health outcomes.<sup>2–5</sup>

Since immune sentinel cells function as individuals in initiating and coordinating immune responses, measuring their capacity to respond specifically to diverse pathogens requires consideration of their substantial cell-to-cell heterogeneity.<sup>6</sup> Prior population-level macrophage transcriptome profiling studies identified a common immune response gene program,<sup>7</sup> as well as gene programs that were in fact highly stimulus specific.<sup>8–10</sup> A plethora of population-level studies have shown that specific responses arise from the signaling and epigenetic networks downstream of receptor-ligand interactions.<sup>11–13</sup> In addition, several studies measuring time-dependent stimulus responses using transcription factor knockouts have delineated the regulation of stimulus-dependent gene programs.<sup>8,13</sup> However, these studies did not quantify how specific macrophages responses actually are, as this requires single-cell measurements to determine the breadth and overlap of single-cell response distributions.

Macrophage heterogeneity itself has recently been studied via single-cell measurements in multiple contexts, such as during polarization or differentiation. For instance, the heterogeneity of macrophage populations after exposure to polarizing cytokines was shown to be greater if simultaneous conflicting cues were provided.<sup>14</sup> Single-cell heterogeneity has also been profiled during monocyte-to-macrophage differentiation in different tissues under different infection conditions, uncovering distinct activation paths in vivo that are altered by microenvironment and disease.<sup>15</sup> Importantly, macrophages must function as responders to immune threats, and thus recent studies have investigated the heterogeneity of single-cell macrophage or monocyte responses to single stimuli, after polarization<sup>16</sup> or after immune training.<sup>17</sup> However, these studies have not yet addressed how the heterogeneity of macrophage gene expression affects the specificity of responses to different to stimuli, and how the resulting stimulus-response specificity may be affected by different microenvironmental contexts.

The response of immune sentinels to stimuli is known to be a function of the microenvironmental context.<sup>18</sup> Polarizing cytokines enhance and diminish the activation of specific immune



Cell Systems Article

genes in an immune stimulus-specific manner by altering signaling networks and epigenetic landscapes.<sup>19–22</sup> Thus, aberrant cytokine contexts that are associated with inflammatory or immunological disease may affect the macrophage's capacity to mount stimulus-specific responses. Indeed, as macrophages circulate and patrol the body, profiling their capacities to mount stimulus-specific responses may report on the inflammatory state of their donor. However, in the absence of a workflow that measures response distributions and quantifies their overlap, the opportunity to leverage the macrophage's sensitivity to immune cytokines for reporting on the function of an individual's innate immune system has not been explored.

Here, we developed the necessary single-cell experimental and computational approaches to assess the functional states of macrophages, via quantification of the specificity of their responses to immune threats. Using information theoretic and machine learning approaches, we found that stimulus-response specificity (SRS) was driven by genes with narrow response distributions that in combination distinguished different ligands. Mechanistically, high SRS was associated with stimulus-specific activation of interferon regulatory factor (IRF) or MAPKp38 pathways, or with differences in the dynamical profiles of NF-κB signaling. We found that SRS was affected by microenvironmental cytokines in a gene-specific manner. This realization prompted the development of the response specificity profile, involving systematic pairwise comparisons, to reveal important functional distinctions in macrophages polarized in vitro or conditioned in vivo by inflammatory conditions of age or obesity.

#### RESULTS

#### Despite single-cell heterogeneity, macrophages produce highly specific gene expression responses to diverse immune stimuli

To quantify the degree of response specificity in macrophages, we developed an experimental workflow using a targeted mRNA sequencing approach<sup>23</sup> that was both cost-effective and reduced the technical noise of genome-wide single-cell RNA sequencing (scRNA-seq) approaches (Figure 1A). We selected a set of 500 macrophage genes via principal-component analysis (PCA) on available macrophage response RNAseg data, which profiled macrophage responses across a time series at 0, 1, 3, and 8 h in response to 14 different viral and bacterial immune components or whole pathogens.<sup>8</sup> The gene loadings of the resulting PCA identified the most stimulus-specific genes in an unsupervised manner (Figure S1A; Table S1; STAR Methods). This bulked RNA-seq dataset. The selected set of 500 genes showed greater enrichment of NF-κB, IRF, and AP1 motifs than 1,502 stimulus-induced genes, which suggested that these three signaling pathways are the primary drivers of response specificity (Figure S1B).

To identify immune stimuli that would best represent response specificity, we further analyzed bulk RNA-seq data from macrophages responding to 14 different pathogen or cytokine ligands to determine the ligands that induce diverse macrophage responses (Figure S1A).<sup>8</sup> Tensor component analysis<sup>24</sup> allowed for an integrated decomposition of the data across all measured time points and stimuli, and showed that each ligand occupied a non-redundant location, indicating a distinct transcriptomic

response at the population level. However, some ligands such as CpG and Pam3CSK (P3C) sat closely adjacent in the tensor-decomposed space (Figure S1C), indicating a greater similarity in their time-dependent expression profiles. From the 14 stimuli, we selected 6 with distinct tensor component weights that represented a spectrum of Gram-positive bacteria (P3C), Gram-negative bacteria (lipopolysaccharide [LPS]), bacterial DNA (CpG), viral nucleic acids (poly(I:C) [PIC]), and host cyto-kines activating either the interferon (IFN- $\beta$ ) or NF- $\kappa$ B (TNF) signaling pathways.

Because gene programs are induced dynamically, the quantification of SRS may depend on the time point at which gene expression measurements are taken. To compare different time points of stimulation, we calculated pairwise stimuli distances at 1, 3, and 8 h. Ligand-pair distances were most distinct at 3 h (Figure S1D), reflecting less impact from secondary signaling mechanisms than at the 8-h time point, and thus we chose the 3-h time point for our analysis of response specificity. A heatmap comparison of the newly generated single-cell data showed good concordance with the published bulk RNA-seq data while revealing substantial cell-to-cell heterogeneity in expression (Figure 1B).

A comparison of the scRNA-seq data from our targeted approach (Rhapsody) to the genome-wide approach (10x Genomics) showed that both had similar concordance in their means to bulk data (Figure S2A) and had comparable distributions to each other (Figure S2B). Notably, the majority of genes captured only by the genome-wide approach were poorly expressed, further supporting the utility of sequencing just the most informative genes through the targeted approach (Figure S2C). For  $\sim$ 90% of the genes measured by both approaches, the targeted approach also had a smaller percentage of cells with drop-out data (i.e., genes with zero reads), as each transcript could be sequenced more deeply (Figure S2D). To avoid heterogeneity in macrophage populations due to different progenitors present in bone marrow, we used a clonal HoxB4-immortalized myeloid progenitor cell line that was differentiated into macrophages with macrophage colony-stimulating factor (M-CSF)-containing medium. The resulting cell populations showed similar bulk<sup>25</sup> and single-cell stimulus-response transcriptomic profiles as bone marrow-derived macrophages (Figures S3A and S3B). Taken together, these results suggest that the targeted sequencing approach provides a reproducible (Figures S3C and S3D; STAR Methods), cost-effective means for measuring heterogeneous single-cell macrophage gene expression in response to diverse immune ligands, making it suitable for quantifying SRS.

To determine how much single-cell heterogeneity affected the stimulus specificity of responses, we next sought to assess the overlap in gene expression response distributions. PCA revealed that IFN- $\beta$ , LPS, and PIC-response distributions were best distinguished, with minimal overlap of their 95% confidence regions on the first two components (44% variance explained), while TNF, CpG, and P3C appeared overlapping (Figure 1C). Uniform manifold approximation and projection (UMAP) on the top 20 components clarified that TNF could be separated from CpG and P3C on lower components (Figure 1D), potentially because the latter activate stronger MAPK and non-oscillatory NF- $\kappa$ B, while the MyD88-mediated CpG and P3C response distributions remained largely indistinguishable (Figure 1E). We



#### Figure 1. Macrophages produce highly specific gene expression responses to diverse immune stimuli

(A) A cost-effective experimental workflow using targeted single-cell RNA sequencing profiles immune response gene expression at single-cell resolution. Bottom left: heatmap of all genes induced with  $log_2(FC) > 2$  in any stimulus from bulk RNA-seq data.<sup>8</sup> Bottom middle: PCA loadings from bulk RNA-seq data. Bottom right: panel genes were selected by scoring each gene based on the summed strength of the loadings from the top 20 principal components. Top right: all genes are captured, and selected genes are amplified via a targeted primer panel.

(B) Comparison of bulk RNA-seq data and single-cell data shows concordance of gene expression clusters and also reveals single-cell heterogeneity. Color bars represent Z scores.

(C) PCA on macrophages stimulated for 3 h indicates there may be a high degree of stimulus specificity despite single-cell heterogeneity, but TNF, P3C, and CpG appear overlapping on the first two components. Ellipses represent 95% confidence intervals based on multivariate t distribution. Colors are the same as in (A). (D) UMAP of stimulated macrophages using the top 20 principal components clarifies that TNF can be separated, but P3C and CpG response distributions remain overlapping.

(E) Calculation of the Bhattacharya distance between pairs of response distributions confirms the similarity of CpG and P3C distributions. IFN-β is the most distinct from all other stimuli but with a response-distribution space closest to that of PIC. Bhattacharya distances were calculated on the top 20 principal components.

(F) A random forest classifier confirms high identifiability of each stimulus condition, with only the bacterial ligands P3C and CpG being confused with each other. Classifier was trained on 70% of all single-cell data (2 replicate experiments from M0 naive macrophage populations) and tested on the remaining 30% of held-out data. Color bar represents sensitivity.

trained a random forest classifier to determine how well the stimulus could be predicted given a single cell's response transcriptome (Figures 1F, S4A, and S4B). We found that IFN- $\beta$  could be perfectly predicted (F1 score = 100%), while the bacterial ligands CpG and P3C were confused with each other and were thus more poorly classified (F1 = 85% and 74%, respectively). The overall prediction accuracy of 88% across all stimuli indicated a high degree of stimulus specificity despite cellto-cell heterogeneity in macrophage responses to ligands (Figure 1F).

#### High stimulus specificity is determined by combinations of individual genes that alone can distinguish only subsets of stimulus pairs

To identify genes that may be driving the observed stimulus specificity, we next employed an information theoretic approach.<sup>26,27</sup> We considered ligand information as transmitted through a channel comprising cell signaling and gene regulatory

networks, both affected by pre-existing biological heterogeneity and the stochasticity of biochemical reactions, to produce heterogeneous gene expression responses (Figure 2A).<sup>28,29</sup> Under this framework, the maximum mutual information (max MI) describes the certainty about the ligand input given the transcriptome output. One bit equates to perfect distinguishability of 2 ligand response distributions ( $2^1 = 2$ ), 2 bits for 4 ligands ( $2^2 = 4$ ), and 2.58 bits for 6 ligands ( $2^{2.58} = 6$ ). Because MI is maximized over all possible input distributions (max MI), this metric provides a comparable absolute quantity for the biological characteristic of response specificity that is independent of the underlying type of data distribution or the mechanisms by which the information is encoded.

We found that 95% of the measured genes on their own conveyed no more than 1 bit of information (Figure 2B; Table S2), which would indicate the ability to distinguish two conditions. A heatmap of the top-ranking individual genes indicated that rather than each gene having several levels of



## Cell Systems Article



Figure 2. Despite the high stimulus specificity of gene expression programs, individual genes are capable of distinguishing only a subset of stimulus pairs

(A) An information theoretic approach treats single-cell signaling and epigenetic mechanisms as a communication channel that passes extracellular information to nuclear target genes.

(B) Distribution of max MI values for single genes shows that only a few genes have a max MI above 1 bit, with 1 bit indicating the ability to distinguish two groups. Highlighted in green are high max MI genes shown in the heatmap in (C).

(C) Heatmap of single-cell gene expression indicates that low max MI is because most genes have binary expression patterns, on or off, rather than a range of levels that would have been able to distinguish multiple groups of stimuli.

(D) Max MI as a function of the indicated number of genes in combination, for the gene combinations providing the highest max MI. The horizontal dotted line indicates the theoretical limit for 6 stimuli.

(E) Gene names for the combinations that allow for the highest max MI for each given number of genes.

(F) Confusion matrices from random forest classifiers for gene combinations indicate that a small number of genes generate fair prediction accuracy, and a classifier trained on 15 genes performs approximately the same as all genes.

(G) The genes comprising the top 2-gene combination, *Cmpk2* and *Nfkbiz*, distinguish complementary stimulus pairs. Error bars represent standard deviations from 10 iterations of 50% bootstrap resampling of the single-cell data.

expression that would separate ligands, most genes displayed an on-or-off expression pattern across single cells, which thus distinguishes only two groups of ligands (Figure 2C). The single gene most informative for distinguishing stimuli was *Cmpk2*, a mitochondria-associated gene with reported antiviral and homeostatic functions,<sup>30,31</sup> which allowed for a max MI of 1.5 bits (Figure 2C; Table S2), corresponding to expression distributions sufficiently narrow to define approximately 3 kinds of stimulusspecific responses.

How then can macrophages achieve high response specificity? One possibility is that a combination of genes that shows low cell-to-cell heterogeneity may specify stimulus-specific responses to multiple stimuli. To assess this possibility, we employed the same information theoretic framework to calculate the max MI provided by the best-performing combinations of genes. Indeed, the best combination of two genes (*Cmpk2* and *Nfkbiz*) already allowed for a max MI of almost 2 bits, with the gain in max MI plateauing to ~2.25 bits as larger combinations were tested (Figure 2D). Interestingly, the majority of genes within the top gene combinations were intracellular proteins controlling nucleotide metabolism (*Cmpk2*),<sup>32</sup> antiviral activity and cell death (*Ifit3, Ifit1, and Mx2*),<sup>33–35</sup> ubiquitination (*Peli1*),<sup>36</sup> mRNA half-life of inflammatory genes (*Zc3h12a*),<sup>37</sup> or phagocytosis (*Swap70*)<sup>38</sup> (Figure 2E). Cytokine genes commonly



Figure 3. The response specificity of cytokine genes is limited by cell-to-cell heterogeneity, despite having distinct expression means (A) Deviation in mean expression among the six stimuli plotted against max MI for each gene. Genes are colored by their average amount of dispersion across the six stimuli (average Fano factor).

(B) Average Fano factor across the six stimuli plotted against max MI shows genes with high or low dispersion. Genes are colored by deviation in mean expression across stimuli (mean-square deviation).

(C) Mean-square deviation vs. average Fano factor over all stimuli, with cytokine genes highlighted.

(D) Expression distributions of cytokine (Cc/5, Tnf, and Cxc/10) and non-cytokine (Cmpk2) genes, showing distinct variance of responses to each stimulus.

(E) Mean expression level vs. expression variance for each stimulus-response distribution. Dotted lines mark where variance equals mean.

measured by antibody-based assays (*Tnf, Cxcl10, Ccl5*, and *Ccl2*) had secondary roles (Table S3). Machine learning classification using different gene combination sizes confirmed these conclusions. Even just the top two genes performed reasonably well at 70% accuracy, the top 5 genes further improved classification accuracy to 78%, and the top 15 genes performed almost as well as all genes at 85% accuracy (vs. 88% for all) (Figure 2F cf. Figure 1F). Gene combinations that worked well together did so by distinguishing complementary ligand pairs (Figure 2G). For example, *Nfkbiz* alone contributed 0.75 bits to distinguishing CpG and PIC, complementing the inability of *Cmpk2* expression to distinguish these ligands (0 bits).

#### The response specificity of cytokine genes is limited by high cell-to-cell heterogeneity, despite having distinct expression means

High response specificity requires not only that mean population gene expression is distinct but also that the cell-to-cell heterogeneity is low, such that distributions of single-cell responses have limited overlap. Hence, we investigated whether SRS was limited by small differences in means or wide distributions. We found that mean-square deviation (MSD), which summarizes differences in means across ligands, correlated more strongly to response specificity (r = 0.8) than the average Fano factor, which summarizes average gene heterogeneity (r = -0.5) (Figure 3A). However, a few outliers were evident: some genes with a high mean difference (e.g., Ccl5 or Cxcl10) showed unimpressive response specificity, with similar max MI to genes with a low mean difference (e.g., Ifi205) (Figure 3A). Plotting the average Fano factor revealed that these outliers differed in their dispersion: Cxc/10 and Cc/5 displayed much higher heterogeneity (i.e., high avg. Fano factors) (Figure 3B). By contrast, metabolic and non-secreted genes such as Cmpk2, Ifi205, and Ifit3 had tight distributions (i.e., unusually low avg. Fano factors) and thus high response specificity.

Interestingly, a pattern emerged where the Fano factors of cytokine/chemokine genes were higher than expected, diminishing their response specificity despite distinct mean



expression levels (Figure 3C). For example, directly plotting the response distributions of cytokines *Ccl5*, *Tnf*, and *Cxcl10* (Figure 3D), as well as of the non-cytokine gene *Cmpk2*, showed that *Ccl5* was induced with high variance expression distributions in response to PIC, P3CSK, and CpG, while *Cxcl10* and *Tnf* distributions also had Fano factors (variance/mean ratio) close to or above one for multiple stimuli, which meant that their distributions were broad in relation to mean expression. By contrast, the Fano factors across stimuli for the non-cytokine gene *Cmpk2* were well below one for all stimuli (Figure 3E). These results suggest that single-cell cytokine/chemokine expression may have limited predictive value in specifying distinct ligand responses due to single-cell heterogeneity.

## The stimulus specificities of immune response genes are due to the selective deployment of IRF and p38 pathways and NF- $\kappa$ B dynamical features

In macrophage responses, four signaling pathways and their downstream gene regulatory factors are combinatorially activated and are responsible for transmitting information about extracellular ligands to the nucleus (Figure 4A). Their target genes have been categorized into five gene regulatory strategies, namely, AP1, NF- $\kappa$ B, IRF, NF- $\kappa$ B|p38, and NF- $\kappa$ B|IRF.<sup>8</sup> We asked which gene regulatory strategies may mediate the high SRS of particular genes, as measured by max MI. We assigned each gene to a regulatory logic through matching mathematical model simulations of possible regulatory logics to available stimulus-response datasets, followed by a curation of prior literature (Figure S5A; Table S4; STAR Methods). The appropriate motifs were enriched for genes assigned to each of the clusters (Figure S5B). We then calculated the max MI of every gene for ligand pairs (Figures 4B-4D). For example, contrasting TNF vs. IFN-β, genes within every regulatory logic group showed high specificity (Figure 4B), a reflection of the fact that each group is activated by only one of the ligands: AP1 and NF-kB and NF-kB|p38 are activated by TNF but not IFN- $\beta$ , whereas IRF genes are activated by IFN- $\beta$  but not TNF.

By contrast, for the P3CSK vs. TNF stimulus pair, only NF- $\kappa$ B| p38 target genes showed specificity because both stimuli activate AP1 and NF- $\kappa$ B, fail to activate IRF, and differ in the extent of p38 activation (Figure 4C). Similarly, for P3CSK vs. PIC, NF- $\kappa$ B|p38 targets showed specificity because PIC does not activate p38.<sup>8</sup> However, for P3CSK vs. PIC, unlike P3CSK vs. TNF, IRF target genes also contributed to SRS, since PIC activates IRF but P3CSK does not (Figure 4D). Of note, the max MI of IRF target genes was on average lower for P3CSK vs. PIC than for TNF vs. IFN- $\beta$ , potentially due to highly heterogeneous activation of IRF by the TRIF signaling pathway.<sup>40</sup>

In addition to combinatorial pathway control, the dynamics of NF-κB activation also specify gene expression.<sup>41</sup> Stimulusspecific NF-κB temporal dynamics involve six NF-κB signaling codons that convey information about the stimulus to target genes: "speed," "peak amplitude," "oscillations," "duration," "total activity," and "early vs. late" activity (Figure 4E).<sup>39</sup> To determine which signaling codons may be associated with the stimulus specificity of NF-κB target genes, we correlated pairwise specificities of NF-κB signaling codons with the pair-

## Cell Systems Article

wise specificities of NF- $\kappa$ B target genes (Figure S5C). We found that NF- $\kappa$ B target genes differed in with which signaling codons they were associated (Figures 4F and S5D). These distinctions may be reflective of genes employing distinct gene regulatory mechanisms such as an incoherent feedforward loop that decodes peak amplitude,<sup>42</sup> long mRNA half-life or slow chromatin opening steps that decode duration,<sup>43</sup> or the requirement for *de novo* enhancers that distinguishes oscillations.<sup>44</sup>

## Cytokine polarization modulates the response specificity of specific genes to specific stimuli

Macrophages show remarkable functional pleiotropy that is dependent on microenvironmental context.<sup>45</sup> Thus, polarization by prior cytokine exposure may alter their capacity for stimulus-specific responses. To test this hypothesis, we polarized macrophages into M1(IFN- $\gamma$ ) and M2(IL-4) states, which represent opposing ends of the macrophage polarization spectrum (Figure S6A),<sup>46</sup> and generated single-cell stimulus-response data for the six ligands (Figure 5A). Polarized M1(IFN- $\gamma$ ) and M2(IL-4) macrophages expressed macrophage marker *Adgre1* (Figure S6B) and the appropriate polarization markers (Figures S6C and S6D). PCA and UMAP projections of response distributions revealed that stimulus responses for polarized macrophages were distinct (Figures 5B and S6E), but M1(IFN- $\gamma$ ) response distributions were more overlapping than those for M0 naive macrophages.

We found that SRS was reduced in both polarization states for every set of the best gene combinations, as calculated by max MI, indicating that polarized macrophages may function more as specialized effectors and less as sentinels that serve a primary role of distinguishing immune threats (Figure 5C; Table S3). Interestingly, we further observed that the genes within each of the best gene sets were different for each polarization state (Figure 5D). This was corroborated by examining one gene. Cxc/10, which was included in the best gene combinations in M0 and M2 (IL-4) conditions but not in M1 (IFN-<sub>γ</sub>). Indeed, in M1 (IFN-γ) macrophages, Cxc/10 was promiscuously rather than stimulus-specifically activated and no longer carried stimulus-specific information about any pairs of stimuli (Figure 5E). This change in SRS could be ascribed to the IRF pathway. Several NF-kBIRF target genes (Cxcl10, Cmpk2, Ifit3, and Trim21) lost specificity in M1(IFN- $\gamma$ ) macrophages (Figures 5F and S6F), reflecting the fact that in the presence of IFN- $\gamma$  conditioning, such genes only require activation of NF-kB to be induced. In fact, using the random forest machine learning model trained on M0 naive macrophages to predict the stimuli seen by M1 or M2 cells, we found that M1 responses were all more likely to be predicted as LPS or PIC, which are IRF-activating stimuli (Figure 5G). Gene ontology attributed this loss of SRS to the biological pathways "response to virus," "response to LPS," and "response to IFN-β" (Figure S6G; Table S5). Motif enrichment analysis also identified IRF target genes as responsible (Figure 5H).

While a global analysis indicated that polarization diminished macrophage response specificity, the genes most affected differed for each macrophage state. In addition, a smaller set of largely NF- $\kappa$ B response genes also showed increased rather than diminished response specificities under

В

1.00

0.75

0.25

(bits)

₹ 0.50

max

IFNβ vs. TNF

Nfkb

Tlr2

## Cell Systems Article



eCmpk2

Trim21 - Cd69

Cxcl10

Peli1e Ifi205 Mx1

Çcl12

Ehd1

elfit1

n30r

eFgl2

IFNβ – TNF

5.0

2.5 0.0

-25

G530011006Rik







Time (hr)

Signaling Codons

Oscillatory

Early

Speed

Duration

Peak

Total

0.00 \* [ 0000 ] \* 18 18 J.X D P3CSK vs. PIC 1.00 Nfkbiz lfit3 0.75 P3C – PIC (bits) Tnf-0 Zc3h12a Cmpk2 olfit3b 2 max MI 0.50 Cxcl2 Ifit2 0 -2 0.25 8 0.00 184 184 1 000 M T WH WH ¢ F Mmp9 Pilra 0.8 Pearson's correlation -II1b Tnf 0.6 Nfkbiz 0.4 0.2 Clec4e 0 -116 -0.2 Ccl5Cxcl10 -0.4 -Peli1 NFκB NFκB | p38 Post Aria Osiliatori Outation NFKB | IRF 200 

Figure 4. Stimulus specificity of immune response genes from the selective deployment of IRF and p38 pathways, and NF-KB dynamical features

(A) Signaling and gene regulatory mechanisms are responsible for generating response specificity. MAPKp38, IRF3, AP1, and NF-κB signaling profiles are activated in response to inflammatory stimuli and act in combination to regulate gene expression, using five identified regulatory logics.<sup>8</sup>

(B) Distribution of max MI values for genes of the five regulatory logics for IFN- $\beta$  vs. TNF.

(C) Distribution of max MI values for genes of the five regulatory logics for P3CSK vs. TNF.

(D) Distribution of max MI values for genes of the five regulatory logics for P3CSK vs. PIC.

(E) Single-cell NF-κB signaling dynamics in response to TNF, P3CSK, LPS, or PIC. Heatmap represents nuclear NF-κB concentration. Information conveyed by NF-κB dynamical activity is transmitted through six NF-κB signaling codons.<sup>39</sup>

(F) Response specificity of NF-κB target genes across all pairs of stimuli vs. Stimulus specificity in NF-κB signaling codons. Color represents the correlation strength between pairwise gene expression and pairwise signaling codon response specificity.

each polarization condition (Figure 5F). Such nuanced findings preclude the use of a single gene or even a single set of genes to quantify SRS over multiple macrophage states. The

response specificity of macrophages thus may not be appropriately characterized by a single number, but rather by a higher dimensional profile.



## Cell Systems Article



Figure 5. Cytokine polarization causes complex modulation of the response specificity of specific genes to specific stimuli

(A) The response specificity experimental assay captures response distributions to six stimuli in M1(IFN-γ) and M2(IL-4) polarized macrophages.

(B) PCA on all M0, M1(IFN-γ), and M2(IL-4) macrophage responses to the six stimuli. Variance explained by PC1 (19.9%) and by PC2 (15.7%). Polarized macrophages produce response distributions distinct from those of M0 macrophages.

(C) Response specificity as calculated by max MI is slightly reduced in both M1(IFN- $\gamma$ ) and M2(IL-4) polarization states, for every dimension of best possible gene combinations.

(D) Genes within the best 15-gene combination for M0, M1(IFN-γ), and M2(IL-4) macrophages are distinct.

(E) Top: distribution of max MI values of individual genes for M0, M1(IFN-γ), and M2(IL-4) macrophage responses. Dotted lines mark the max MI of *Cxc/10* for each macrophage state: red (M0), green (M1), and blue (M2). Bottom: *Cxc/10* stimulus-response distributions for M0 (left), M1 (center), and M2 (right) macrophage states are shown.

(F) Scatterplot of differences in max MI values between M0 vs. M1(IFN-γ) or M2(IL-4) macrophage responses. Genes are colored by assigned gene regulatory logic. (G) Confusion matrices representing the accuracy of stimulus prediction, given single-cell gene expression profiles for M1 and M2 macrophages, using the random forest model trained on naive M0 macrophages. Color bar represents sensitivity.

(H) Motifs that lose specificity in M1 or M2 macrophages, compared with M0 naive macrophages. Left: enrichments of all potential motifs. Right: barplot of most significant enrichment for each motif category.

#### The response specificity profile of stimulus pairs assesses the functional state of macrophages and readily distinguishes M0 vs. M1 vs. M2 macrophages

We noted that while different macrophage states could be identified by profiling their steady-state transcriptomes (Figure S6A), their responses to stimuli were even more distinguishable (Figure 6A). Quantifying the distance between both the stimulated and unstimulated distributions, we found that in all comparisons, specific stimuli revealed differences that were not as evident from steady-state measurements. This was especially evident for M0 vs. M2 macrophages, whose response distributions to P3C and CpG were particularly more distinct than could be ascertained from steady-state distributions (Figure 6B). This illustrates that for assessing the functional state of macrophages, measurements of multiple stimulus responses provide nonredundant information.

To quantitatively assess the response specificity provided by all 6 stimuli, we developed an approach to profile the max MI of all 15 stimulus pairs within the macrophage response landscape characterized by a PCA projection of all single-cell stimulus-response transcriptomes (Figure 6C; Table S6; STAR Methods). The resulting response specificity profile of M0, M1(IFN- $\gamma$ ), and M2(IL-4) macrophages showed specific differences for some of the stimulus pairs tested (Figure 6C).



Figure 6. The response specificity profile of stimulus pairs assesses the functional state of macrophages, and readily distinguishes M0 vs. M1 vs. M2

(A) PCA of stimulus responses vs. resting-state expression profiles of M0 vs. M1 vs. M2 macrophages.

(B) Quantification of differences in response distributions vs. resting-state distributions among the three macrophage states, using Bhattacharyya distance on the first three components. Colors are the same as in (A).

(C) Response specificity profiles measured by max MI of all 15 stimulus pairs, as defined by response distributions to 6 stimuli. M0, M1, and M2 macrophage responses each show specific differences in select pairs. Error bars represent standard deviations from 50 iterations of 50% bootstrap resampling.

(D) Across all pairs, the difference in max MI from M0 responses is calculated to derive  $\Delta$ RSI, a characteristic signature of the functional state of the macrophage. (E) Top: a summary of the pairwise profile into a single number, the  $\Delta$ RSI. Bottom:  $\Delta$ RSI provides a clearer indication of differences than calculating the MI on all stimulus-response datasets together. Error bars for  $\Delta$ RSI represent the propagation of the original standard deviations in max MI calculations.

Particularly, M1(IFN- $\gamma$ ) macrophages were impaired in distinguishing bacterial stimuli (LPS-CpG and LPS-P3C), whereas M2(IL-4) macrophages were not. Meanwhile, M2(IL-4) macrophages were more impaired in distinguishing host cytokine vs. bacteria (TNF-CpG and TNF-P3C). In both polarization states, CpG-P3CSK specificity, which had been the least distinguishable stimulus pair for M0 macrophages, was enhanced.

Calculating the difference in max MI from the M0 state, and thereby the delta response specificity profile, provided a characteristic signature of the functional state of the macrophage (Figure 6D). This profile could also be summarized succinctly in a single number, the delta response specificity index ( $\Delta$ RSI), which provided a clearer indication of differences among polarization states than the MI calculated on all stimuli together (Figure 6E). We found that the overall  $\Delta$ RSI of M2(IL-4) macrophages was greater than that of M1(IFN- $\gamma$ ) macrophages, as a result of M2(IL-4) macrophages showing a greater loss of specificity in distinguishing host vs. bacterial ligands. Thus, the response specificity profile and  $\Delta$ RSI revealed a signature of pairwise SRS scores associated with the function of each macrophage type, whether naive, M1(IFN- $\gamma$ ), or M2(IL-4).

## Peritoneal macrophages from old and obese mice show distinctive alterations in their response specificity profiles

Next, we tested whether the response specificity profile might reveal aberrations in macrophages derived from mice with conditions associated with inflammatory disease. We took



Figure 7. Old and obese mouse peritoneal macrophages show distinctive alterations in their response specificity profile

(A) Measuring response specificity for three mouse models using peritoneal macrophages: healthy low-fat diet mice (16 weeks), old mice (>90 weeks), and highfat diet mice (16 weeks). Cells from two mice were aggregated for each condition.

(B) tSNE visualization of peritoneal macrophage responses to stimuli for each of the mouse models, healthy low-fat diet (LFD), aged (OLD), and high-fat diet (HFD), using the top 20 principal components.

(C) Left: macrophage responses from healthy and diseased mouse models are scored by the response specificity profile. Error bars represent standard deviations from 50 iterations of 50% bootstrap resampling. Right: subtracting each profile from the response specificity profile of M0 macrophages highlights pairwise differences in max MI values.

(D) Square root of the sum of the square deviation from M0 across all stimulus pairs was calculated to obtain the single-value  $\Delta$ RSI (Figure 6B). For both polarized macrophages and peritoneal macrophages, stimulus pairs that included CpG were not included due to CpG not being used in the response specificity assay on peritoneal macrophages. Higher numbers for  $\Delta$ RSI indicate a larger deviation from M0. Error bars represent the propagation of the original standard deviations in max MI calculations.

(E) Scatterplot of differences in max MI values between LFD vs. OLD or HFD macrophage responses. Genes are colored by the absolute max MI quantity for each gene in the LFD mouse model.

peritoneal macrophages (PMs) from healthy mice (17 weeks old), aged mice (90 weeks old), and high-fat-diet obese mice (17 weeks old) and performed the response specificity workflow using 5 ligands: LPS, TNF, PIC, P3C, and IFN-β (Figure 7A). Visualization of the resulting single-cell data suggested that old mice had aberrant IFN-β-response distributions, while obese mice had aberrant PIC-response distributions (Figure 7B). We calculated the response specificities for the data (Figure 7C) and noted that the PMs from young, healthy mice had an overall ΔRSI most similar to naive M0 macrophages (Figure 7D). A closer inspection of the response specificity profile showed that PMs from old mice showed the most diminished specificity for interferon-activating stimulus pairs (e.g., LPS-PIC), akin to what we observed with M1(IFN-y) macrophages. PMs from high-fatdiet, obese mice had decreased specificity across all stimulus pairs, but particularly in distinguishing cytokine vs. viral (TNF-PIC) responses. The differences observed through calculation of the response specificity profile on stimulus subsets suggest an importance to comparing multiple subsets of stimuli in evaluating innate immune function.

To identify individual genes that showed particularly high losses of SRS in these *in-vivo*-conditioned macrophages, we compared max MI values for each gene (Figure 7E). We found that macrophages from both old and obese mice lost specificity in the cytokine *Tnf*, but also in the metabolic gene *Acod1*,<sup>47,48</sup> and in upstream TLR signaling network proteins such as *Peli1*,<sup>36</sup> *Nfkbiz*,<sup>49,50</sup> and *PhIda1*.<sup>51</sup> Normal function of these genes has been implicated in resistance to septic shock, macrophage response to atherosclerosis, and protection from autoimmune disease. Interestingly, a couple of genes showed higher stimulus specificity in these dysregulated microenvironments. In old mice, genes with the highest gains in specificity were the cytokine *Cxcl10* and cytokine regulator *Aw112010*, which are required for mucosal immunity,<sup>52</sup> and in obese mice, *Cav1* and

## Cell Systems Article

*Slamf8*, which play roles in macrophage differentiation and migration.<sup>53,54</sup> This suggests that macrophage responses in unhealthy mice deviate from those in healthy mice through both losses and gains of stimulus specificity in individual genes: losses indicating responses that are too promiscuous, potentially causing conditions for autoimmunity to arise; and gains indicating responses that are too restricted to a particular stimulus, potentially diminishing core functions of macrophages in innate defenses or resolution of tissue inflammation and damage.

As PMs of mice of different ages may be composed of different subpopulations, we next evaluated to what extent the changes in SRS may be due to subpopulations of macrophages differentially present in diseased vs. healthy mice. Based on scRNA-seq of steady-state populations,<sup>55</sup> we found that all subclusters of macrophages were found in both young and old mice (Figures S7A–S7E), although in slightly different proportions (Figure S7F). Using marker genes to match these clusters to the stimulus-response data, we likewise found these subgroups represented across healthy, aged, and obese mice, although again in different proportions (Figure S7G).

Taken together, quantifying the SRS of macrophages revealed that this functional hallmark of immune sentinel cells is affected not only by polarizing cytokines used in pre-conditioning regimes *in vitro* but also by the microenvironments *in vivo* that are evidently distinct in obese and old mice. It is possible that response specificity profiles of PMs capture altered responses of both distinct subpopulations of macrophages that are differentially represented in inflammatory conditions and distinct functional states of the same subpopulation. Regardless, the observed differences in SRS suggest that quantifying post-stimulation single-cell response distributions could be valuable for assessing innate immune function.

#### DISCUSSION

Mounting stimulus-appropriate immune responses is a key property of healthy macrophage function.7,8,56-58 Macrophage response specificity is a function of the stimulus-specific engagement of signaling pathways and may be diminished by molecular network noise that results in cell-to-cell heterogeneity. While response specificity of the macrophage NF-kB signaling pathway has been characterized, 39,59 the response specificity of immune gene expression responses arising from all macrophage signaling pathways has not yet been quantified. By developing the experimental and computational tools to do so, we found that the high gene expression specificity observed was generated by sufficiently narrow response distributions in combinations of genes that respond with distinct patterns across stimuli, but the contribution of often-measured cytokine genes was limited by high cell-to-cell heterogeneity of expression. Mechanistically, we found that SRS is generated by stimulusspecific activation of IRF or MAPKp38 signaling, or by differences in NF-κB dynamics. Loss of IRF gene specificity by microenvironmental polarization was the key driver in altering response specificity profiles. Given that response specificity is context-dependent, we profiled PMs from old and obese mice, revealing highly specific changes in the response specificity profile that correlated with a different health status. These findings may prompt further studies to investigate whether macrophage



response specificity could be a means to characterize the innate immune health of human donors.

Our ability to measure and subsequently quantify response specificity was enabled by a quantitative assay for costeffective, reliable scRNA-seq.<sup>23,60</sup> The targeted sequencing approach we pursued here resulted in less technical noise than genome-wide approaches, due to improved reverse-transcriptase efficiency and increased sequencing depth per gene.<sup>23</sup> However, even with technical improvements, the remaining measurement noise<sup>61</sup> still may result in underestimates of the true response specificity. Future efforts on smaller gene lists may allow for the use of even less noisy measurement approaches like single-molecule fluorescence *in situ* hybridization (smFISH),<sup>62,63</sup> which capitalize on the smaller sets of informative genes identified in this study.

Measuring the responses from macrophages rather than their steady-state transcriptomes provided additional levels of information. First, we found that stimulus-response transcriptomes may reveal differences in macrophage populations that are not apparent in the steady state. This may be because exposure histories and cytokine contexts change signaling pathways and chromatin states that are not reflected in steady-state mRNA abundances. Measuring stimulus responses also allowed us to evaluate multiple distributions for each macrophage type, rather than the single distribution provided by profiling the steady state. Emergent from these multiple distributions is response specificity, which evaluates the relative extent of distribution overlap and reports on the ability of each macrophage type to distinguish among immune threats. Not only is response specificity an important biological hallmark property of macrophages,<sup>58</sup> but also from the workflow perspective, being based on multiple measurements that may be compared with one another, it makes this analysis metric more resistant to technical noise or batch effect, potentially enabling better comparisons across studies.

The information theoretic approach we used here to quantify SRS has been previously employed to quantify the information transmission capacity of signaling pathways.<sup>28,39,59,64-69</sup> In fact, we found that SRS of individual genes can be traced either to a single pathway or by multiple pathways,<sup>8</sup> for example, the NF-kB target gene Tnf whose mRNA half-life is regulated by stimulus-induced MAPKp38.70 This combinatorial control explains why in principle some single genes like Tnf can hold more information than available from a single signaling pathway. But even single pathway genes may show response specificity owing to their ability to distinguish different dynamical characteristics, such as NF-kB signaling codons. Interestingly, we observed that such genes correlated strongly to identifiable signaling codons, indicating that their gene regulatory strategies are able to decode the information present in the stimulus-specific deployment of the signaling codon.

Theoretically, information loss from signaling to gene expression is minimal without noise, <sup>71,72</sup> while with noise, maintaining minimal information loss is only possible under select optimal promoter or chromatin conditions.<sup>73–75</sup> Specifically, to achieve high gene expression response specificity, signaling information must be interpreted by gene regulatory strategies<sup>76,77</sup> without amplifying the cell-to-cell heterogeneity in signaling activity,<sup>78</sup> and also without introducing further heterogeneity through



## Cell Systems Article

pre-existing chromatin heterogeneity<sup>79,80</sup> or molecular noise.<sup>81–84</sup> In this context, it is not surprising that the SRS of most individual genes was low. However, as macrophages do not rely on only a single gene to mount a biological response to a specific immune threat, even with information loss at each particular promoter, the overall SRS observed through combinations of a few genes from complementary pathways was still high.

Within the body, macrophages are exposed to polarizing microenvironments in physiological scenarios,<sup>85</sup> as well as in pathological inflammatory contexts such as aging or obesity.<sup>86–88</sup> As circulating cells they are potentially reporters of even localized infections.<sup>89</sup> Indeed, profiling the transcriptome or epigenome of circulating cells or macrophages has revealed molecular markers or signatures that are prognostic for therapeutic efficacy<sup>90,91</sup> or alternative disease courses, such as in persistent infectious or inflammatory diseases<sup>15,92</sup> or in cardiovascular and autoimmune diseases.93-97 However, as seen in human COVID-19 studies, it can be unclear which individuals have poor or vigorous immune health until they are challenged by infection.<sup>98–102</sup> Here, we considered that macrophage functions are deployed in response to immune threats and that stimulus responses are a function not only of the steady-state transcriptome or epigenome but also the dynamics of signaling complexes, membranes, and transport rates. We reasoned that macrophage responses to different stimuli reveal a functional pleiotropy not evident at steady state and that the stimulus-specific deployment of functions is key to healthy immunity. Indeed, we found that macrophages conditioned in vitro by defined polarizing cytokines, as well as macrophages isolated from obese or old mice, showed distinctly altered response specificity profiles due to both cytokine genes and regulators within macrophage metabolic and signaling pathways.

In developing the response specificity profile, we found that analyzing single genes and pairs of stimuli provided more insight than aggregating the data together. For example, response specificity for each pair of stimuli differed for each polarization condition, but this information is lost when calculating MI for all stimuli at once. Instead, an aggregate score of alterations in the response specificity profile ( $\Delta$ RSI) provides the first indication of differences and the full response specificity profile pinpointed aberrancies in select stimulus pairs that may be a diagnostic for a specific condition, such as macrophages from obese mice confusing TNF and PIC responses. For initial surveys of response specificity profiles, measuring the expression of a large number of genes in response to multiple stimuli is important, as the most informative genes and stimulus comparisons are different across various macrophage states.<sup>103</sup> We employed a PCA approach to characterize the response landscape and identify genes that are important across conditions, using the gene weights in principal components. Indeed, the response specificity of individual genes differed greatly between the two disease models tested here, emphasizing the importance of gene-by-gene analysis in characterizing the response specificity profile to make biologically meaningful predictions.

Characterizing the macrophage response specificity profile may prove useful in clinical scenarios. Many steady-state metrics of immune health exist, such as the complete blood count that is a mainstay of clinical lab tests. However, tests for functional immune responses, already used in select clinical scenarios such as tuberculosis testing or allergy testing, are also rapidly emerging.<sup>104</sup> Multiple such assays rely on *ex vivo* stimulation of extracted clinical samples to diagnose immunosuppression<sup>105</sup> or phagocytic ability<sup>106</sup> and have included transcriptomic profiling studies of stimulus responses on peripheral blood, identifying inter-individual variation among healthy donors and the genes driving those differences.<sup>107</sup> As seen with existing assays, the extent to which stimulus-response measurements provide more reliably prognostic information than steady-state molecular profiling may depend on the health condition being studied.

Our assay of single-cell macrophage responses may identify outlier response cells within each donor sample through the response specificity profile, which may be associated with risk for aberrant inflammatory responses or diminished innate immune defenses, or which may be reflective of an ongoing inflammatory or infectious condition that is not otherwise presented. The response specificity profile allows for new samples to be compared readily with a healthy range, a property that could be the basis for a clinically deployable measure of innate immune health. The stimulus-response data may also identify specific genes with aberrant response distributions in patient cohorts or in individual donors. Identifying such genes may provide cost-effective prognostic markers for specific cohorts or may point toward the underlying etiology of poor immune function. To realize this promise, large-scale clinical studies will be required to establish connections between response specificity profiles and risk for disease. However, as a quantifiable property of macrophage function that changes with conditioning cytokines or states of health and disease, the macrophage response specificity profile may be a viable approach for measuring the health of innate immune function in the clinic.

#### **STAR**\*METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- Macrophage Cell Culture
- Peritoneal Macrophage Experiments
- METHOD DETAILS
  - Gene Panel Selection Algorithm
  - Selection of Stimuli for Response Specificity Assay
  - Rhapsody scRNAseq
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Motif Enrichment
  - Gene Ontology
  - scRNAseq Data Processing
  - $\odot\,$  Assignment of Genes to Regulatory Mechanisms
  - Machine Learning Models and Feature Importance
  - Mutual Information Analyses
  - Response Specificity Profile





#### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <a href="https://doi.org/10.1016/j.cels.2022.12.012">https://doi.org/10.1016/j.cels.2022.12.012</a>.

#### ACKNOWLEDGMENTS

We thank lab members Ying Tang and Diane Lefaudeux for their helpful comments and Supriya Sen and Carolina Chavez for reagents and experimental and analysis assistance. We thank Jessica Li (UCLA) for their support and advice. K.M.S. was supported by the UCLA Medical Scientist Training Program (T32-GM008042) and Systems and Integrative Biology Training Grant (T32-GM008185). The work was funded by R01Al127864 and R01GM12507 to A.H. Cartoon mouse icons were created with BioRender. Sequencing data from this paper was performed with the services of the UCLA Technology Center for Genomics and Bioinformatics Sequencing Core.

#### **AUTHOR CONTRIBUTIONS**

K.M.S. and A.H. designed the study. K.M.S. performed the experiments and conducted the analyses. A.A.G. helped piloting the experimental approach and assisted with the machine learning analysis. K.M.S. and A.H. wrote the paper. A.H. funded the work. All authors reviewed the manuscript.

#### **DECLARATION OF INTERESTS**

The authors are inventors on a provisional patent application related to this work, filed by UCLA.

#### **INCLUSION AND DIVERSITY**

While citing references scientifically relevant for this work, we also actively worked to promote gender balance in our reference list.

Received: June 1, 2022 Revised: October 5, 2022 Accepted: December 22, 2022 Published: January 18, 2023

#### REFERENCES

- Murray, P.J., and Wynn, T.A. (2011). Protective and pathogenic functions of macrophage subsets. Nat. Rev. Immunol. *11*, 723–737. https://doi. org/10.1038/nri3073.
- Alleva, D.G., Pavlovich, R.P., Grant, C., Kaser, S.B., and Beller, D.I. (2000). Aberrant macrophage cytokine production is a conserved feature among autoimmune-prone mouse strains: elevated interleukin (IL)-12 and an imbalance in tumor necrosis factor-alpha and IL-10 define a unique cytokine profile in macrophages from young nonobese diabetic mice. Diabetes *49*, 1106–1115. https://doi.org/10.2337/diabetes.49. 7.1106.
- Galani, I.-E., Rovina, N., Lampropoulou, V., Triantafyllia, V., Manioudaki, M., Pavlos, E., Koukaki, E., Fragkou, P.C., Panou, V., Rapti, V., et al. (2021). Untuned antiviral immunity in COVID-19 revealed by temporal type I/III interferon patterns and flu comparison. Nat. Immunol. 22, 32–40. https://doi.org/10.1038/s41590-020-00840-x.
- McKechnie, J.L., and Blish, C.A. (2020). The innate immune system: fighting on the front lines or fanning the flames of COVID-19? Cell Host Microbe 27, 863–869. https://doi.org/10.1016/j.chom.2020.05.009.
- Moore, J.B., and June, C.H. (2020). Cytokine release syndrome in severe COVID-19. Science 368, 473–474. https://doi.org/10.1126/science. abb8925.
- Muldoon, J.J., Chuang, Y., Bagheri, N., and Leonard, J.N. (2020). Macrophages employ quorum licensing to regulate collective activation. Nat. Commun. *11*, 878. https://doi.org/10.1038/s41467-020-14547-y.
- 7. Nau, G.J., Richmond, J.F.L., Schlesinger, A., Jennings, E.G., Lander, E.S., and Young, R.A. (2002). Human macrophage activation programs

induced by bacterial pathogens. Proc. Natl. Acad. Sci. USA 99, 1503–1508. https://doi.org/10.1073/pnas.022649799.

- Cheng, C.S., Behar, M.S., Suryawanshi, G.W., Feldman, K.E., Spreafico, R., and Hoffmann, A. (2017). Iterative modeling reveals evidence of sequential transcriptional control mechanisms. Cell Syst. 4, 330– 343.e5. https://doi.org/10.1016/j.cels.2017.01.012.
- Sheu, K.M., Luecke, S., and Hoffmann, A. (2019). Stimulus-specificity in the responses of immune sentinel cells. Curr. Opin. Syst. Biol. 18, 53–61. https://doi.org/10.1016/j.coisb.2019.10.011.
- 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. https://doi.org/10. 1016/j.immuni.2014.01.006.
- Gottschalk, R.A., Martins, A.J., Angermann, B.R., Dutta, B., Ng, C.E., Uderhardt, S., Tsang, J.S., Fraser, I.D.C., Meier-Schellersheim, M., and Germain, R.N. (2016). Distinct NF-kB and MAPK activation thresholds uncouple steady-state microbe sensing from anti-pathogen inflammatory responses. Cell Syst. *2*, 378–390. https://doi.org/10.1016/j.cels. 2016.04.016.
- Luecke, S., Sheu, K.M., and Hoffmann, A. (2021). Stimulus-specific responses in innate immunity: multilayered regulatory circuits. Immunity 54, 1915–1932. https://doi.org/10.1016/j.immuni.2021.08.018.
- Tong, A.J., Liu, X., Thomas, B.J., Lissner, M.M., Baker, M.R., Senagolage, M.D., Allred, A.L., Barish, G.D., and Smale, S.T. (2016). A stringent systems approach uncovers gene-specific mechanisms regulating inflammation. Cell *165*, 165–179. https://doi.org/10.1016/j.cell. 2016.01.020.
- Muñoz-Rojas, A.R., Kelsey, I., Pappalardo, J.L., Chen, M., and Miller-Jensen, K. (2021). Co-stimulation with opposing macrophage polarization cues leads to orthogonal secretion programs in individual cells. Nat. Commun. 12, 301. https://doi.org/10.1038/s41467-020-20540-2.
- Sanin, D.E., Ge, Y., Marinkovic, E., Kabat, A.M., Castoldi, A., Caputa, G., Grzes, K.M., Curtis, J.D., Thompson, E.A., Willenborg, S., et al. (2022). A common framework of monocyte-derived macrophage activation. Sci. Immunol. 7, eabl7482. https://doi.org/10.1126/sciimmunol.abl7482.
- Dichtl, S., Sanin, D.E., Koss, C.K., Willenborg, S., Petzold, A., Tanzer, M.C., Dahl, A., Kabat, A.M., Lindenthal, L., Zeitler, L., et al. (2022). Gene-selective transcription promotes the inhibition of tissue reparative macrophages by TNF. Life Sci. Alliance 5, e202101315. https://doi.org/ 10.26508/lsa.202101315.
- Zhang, B., Moorlag, S.J.C.F.M., Dominguez-Andres, J., Bulut, Ö., Kilic, G., Liu, Z., van Crevel, R., Xu, C.J., Joosten, L.A.B., Netea, M.G., et al. (2022). Single-cell RNA sequencing reveals induction of distinct trained-immunity programs in human monocytes. J. Clin. Invest. *132*, e147719. https://doi.org/10.1172/JCl147719.
- Ginhoux, F., Schultze, J.L., Murray, P.J., Ochando, J., and Biswas, S.K. (2016). New insights into the multidimensional concept of macrophage ontogeny, activation and function. Nat. Immunol. *17*, 34–40. https://doi. org/10.1038/ni.3324.
- Cheng, Q., Behzadi, F., Sen, S., Ohta, S., Spreafico, R., Teles, R., Modlin, R.L., and Hoffmann, A. (2019). Sequential conditioning-stimulation reveals distinct gene- and stimulus-specific effects of Type I and II IFN on human macrophage functions. Sci. Rep. 9, 5288. https://doi.org/10. 1038/s41598-019-40503-y.
- 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. https://doi.org/10.1038/s41467-019-11147-3.
- Lawrence, T., and Natoli, G. (2011). Transcriptional regulation of macrophage polarization: enabling diversity with identity. Nat. Rev. Immunol. *11*, 750–761. https://doi.org/10.1038/nri3088.
- Mitchell, S., Mercado, E.L., Adelaja, A., Ho, J.Q., Cheng, Q.J., Ghosh, G., and Hoffmann, A. (2019). An NFκB activity calculator to delineate signaling crosstalk: type I and II interferons enhance NFκB via distinct



mechanisms. Front. Immunol. 10, 1425. https://doi.org/10.3389/fimmu. 2019.01425.

- Shum, E.Y., Walczak, E.M., Chang, C., and Christina Fan, H. (2019). Quantitation of mRNA transcripts and proteins using the BD Rhapsody<sup>TM</sup> single-cell analysis system. Adv. Exp. Med. Biol. *1129*, 63–79. https://doi.org/10.1007/978-981-13-6037-4\_5.
- Kolda, T.G., and Bader, B.W. (2009). Tensor decompositions and applications. SIAM Rev. 51, 455–500. https://doi.org/10.1137/07070111X.
- Singh, A., Sen, S., Adelaja, A., and Hoffmann, A. (2022). Stimulus-Response signaling dynamics characterize macrophage polarization states. Preprint at bioRxiv. https://doi.org/10.1101/2022.03.27.485991.
- 26. Cover, T.M., and Thomas, J.A. (2012). Elements of Information Theory (John Wiley & Sons).
- Shannon, C.E. (1948). A mathematical theory of communication. Bell Syst. Tech. J. 27, 379–423. https://doi.org/10.1002/j.1538-7305.1948. tb01338.x.
- Suderman, R., Bachman, J.A., Smith, A., Sorger, P.K., and Deeds, E.J. (2017). Fundamental trade-offs between information flow in single cells and cellular populations. Proc. Natl. Acad. Sci. USA *114*, 5755–5760. https://doi.org/10.1073/pnas.1615660114.
- Tkačik, G., Callan, C.G., and Bialek, W. (2008). Information flow and optimization in transcriptional regulation. Proc. Natl. Acad. Sci. USA 105, 12265–12270. https://doi.org/10.1073/pnas.0806077105.
- Lai, J.H., Wu, D.W., Wu, C.H., Hung, L.F., Huang, C.Y., Ka, S.M., Chen, A., Chang, Z.F., and Ho, L.J. (2021). Mitochondrial CMPK2 mediates immunomodulatory and antiviral activities through IFN-dependent and IFN-independent pathways. iScience 24, 102498. https://doi.org/10. 1016/j.isci.2021.102498.
- Arumugam, P., Chauhan, M., Rajeev, T., Chakraborty, R., Shankaran, D., Ramalingam, S., Gandotra, S., and Rao, V. (2021). The mitochondrial gene CMPK2 functions as a rheostat for macrophage homeostasis in inflammation. Preprint at bioRxiv. https://doi.org/10.1101/2021.11.01. 466732.
- Xu, Y., Johansson, M., and Karlsson, A. (2008). Human UMP-CMP kinase 2, a novel nucleoside monophosphate kinase localized in mitochondria. J. Biol. Chem. 283, 1563–1571. https://doi.org/10.1074/jbc. M707997200.
- Hsu, Y.L., Shi, S.F., Wu, W.L., Ho, L.J., and Lai, J.H. (2013). Protective roles of interferon-induced protein with tetratricopeptide repeats 3 (IFIT3) in dengue virus infection of human lung epithelial cells. PLoS One 8, e79518. https://doi.org/10.1371/journal.pone.0079518.
- Pidugu, V.K., Pidugu, H.B., Wu, M.M., Liu, C.J., and Lee, T.C. (2019). Emerging functions of human IFIT proteins in cancer. Front. Mol. Biosci. 6, 148. https://doi.org/10.3389/fmolb.2019.00148.
- Betancor, G., Jimenez-Guardeño, J.M., Lynham, S., Antrobus, R., Khan, H., Sobala, A., Dicks, M.D.J., and Malim, M.H. (2021). MX2mediated innate immunity against HIV-1 is regulated by serine phosphorylation. Nat. Microbiol. 6, 1031–1042. https://doi.org/10.1038/ s41564-021-00937-5.
- Chang, M., Jin, W., and Sun, S.C. (2009). Peli1 facilitates TRIF-dependent toll-like receptor signaling and proinflammatory cytokine production. Nat. Immunol. 10, 1089–1095. https://doi.org/10.1038/ni.1777.
- Matsushita, K., Takeuchi, O., Standley, D.M., Kumagai, Y., Kawagoe, T., Miyake, T., Satoh, T., Kato, H., Tsujimura, T., Nakamura, H., et al. (2009). Zc3h12a is an RNase essential for controlling immune responses by regulating mRNA decay. Nature 458, 1185–1190. https://doi.org/10. 1038/nature07924.
- Baranov, M.V., Revelo, N.H., Dingjan, I., Maraspini, R., Ter Beest, M., Honigmann, A., and van den Bogaart, G. (2016). SWAP70 organizes the actin cytoskeleton and is essential for phagocytosis. Cell Rep. 17, 1518–1531. https://doi.org/10.1016/j.celrep.2016.10.021.
- Adelaja, A., Taylor, B., Sheu, K.M., Liu, Y., Luecke, S., and Hoffmann, A. (2021). Six distinct NFκB signaling codons convey discrete information to distinguish stimuli and enable appropriate macrophage responses.

Immunity 54, 916–930.e7. https://doi.org/10.1016/j.immuni.2021. 04.011.

**Cell Systems** 

- Cheng, Z., Taylor, B., Ourthiague, D.R., and Hoffmann, A. (2015). Distinct single-cell signaling characteristics are conferred by the MyD88 and TRIF pathways during TLR4 activation. Sci. Signal. 8, ra69. https://doi.org/10. 1126/scisignal.aaa5208.
- Behar, M., and Hoffmann, A. (2010). Understanding the temporal codes of intra-cellular signals. Curr. Opin. Genet. Dev. 20, 684–693. https://doi. org/10.1016/j.gde.2010.09.007.
- Lee, R.E., Walker, S.R., Savery, K., Frank, D.A., and Gaudet, S. (2014). Fold change of nuclear NF-kappaB determines TNF-induced transcription in single cells. Mol. Cell 53, 867–879. https://doi.org/10.1016/j.molcel.2014.01.026.
- Sen, S., Cheng, Z., Sheu, K.M., Chen, Y.H., and Hoffmann, A. (2020). Gene Regulatory Strategies that Decode the Duration of NFκB Dynamics Contribute to LPS- versus TNF-Specific Gene Expression. Cell Syst. 10, 169–182.e5. https://doi.org/10.1016/j.cels.2019.12.004.
- Cheng, Q.J., Ohta, S., Sheu, K.M., Spreafico, R., Adelaja, A., Taylor, B., and Hoffmann, A. (2021). NF-κB dynamics determine the stimulus specificity of epigenomic reprogramming in macrophages. Science 372, 1349–1353. https://doi.org/10.1126/science.abc0269.
- Murray, P.J. (2017). Macrophage polarization. Annu. Rev. Physiol. 79, 541–566. https://doi.org/10.1146/annurev-physiol-022516-034339.
- Murray, P.J., Allen, J.E., Biswas, S.K., Fisher, E.A., Gilroy, D.W., Goerdt, S., Gordon, S., Hamilton, J.A., Ivashkiv, L.B., Lawrence, T., et al. (2014). Macrophage activation and polarization: nomenclature and experimental guidelines. Immunity *41*, 14–20. https://doi.org/10.1016/j.immuni.2014. 06.008.
- 47. Li, Y., Zhang, P., Wang, C., Han, C., Meng, J., Liu, X., Xu, S., Li, N., Wang, Q., Shi, X., et al. (2013). Immune responsive gene 1 (IRG1) promotes endotoxin tolerance by increasing A20 expression in macrophages through reactive oxygen species. J. Biol. Chem. 288, 16225–16234. https://doi.org/10.1074/jbc.M113.454538.
- Sun, K.A., Li, Y., Meliton, A.Y., Woods, P.S., Kimmig, L.M., Cetin-Atalay, R., Hamanaka, R.B., and Mutlu, G.M. (2020). Endogenous itaconate is not required for particulate matter-induced NRF2 expression or inflammatory response. eLife 9, e54877. https://doi.org/10.7554/eLife.54877.
- Kakiuchi, N., Yoshida, K., Uchino, M., Kihara, T., Akaki, K., Inoue, Y., Kawada, K., Nagayama, S., Yokoyama, A., Yamamoto, S., et al. (2019). Frequent mutations that converge on the NFKBIZ pathway in ulcerative colitis. Nature 577, 260–265. https://doi.org/10.1038/s41586-019-1856-1.
- 50. Okuma, A., Hoshino, K., Ohba, T., Fukushi, S., Aiba, S., Akira, S., Ono, M., Kaisho, T., and Muta, T. (2013). Enhanced apoptosis by disruption of the STAT3-I<sub>κ</sub>B-ζ signaling pathway in epithelial cells induces Sjögren's syndrome-like autoimmune disease. Immunity *38*, 450–460. https://doi.org/10.1016/j.immuni.2012.11.016.
- Lyu, J.H., Huang, B., Park, D.W., and Baek, S.H. (2016). Regulation of PHLDA1 expression by JAK2-ERK1/2-STAT3 signaling pathway. J. Cell. Biochem. *117*, 483–490. https://doi.org/10.1002/jcb.25296.
- Jackson, R., Kroehling, L., Khitun, A., Bailis, W., Jarret, A., York, A.G., Khan, O.M., Brewer, J.R., Skadow, M.H., Duizer, C., et al. (2018). The translation of non-canonical open reading frames controls mucosal immunity. Nature 564, 434–438. https://doi.org/10.1038/s41586-018-0794-7.
- Fu, Y., Moore, X.-L., Lee, M.K.S., Fernández-Rojo, M.A., Parat, M.O., Parton, R.G., Meikle, P.J., Sviridov, D., and Chin-Dusting, J.P.F. (2012). Caveolin-1 plays a critical role in the differentiation of monocytes into macrophages. Arterioscler. Thromb. Vasc. Biol. 32, e117–e125. https:// doi.org/10.1161/ATVBAHA.112.254151.
- Wang, G., van Driel, B.J., Liao, G., O'Keeffe, M.S., Halibozek, P.J., Flipse, J., Yigit, B., Azcutia, V., Luscinskas, F.W., Wang, N., et al. (2015). Migration of myeloid cells during inflammation is differentially regulated by the cell surface receptors Slamf1 and Slamf8. PLoS One 10, e0121968. https://doi.org/10.1371/journal.pone.0121968.

## Cell Systems Article



- Hajishengallis, G., and Lambris, J.D. (2011). Microbial manipulation of receptor crosstalk in innate immunity. Nat. Rev. Immunol. *11*, 187–200. https://doi.org/10.1038/nri2918.
- Rivera, A., Siracusa, M.C., Yap, G.S., and Gause, W.C. (2016). Innate cell communication kick-starts pathogen-specific immunity. Nat. Immunol. *17*, 356–363. https://doi.org/10.1038/ni.3375.
- Sheu, K.M., and Hoffmann, A. (2022). Functional hallmarks of healthy macrophage responses: their regulatory basis and disease relevance. Annu. Rev. Immunol. 40, 295–321. https://doi.org/10.1146/annurev-immunol-101320-031555.
- Selimkhanov, J., Taylor, B., Yao, J., Pilko, A., Albeck, J., Hoffmann, A., Tsimring, L., and Wollman, R. (2014). Systems biology. Accurate information transmission through dynamic biochemical signaling networks. Science 346, 1370–1373. https://doi.org/10.1126/science.1254933.
- Mair, F., Erickson, J.R., Voillet, V., Simoni, Y., Bi, T., Tyznik, A.J., Martin, J., Gottardo, R., Newell, E.W., and Prlic, M. (2020). A targeted multi-omic analysis approach measures protein expression and low-abundance transcripts on the single-cell level. Cell Rep. *31*, 107499. https://doi. org/10.1016/j.celrep.2020.03.063.
- Jiang, R., Sun, T., Song, D., and Li, J.J. (2022). Statistics or biology: the zero-inflation controversy about scRNA-seq data. Genome Biol. 23, 31. https://doi.org/10.1186/s13059-022-02601-5.
- Foreman, R., and Wollman, R. (2020). Mammalian gene expression variability is explained by underlying cell state. Mol. Syst. Biol. 16, e9146. https://doi.org/10.15252/msb.20199146.
- Kim, J.K., Kolodziejczyk, A.A., Ilicic, T., Teichmann, S.A., and Marioni, J.C. (2015). Characterizing noise structure in single-cell RNA-seq distinguishes genuine from technical stochastic allelic expression. Nat. Commun. 6, 8687. https://doi.org/10.1038/ncomms9687.
- Brennan, M.D., Cheong, R., and Levchenko, A. (2012). Systems biology. How information theory handles cell signaling and uncertainty. Science 338, 334–335. https://doi.org/10.1126/science.1227946.
- Jetka, T., Nienałtowski, K., Filippi, S., Stumpf, M.P.H., and Komorowski, M. (2018). An information-theoretic framework for deciphering pleiotropic and noisy biochemical signaling. Nat. Commun. 9, 4591. https:// doi.org/10.1038/s41467-018-07085-1.
- Jetka, T., Nienałtowski, K., Winarski, T., Błoński, S., and Komorowski, M. (2019). Information-theoretic analysis of multivariate single-cell signaling responses. PLOS Comput. Biol. *15*, e1007132. https://doi.org/10.1371/ journal.pcbi.1007132.
- Komorowski, M., and Tawfik, D.S. (2019). The limited information capacity of cross-reactive sensors drives the evolutionary expansion of signaling. Cell Syst. 8, 76–85.e6. https://doi.org/10.1016/j.cels.2018. 12.006.
- Tang, Y., Adelaja, A., Ye, F.X.-F., Deeds, E., Wollman, R., and Hoffmann, A. (2021). Quantifying information accumulation encoded in the dynamics of biochemical signaling. Nat. Commun. *12*, 1272. https://doi. org/10.1038/s41467-021-21562-0.
- Uda, S., Saito, T.H., Kudo, T., Kokaji, T., Tsuchiya, T., Kubota, H., Komori, Y., Ozaki, Y., and Kuroda, S. (2013). Robustness and compensation of information transmission of signaling pathways. Science 341, 558–561. https://doi.org/10.1126/science.1234511.
- Caldwell, A.B., Cheng, Z., Vargas, J.D., Birnbaum, H.A., and Hoffmann, A. (2014). Network dynamics determine the autocrine and paracrine signaling functions of TNF. Genes Dev. 28, 2120–2133. https://doi.org/ 10.1101/gad.244749.114.



- Tkačik, G., and Walczak, A.M. (2011). Information transmission in genetic regulatory networks: a review. J. Phys. Condens. Matter 23, 153102. https://doi.org/10.1088/0953-8984/23/15/153102.
- Tkačik, G., Callan, C.G., and Bialek, W. (2008). Information capacity of genetic regulatory elements. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. 78, 011910. https://doi.org/10.1103/PhysRevE.78.011910.
- Maity, A., and Wollman, R. (2020). Information transmission from NFkB signaling dynamics to gene expression. PLoS Comp. Biol. 16, e1008011. https://doi.org/10.1371/journal.pcbi.1008011.
- Mukund, A., and Bintu, L. (2022). Temporal signaling, population control, and information processing through chromatin-mediated gene regulation. J. Theor. Biol. 535, 110977. https://doi.org/10.1016/j.jtbi.2021. 110977.
- Rieckh, G., and Tkačik, G. (2014). Noise and information transmission in promoters with multiple internal states. Biophys. J. 106, 1194–1204. https://doi.org/10.1016/j.bpj.2014.01.014.
- Bintu, L., Yong, J., Antebi, Y.E., McCue, K., Kazuki, Y., Uno, N., Oshimura, M., and Elowitz, M.B. (2016). Dynamics of epigenetic regulation at the single-cell level. Science 351, 720–724. https://doi.org/10. 1126/science.aab2956.
- 77. Lavin, Y., Winter, D., Blecher-Gonen, R., David, E., Keren-Shaul, H., Merad, M., Jung, S., and Amit, I. (2014). Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. Cell 159, 1312–1326. https://doi.org/10.1016/j.cell.2014.11.018.
- Davies, A.E., Pargett, M., Siebert, S., Gillies, T.E., Choi, Y., Tobin, S.J., Ram, A.R., Murthy, V., Juliano, C., Quon, G., et al. (2020). Systems-level properties of EGFR-RAS-ERK signaling amplify local signals to generate dynamic gene expression heterogeneity. Cell Syst. *11*, 161–175.e5. https://doi.org/10.1016/j.cels.2020.07.004.
- Faure, A.J., Schmiedel, J.M., and Lehner, B. (2017). Systematic analysis of the determinants of gene expression noise in embryonic stem cells. Cell Syst. 5, 471–484.e4. https://doi.org/10.1016/j.cels.2017.10.003.
- Topolewski, P., Zakrzewska, K.E., Walczak, J., Nienałtowski, K., Müller-Newen, G., Singh, A., and Komorowski, M. (2022). Phenotypic variability, not noise, accounts for most of the cell-to-cell heterogeneity in IFN-γ and oncostatin M signaling responses. Sci. Signal. *15*, eabd9303. https://doi. org/10.1126/scisignal.abd9303.
- Brown, C.R., and Boeger, H. (2014). Nucleosomal promoter variation generates gene expression noise. Proc. Natl. Acad. Sci. USA *111*, 17893–17898. https://doi.org/10.1073/pnas.1417527111.
- Brown, C.R., Mao, C., Falkovskaia, E., Jurica, M.S., and Boeger, H. (2013). Linking stochastic fluctuations in chromatin structure and gene expression. PLoS Biol. *11*, e1001621. https://doi.org/10.1371/journal. pbio.1001621.
- Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene expression in a single cell. Science 297, 1183–1186. https://doi.org/10.1126/science.1070919.
- Swain, P.S., Elowitz, M.B., and Siggia, E.D. (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. Proc. Natl. Acad. Sci. USA 99, 12795–12800. https://doi.org/10.1073/pnas.162041399.
- Okabe, Y., and Medzhitov, R. (2014). Tissue-specific signals control reversible program of localization and functional polarization of macrophages. Cell 157, 832–844. https://doi.org/10.1016/j.cell.2014.04.016.
- van Beek, A.A., Van den Bossche, J., Mastroberardino, P.G., de Winther, M.P.J., and Leenen, P.J.M. (2019). Metabolic alterations in aging macrophages: ingredients for inflammaging? Trends Immunol. 40, 113–127. https://doi.org/10.1016/j.it.2018.12.007.
- Cui, C.Y., Driscoll, R.K., Piao, Y., Chia, C.W., Gorospe, M., and Ferrucci, L. (2019). Skewed macrophage polarization in aging skeletal muscle. Aging Cell *18*, e13032. https://doi.org/10.1111/acel.13032.
- Weisberg, S.P., McCann, D., Desai, M., Rosenbaum, M., Leibel, R.L., and Ferrante, A.W. (2003). Obesity is associated with macrophage accumulation in adipose tissue. J. Clin. Invest. *112*, 1796–1808. https://doi.org/ 10.1172/JCI19246.



- Galli, S.J., Borregaard, N., and Wynn, T.A. (2011). Phenotypic and functional plasticity of cells of innate immunity: macrophages, mast cells and neutrophils. Nat. Immunol. *12*, 1035–1044. https://doi.org/10.1038/ ni.2109.
- Cassetta, L., Fragkogianni, S., Sims, A.H., Swierczak, A., Forrester, L.M., Zhang, H., Soong, D.Y.H., Cotechini, T., Anur, P., Lin, E.Y., et al. (2019). Human tumor-associated macrophage and monocyte transcriptional landscapes reveal cancer-specific reprogramming, biomarkers, and therapeutic targets. Cancer Cell 35, 588–602.e10. https://doi.org/10. 1016/j.ccell.2019.02.009.
- Cohen, Y.C., Zada, M., Wang, S.-Y., Bornstein, C., David, E., Moshe, A., Li, B., Shlomi-Loubaton, S., Gatt, M.E., Gur, C., et al. (2021). Identification of resistance pathways and therapeutic targets in relapsed multiple myeloma patients through single-cell sequencing. Nat. Med. 27, 491–503. https://doi.org/10.1038/s41591-021-01232-w.
- Chang, Y.-L., Rossetti, M., Gjertson, D.W., Rubbi, L., Thompson, M., Montoya, D.J., Morselli, M., Ruffin, F., Hoffmann, A., Pellegrini, M., et al. (2021). Human DNA methylation signatures differentiate persistent from resolving MRSA bacteremia. Proc. Natl. Acad. Sci. USA *118*. e2000663118. https://doi.org/10.1073/pnas.2000663118.
- Burel, J.G., Pomaznoy, M., Lindestam Arlehamn, C.S., Weiskopf, D., da Silva Antunes, R., Jung, Y., Babor, M., Schulten, V., Seumois, G., Greenbaum, J.A., et al. (2019). Circulating T cell-monocyte complexes are markers of immune perturbations. eLife 8, e46045. https://doi.org/ 10.7554/eLife.46045.
- Gratchev, A., Sobenin, I., Orekhov, A., and Kzhyshkowska, J. (2012). Monocytes as a diagnostic marker of cardiovascular diseases. Immunobiology 217, 476–482. https://doi.org/10.1016/j.imbio.2012. 01.008.
- 95. Kapellos, T.S., Bonaguro, L., Gemünd, I., Reusch, N., Saglam, A., Hinkley, E.R., and Schultze, J.L. (2019). Human monocyte subsets and phenotypes in major chronic inflammatory diseases. Front. Immunol. 10, 2035.
- Mathai, S.K., Gulati, M., Peng, X., Russell, T.R., Shaw, A.C., Rubinowitz, A.N., Murray, L.A., Siner, J.M., Antin-Ozerkis, D.E., Montgomery, R.R., et al. (2010). Circulating monocytes from systemic sclerosis patients with interstitial lung disease show an enhanced profibrotic phenotype. Lab. Invest. 90, 812–823. https://doi.org/10.1038/labinvest.2010.73.
- Patel, V.K., Williams, H., Li, S.C.H., Fletcher, J.P., and Medbury, H.J. (2017). Monocyte inflammatory profile is specific for individuals and associated with altered blood lipid levels. Atherosclerosis 263, 15–23. https://doi.org/10.1016/j.atherosclerosis.2017.05.026.
- Brodin, P. (2021). Immune determinants of COVID-19 disease presentation and severity. Nat. Med. 27, 28–33. https://doi.org/10.1038/s41591-020-01202-8.
- Brodin, P. (2022). SARS-CoV-2 infections in children: understanding diverse outcomes. Immunity 55, 201–209. https://doi.org/10.1016/j.immuni.2022.01.014.
- Laing, A.G., Lorenc, A., del Molino del Barrio, I., Das, A., Fish, M., Monin, L., Muñoz-Ruiz, M., McKenzie, D.R., Hayday, T.S., Francos-Quijorna, I., et al. (2020). A dynamic COVID-19 immune signature includes associations with poor prognosis. Nat. Med. 26, 1623–1635. https://doi.org/10. 1038/s41591-020-1038-6.
- 101. Liu, C., Martins, A.J., Lau, W.W., Rachmaninoff, N., Chen, J., Imberti, L., Mostaghimi, D., Fink, D.L., Burbelo, P.D., Dobbs, K., et al. (2021). Timeresolved systems immunology reveals a late juncture linked to fatal COVID-19. Cell *184*, 1836–1857.e22. https://doi.org/10.1016/j.cell. 2021.02.018.
- Scully, E.P., Haverfield, J., Ursin, R.L., Tannenbaum, C., and Klein, S.L. (2020). Considering how biological sex impacts immune responses and

COVID-19 outcomes. Nat. Rev. Immunol. 20, 442–447. https://doi.org/ 10.1038/s41577-020-0348-8.

**Cell Systems** 

- 103. Carvalho, K., Rebboah, E., Jansen, C., Williams, K., Dowey, A., McGill, C., and Mortazavi, A. (2021). Uncovering the gene regulatory networks underlying macrophage polarization through comparative analysis of bulk and single-cell data. Preprint at bioRxiv. https://doi.org/10.1101/ 2021.01.20.427499.
- 104. Albert-Vega, C., Tawfik, D.M., Trouillet-Assant, S., Vachot, L., Mallet, F., and Textoris, J. (2018). Immune functional assays, from custom to standardized tests for precision medicine. Front. Immunol. 9, 2367.
- 105. Boomer, J.S., To, K., Chang, K.C., Takasu, O., Osborne, D.F., Walton, A.H., Bricker, T.L., Jarman, S.D., Kreisel, D., Krupnick, A.S., et al. (2011). Immunosuppression in patients who die of sepsis and multiple organ failure. JAMA 306, 2594–2605. https://doi.org/10.1001/jama. 2011.1829.
- 106. Morris, A.C., Brittan, M., Wilkinson, T.S., McAuley, D.F., Antonelli, J., McCulloch, C., Barr, L.C., McDonald, N.A., Dhaliwal, K., Jones, R.O., et al. (2011). C5a-mediated neutrophil dysfunction is RhoA-dependent and predicts infection in critically ill patients. Blood *117*, 5178–5188. https://doi.org/10.1182/blood-2010-08-304667.
- Urrutia, A., Duffy, D., Rouilly, V., Posseme, C., Djebali, R., Illanes, G., Libri, V., Albaud, B., Gentien, D., Piasecka, B., et al. (2016). Standardized whole-blood transcriptional profiling enables the deconvolution of complex induced immune responses. Cell Rep. *16*, 2777–2791. https://doi.org/10.1016/j.celrep.2016.08.011.
- Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. J. Stat. Software 28, 1–26. https://doi.org/10.18637/jss. v028.i05.
- 109. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888– 1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.
- 110. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140. https://doi.org/10. 1093/bioinformatics/btp616.
- 111. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38, 576–589. https://doi.org/10.1016/j.molcel.2010.05.004.
- Li, J., Bien, J., and Wells, M.T. (2018). rTensor: an R package for multidimensional array (tensor) unfolding, multiplication, and decomposition. J. Stat. Software 87, 1–31. https://doi.org/10.18637/jss.v087.i10.
- 113. Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS A J. Integr. Biol. 16, 284–287. https://doi.org/10.1089/omi. 2011.0118.
- 114. Lin, L., Song, M., Jiang, Y., Zhao, X., Wang, H., and Zhang, L. (2020). Normalizing single-cell RNA sequencing data with internal spike-in-like genes. NAR Genom. Bioinform. 2, Iqaa059. https://doi.org/10.1093/nargab/Iqaa059.
- Tucker, L.R. (1966). Some mathematical notes on three-mode factor analysis. Psychometrika 31, 279–311. https://doi.org/10.1007/ BF02289464.
- Wang, N., Lefaudeux, D., Mazumder, A., Li, J.J., and Hoffmann, A. (2021). Identifying the combinatorial control of signal-dependent transcription factors. PLoS Comp. Biol. 17, e1009095. https://doi.org/10.1371/journal.pcbi.1009095.





#### **STAR**\***METHODS**

#### **KEY RESOURCES TABLE**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse Single Cell Sample Multiplexing Kit	BD Rhapsody	633793
Chemicals, peptides, and recombinant proteins		
LPS	Sigma, B5:055	L2880
murine TNF	Roche	11271156001
Pam3CSK4	Invivogen	tlrl-pms
low MW polyinosine-polycytidylic acid (Poly(I:C))	Invivogen	tlrl-picw
synthetic CpG ODN 1668	Invivogen	tlrl-1668
murine IFNβ	R&D	12401-1
murine IFN $\gamma$	R&D	485-MI
murine IL-4	R&D	404-ML
Critical commercial assays		
BD Rhapsody Express Single-Cell Analysis system	BD	633702
Targeted mRNA and AbSeq Reagent Kit 4 Pack	BD	633771
BD Rhapsody Cartridge Reagent Kit	BD	633731
BD Rhapsody Cartridge Kit	BD	633733
BD Rhapsody cDNA Kit	BD	633773
BD Rhapsody P5000M pipette	BD	633705
BD™ Stain Buffer (FBS)	BD Pharmigen	554656
Deposited data		
BD Rhapsody scRNAseq processed data	This paper	Zenodo: https://doi.org/10.5281/zenodo.7296165 https://github.com/signalingsystemslab/ ResponseSpecificity
BD Rhapsody scRNAseq fastqs	This paper	GSE220970
BMDM bulk RNAseq data	Cheng et al. <sup>8</sup>	GSE68318
Single cell NFκB signaling dynamics	Adelaja et al. <sup>39</sup>	https://doi.org/10.17632/6wksmvh5p4.1
10x scRNAseq of macrophages	Adelaja et al. <sup>39</sup>	GSE162992
10x scRNAseq of peritoneal macrophages	Mogilenko et al. <sup>55</sup>	GSE145562
Experimental models: Cell lines		
Immortalized Myeloid Progenitor-derived macrophages	Singh et al. <sup>25</sup>	N/A
Experimental models: Organisms/strains		
C57BL/6J 90wks old	Jackson Labs	000664
C57BL/6J/DIO high fat diet (60% fat diet)	Jackson Labs	380050
C57BL/6J/DIO controls (10% fat diet)	Jackson Labs	380056
Oligonucleotides		
Rhapsody Custom Panel: ID 1330	BD	633743
Rhapsody Custom Panel: ID 1331	BD	633743
Rhapsody Custom Panel: ID 1332	BD	633743
Rhapsody Custom Panel: ID 1334	BD	633743
Rhapsody Custom Panel: ID 1341	BD	633743
Software and algorithms		
CARET	Kuhn et al. <sup>108</sup>	http://caret.r-forge.r-project.org/
Seurat	Stuart et al. <sup>109</sup>	https://www.rdocumentation.org/packages/ Seurat/versions/3.1.4
BD Rhapsody Targeted Analysis Pipeline (version v1.0)	Shum et al. <sup>23</sup>	https://www.sevenbridges.com/

(Continued on next page)



## Cell Systems Article

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
edgeR	Robinson et al. <sup>110</sup>	https://bioconductor.org/packages/release/ bioc/html/edgeR.html
HOMER	Heinz et al. <sup>111</sup>	http://homer.ucsd.edu/homer/
SLEMI	Jetka et al. <sup>66</sup>	https://cran.r-project.org/web/packages/ SLEMI/index.html
rtensor	Li et al. <sup>112</sup>	https://cran.r-project.org/web/packages/ rTensor/index.html
clusterProfiler	Yu et al. <sup>113</sup>	https://bioconductor.org/packages/release/ bioc/html/clusterProfiler.html
ISnorm	Lin et al. <sup>114</sup>	PMID 33575610, PMC7671304
Code for machine learning and Response Specificity Profile	This paper	https://doi.org/10.5281/zenodo.7296165 https://github.com/signalingsystemslab/ ResponseSpecificity

#### **RESOURCE AVAILABILITY**

#### Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Alexander Hoffmann@ucla.edu).

#### **Materials Availability**

This study did not generate new materials.

#### **Data and Code Availability**

- Data have been deposited at Zenodo and GEO and are publicly available as of the date of publication. DOIs are listed in the key resources table.
- All original code has been deposited on the Github repository https://github.com/signalingsystemslab/ResponseSpecificity and is publicly available as of the date of publication. The DOI of an archived version is listed in the key resources table. This includes code for data processing and analysis, and processed single cell data, including raw and normalized counts.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

#### **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

#### **Macrophage Cell Culture**

Macrophages were obtained by differentiating immortalized myeloid progenitors (HoxB4 iMPs).<sup>25</sup> The choice of using HoxB4 immortalized progenitors was multifold: First, Singh et al.<sup>22</sup> showed that HoxB4-iMPDMs (immortalized myeloid progenitor-derived macrophages) had gene expression responses more similar to BMDMs than RAW cells to BMDMs, thus making the HoxB4-iMPDMs a better choice to model primary macrophages than the current most popularly used murine macrophage line, RAW 264.7. Second, using iMPDMs removed the need for using large numbers of mice, making it more economical and animal-friendly. Third, while introduction of HoxB4 may be a concern in terms of the function of the macrophage, the near-clonal nature of the precursor cells rather than the mixture of various precursors in bone marrow may in fact be an advantage. HoxB4-iMPS were differentiated in DMEM/10% FBS + 30% L929 supernatant for a total of 10 days: iMPs were initially thawed into 10cm non-adherent petri dishes for 3 days. On day 0 of differentiation, cells were then washed once and transferred into differentiation media (DMEM/10% FBS, 30% L929 supernatant, 1% PS, 1% L-Glut, b-Me (1:1000)). Cells were replated into 6cm plates with new media on day 7, at a density of ~20k cells/cm<sup>2</sup>. On day 10, the iMP-derived macrophages (iMPDMs) were stimulated with 100ng/mL lipopolysaccharide (LPS, Sigma Aldrich), 10ng/mL murine TNF, and 50µg/mL low molecular weight polyinosine-polycytidylic acid (Poly(I:C)), 100nM synthetic CpG ODN 1668 (CpG), 500U/ml IFN $\beta$ , or media only Untreated control. For polarized macrophages, cells were incubated in 50ng/ml IFN $\gamma$  or 50ng/ml IL4 for 24 hours prior to stimulation on day 10.

#### **Peritoneal Macrophage Experiments**

All mouse work was done following institutional approval under UCLA's accreditation by Association for Assessment and Accreditation of Laboratory Animal Care International (AAALAC). C57BI/6 mice were obtained from Jackson labs. Two male mice were combined for each condition in order to obtain sufficient numbers of cells for the assay: 90wks old (000664 C57BL/6J), 17wks old (380050 C57BL/6J/DIO high fat diet (60% fat diet)), 17wks old (380056 C57BL/6J/DIO controls (10% fat diet)). Peritoneal macrophages were





extracted by injecting 10mL PBS +1% FBS into the peritoneal space, shaking gently, and then pulling out as much fluid as possible, typically  $\sim$ 8ml. Macrophages were plated in DMEM +10%FBS and allowed to rest 24hrs. Floating cells after that time were washed away, and remaining adherent macrophages were stimulated with the same ligand concentrations as for iMPDMs: 100ng/mL lipopolysaccharide (LPS, Sigma Aldrich), 10ng/mL murine TNF (R&D), 50µg/mL low molecular weight polyinosine-polycytidylic acid (Poly(I:C)), 500U/ml IFN $\beta$ , or media only Untreated control. Cells were washed once with cold PBS after 3hrs of stimulation and lifted into suspension for the Rhapsody scRNAseq assay. The investigators were not blinded the identity of the animal models during the experiments or outcome assessment.

#### **METHOD DETAILS**

#### **Gene Panel Selection Algorithm**

To select genes for single-cell targeted gene profiling, we analyzed existing bulk transcriptomic profiling of macrophage responses. Bulk RNAseq data from Cheng et al.<sup>8</sup> was obtained from GEO GSE68318. Counts were converted to counts per million (cpm) using the package *edgeR*,<sup>110</sup> and genes with cpm>4 in at least three samples were retained. Induced genes were gathered by calculating fold changes at each of the 14 stimulus conditions available in the dataset, at each timepoint, against the unstimulated controls. Genes were retained as induced genes if they met the threshold of log<sub>2</sub>(fold change)>2 and p-value <  $10^{-5}$ , which resulted in 1502 genes.

Because PCA identifies a new basis that maximizes variance within the rotated data, it was ideal for identifying genes that varied in expression level across different stimuli. PCA was performed centered and unscaled on the induced genes across all time points for the 14 stimuli in the dataset. The loadings matrix obtained from the PCA was used to calculate a rank score for each gene. The rank was computed as the radial distance of each gene *j* from the origin, over the top 20 PCs: score<sub>*j*</sub> =  $\sum_{x=1}^{20} (PCx_j)^2$ , where  $PCx_j$  is the component *x* loadings value for gene *j*. The top 480 ranked genes were included in the panel, and the remaining 20 genes were manually selected to add genes such as cell type markers, macrophage polarization markers, and transcription factors (Table S1). As a visual confirmation of the approach, k-means clustering was performed on all induced genes and loadings were colored by cluster. As expected, genes with the highest absolute loadings values in each principal component tended to be spread across different clusters, and the top genes in each principal component exhibited distinct patterns across stimuli.

#### Selection of Stimuli for Response Specificity Assay

To identify an optimal set of the most distinct stimuli from the set of 14 used in the bulk transcriptomic data, tensor components analysis (TCA) was performed. TCA is a higher dimensional parallel of PCA – whereas PCA is performed on a *genes×sample* matrix, TCA is performed on a higher-order tensor by folding the gene expression matrix into a *genes×stimuli×timepoint* tensor. The bulk RNAseq data consisted of *N* genes over *S* stimuli with *T* timepoints per stimulus, which formed a third-order tensor *X* of dimensions  $N \times S \times T$  (a three-dimensional array). Tucker decomposition<sup>115</sup> was performed using the package *rTensor*. This decomposes the tensor into a core tensor *G* of dimensions  $R_1 \times R_2 \times R_3$ , multiplied by a matrix  $U^{(i)}$  along each mode:

$$X = G \times {}_{1}U^{(1)} \times {}_{2}U^{(2)} \times {}_{3}U^{(3)}$$

The first five components, which explained 92% of the variance in the data, were retained. The stimulus loadings matrix  $U^{(2)}$  of dimensions  $S \times R_2$  was hierarchically clustered on the first five components, and stimuli that each occupied separate branches of the hierarchical tree were selected. Stimuli that represented whole bacterial organisms or viruses were not selected, in favor of isolated bacterial or viral components.

#### **Rhapsody scRNAseq**

To collect the adherent macrophages for scRNAseq using the Rhapsody platform, macrophage cells were washed 1x with cold PBS, then lifted into suspension by incubating at 37C for 5 minutes with Accutase, which resulted in cell viability typically >85%. Cells were centrifuged at 4C, 400g for 5 minutes, and resuspended in PBS + 2% FBS. Cells were hash-tagged with anti-CD45-hashtags (BD Rhapsody # 633793) and loaded onto the cartridge following manufacturer's instructions (BD Rhapsody # 633771), with the following modifications, which helped ensure sufficient cell viability for the subsequent steps: Incubation with hashtags was performed for 30mins on ice, instead of 20mins at room temperature; only two washes were performed after hashtag incubation to minimize cell loss. Each cartridge was then loaded with a total of  $\sim$ 36k cells across 12 hash-tagged samples ( $\sim$ 3k cells/sample). Libraries were prepared according to manufacturer's instructions (BD Rhapsody # 633771) and sequenced 2x100 on Novaseg 6000.

#### **QUANTIFICATION AND STATISTICAL ANALYSIS**

#### **Motif Enrichment**

Motif enrichment of induced genes and selected genes was performed using HOMER,<sup>111</sup> with a motif search range of -1000 to +100 of the TSS of each gene. Individual motif hits were placed into five categories: bZIP (AP1 family TF motifs), IRF (IRF and ISRE motifs), RHD (Rel Homology Domain NFκB family motifs), ETS (Erythroblast Transformation Specific family TF motifs), and Zf (Zinc finger





motifs). To summarize the overall enrichment of particular transcription factor families within the gene sets, the average -ln(p value) of motifs in each category was calculated, and a second log transform was taken for plot visualization.

#### **Gene Ontology**

Gene ontology on selected genes versus unselected genes was performed using *clusterProfiler*<sup>113</sup> against a background of all genes. Cutoff values of p-value < 0.01, Benjamini-Hochberg q-value <0.05, and minimum gene set size >5 were used. Ontologies were grouped if they had a similarity proportion greater than 0.7. The top three Biological Processes ontology terms for each group were plotted.

#### scRNAseq Data Processing

Raw fastq files were processed using the BD Rhapsody<sup>™</sup> Targeted Analysis Pipeline (version v1.0)<sup>23</sup> hosted on Seven Bridges Genomics. Distribution-Based Error Correction (DBEC)-adjusted UMI counts (molecules per cell) were used in the downstream analysis. Multiplets, cells with undetermined barcodes, and cells with less than 80 features were removed from the analysis. Due to the selected 500 gene panel comprised of largely inducible genes, the assumption that the total number of RNAs per cell is constant does not hold. Counts were therefore normalized using the package *ISnorm*,<sup>114</sup> rather than the more standard approach of dividing by total counts per cell. PCA was performed centered and unscaled using the R function *prcomp*, and UMAP and tSNE were performed on the top 20 PCs.

#### Assignment of Genes to Regulatory Mechanisms

We first pursued a data-driven approach for the assignment, by examining time-series cell population average data from macrophages stimulated with the six stimuli. We then simulated the 7 gene regulatory logics identified by Cheng et al.,<sup>8</sup> collapsing short and long mRNA half-life clusters. Each of the ordinary differential equations used for the 7 regulatory logics followed the same general form:

$$\frac{dmRNA}{\{dt\}} = k_{syn}f(t) - k_{deg} * mRNA,$$

where for single transcription factor logic gates<sup>8</sup>:

$$f(t) = (1 - k_0) * \frac{(K_D * [TF(t)])^n}{1 + (K_D * [TF(t)])^n} + k_0$$

and for two transcription factor OR logic gates<sup>8</sup>:

$$f(t) = (1 - k_0) * \frac{(K_{D_1} * TF1(t))^n + (K_{D_2} * TF2(t))^n + (K_{D_1} * K_{D_2} * TF1(t) * TF2(t))^n}{1 + (K_{D_1} * TF1(t))^n + (K_{D_2} * TF2(t))^n + (K_{D_1} * K_{D_2} * TF1(t) * TF2(t))^n)} + k_0$$

We matched regulatory logics to each gene by assigning the GRS with the lowest RMSD between model and experimental data (see https://github.com/signalingsystemslab/ResponseSpecificity; Cheng et al.<sup>8</sup> and Wang et al.<sup>116</sup>). We then manually curated the model assignments based on evidence provided by the literature, resulting in assignments given in Table S4.

#### Machine Learning Models and Feature Importance

Machine learning classification models were trained using scRNAseq data from naïve macrophages. The data was split 70%/30% into a training group and a testing group. Using only the training data, a random forest model was trained using repeated 10-fold cross validation, with 3 repeats. The parameter *mtry*, which is the number of variables randomly selected as candidate features for each decision tree split, was set to  $\sqrt{\text{(total number of features)}}$ . As alternative models, weighted k-nearest neighbors and neural network model were also trained on the same dataset. Random forest, weighted kNN, and neural network models were implemented using the R package *caret (classification and regression training)*.<sup>108</sup> After the model was trained, the remaining held-out data was tested, with each cell assigned a soft probability prediction for each ligand. The highest probability ligand was the final prediction. Ensemble modeling was performed by majority voting, taking the predicted stimuli from each of the individual models and choosing the most common stimulus predicted for each cell among the different machine learning models. Macrophages of other polarization conditions, M1(IFN<sub>Y</sub>) and M2(IL4), were tested using the model trained on M0 naïve macrophages.

Feature importance was extracted from the trained random forest model, which is calculated by how much information is lost at each node/split of the decision trees. We calculated this value based on Gini impurity:  $\sum_{i=1}^{C} (freq_i \times (1 - freq_i)))$ , across all unique category labels *C*. The feature importance is then the product (decrease in node impurity) \* (probability of reaching that node), scaled so the top feature has a value of 100.

#### **Mutual Information Analyses**

An information theoretic approach was used to identify either individual genes or combinations of genes providing the highest maximum mutual information between ligand identity and gene expression. Error bars on mutual information calculations were done using 10 bootstraps on 50% of the data. Estimation of maximum mutual information was implemented using the R package





*SLEMI*,<sup>66</sup> which uses a statistical learning-based approach to more accurately and more efficiently calculate maximum mutual information for data types with higher dimensional outputs. The max MI was first calculated for each gene individually, using all stimuli (highest theoretical max MI = 2.58bits), as well as for all pairs of stimuli (highest theoretical max MI = 1.0bits). To relate the max MI to the either the mean and variance of each gene, max MI was plotted against either the average Fano factor (*avg. FF* =

 $\left(\sum_{i=1}^{S} \left(\frac{a_i}{\mu_i}\right)\right) / S$ ) across all stimuli S for each gene, or the mean squared deviation (MSD =  $\frac{\sum_{i=1}^{S} (x_i - \bar{x})^2}{S}$ ).

To estimate the maximum mutual information of the best combination of 1, 2, 3, ..., *N* genes, we first started from a list of the top 20 genes that individually had the best max MI value. For each of these single dimension channels, we scanned every combination of two genes, and again ranked the best combination of two genes and retained the top 20. This process was repeated for each additional gene until the gain in max MI for each additional gene leveled off. Retaining only the top 20 sets at each dimension made the calculation more computationally feasible, while still allowing the possibility for gene combinations that are not simply additive of the previous dimension's highest max MI combination.

For gene-specific pairwise calculation of MI, max MI between pairs of stimuli was calculated for each gene at 3hrs using the singlecell RNAseq data. Max MI was plotted against gene regulatory groups between the two stimuli. Correlation coefficients were calculated using the max MI values for signaling pairs for every NFkB signaling codon, and the max MI values for the corresponding gene expression pairs for every NFkB target gene. Genes without any correlation p-value < 0.25 across the six codons were removed from the display. The Pearson's coefficient was plotted on the heatmap, and genes were hierarchically clustered using complete linkage.

#### **Response Specificity Profile**

The Response Specificity Profile is a collection of values that collectively summarize the distinguishability of macrophage responses to different stimuli. The response landscape on which the Response Specificity Profile is calculated was obtained first by principal component analysis performed centered and unscaled on all stimuli across M0, M1(IFN $\gamma$ ), and M2(IL4) cells. This initial matrix can be written as = N × P, where N is the total number of measured genes and P is the stimulated single cells of different macrophage subtypes. This matrix rotation of **M** by PCA represents the landscape of physiological macrophage responses. Calculation of maximum mutual information using the PC scores was then performed for all possible pairs of stimuli. The advantage of using PC scores to perform mutual information analyses lies in the reduction of noise that would otherwise result in overfitting. Overfitting due to the large number of features was observed to result in saturation of the maximum mutual information values to the theoretical maximum for all pairs. Max MI was therefore calculated on the top three PCs (capturing 42% of the variance in the data), using a truncation based on the PCA scree plot, as each subsequent component after Component 5 added only ~1% more to the variance explained. All error bars were generated by 50 iterations of 50% bootstrap resampling of the complete single cell dataset.

The Response Specificity Profile of new samples were evaluated by projection of the new gene expression data onto the dimensionality-reduced space. Letting the initial PCA be defined as

 $S = W^T \times M$ , where S is the  $r \times P$  scores matrix, and W is the  $N \times r$  loadings matrix, the scores of the new projected data is then given by

$$S_{new} = W^T \times M_{new}$$

Since cells from disease models are projected into the same basis, scores from any new projected data now sit in the same lowerdimensional space and can be compared to the Response Specificity of samples within initially defined response landscape. For new samples, the maximum mutual information between ligand and transcriptomic output was again calculated for all available pairs of ligands using the PC scores. The summary score delta Response Specificity Index ( $\Delta$ RSI) was generated by the following:

 $\Delta RSI = \sqrt{\sum (RSI_p - RSI_p^{MO})^2}$  across all pairs of stimuli *p*.

Cell Systems, Volume 14

## Supplemental information

## Quantifying stimulus-response specificity

## to probe the functional state of macrophages

Katherine M. Sheu, Aditya A. Guru, and Alexander Hoffmann



Figure S1. Selection of targeted gene panel and stimuli. Related to Figure 1.

- A) Top: Heatmap of all genes induced with log<sub>2</sub>(FC)>2 in any stimulus from bulk RNAseq data (Cheng et al, 2017). Middle: PCA loadings from bulk RNAseq data. Genes were selected by scoring each gene based on loadings from the top 20 PCs. Bottom: Selected genes are marked (green) on the full heatmap.
- B) Enriched motifs found in all 1502 induced genes vs. the 500 genes selected for target amplification. bZIP, basic leucine zipper; ETS, Erythroblast Transformation Specific; IRF, Interferon Regulatory Factor; RHD, Rel-Homology Domain.
- C) Tensor components analysis on all stimuli and all timepoints. Red boxes highlight stimuli chosen.
- D) PCA performed on all data points together. Pairwise Euclidean distances between points in PCA space are calculated for each timepoint.



## Figure S2. Comparison of single-cell RNAseq platforms, Rhapsody vs. 10x. Related to Figure 1.

- A) Heatmap of bulk RNAseq, and pseudobulk values from Rhapsody or 10x, for the same set of genes.
- B) Violin plots of scRNAseq data for a few example genes for Rhapsody vs 10x.
- C) Distributions of genes measured per cell (top row), or counts per gene (bottom row), comparing Rhapsody (left), 10x for all genes (middle), and 10x for the 500 genes in the Rhapsody panel (right). For genes detected by both platforms, there are similar distributions in the number of genes per cell and the number of read counts per gene, and most genes outside the custom panel had low or zero read counts
- D) Left: Violin plot of percentage of cells with 0 counts for each gene in the 10x or Rhapsody platforms. 100% indicates that all cells are measured at 0 for that gene. Right: Scatterplot comparing 10x vs Rhapsody "percentage of 0's per gene", across all common genes.

Figure S3



## Figure S3. Reproducibility of the macrophage experimental system. Related to Figure 1.

- A) Scatterplot of pseudobulk values across 4 stimuli at the 8hr timepoint for iMPDMs (immortalized myeloid progenitor-derived macrophages) vs. BMDMs (bone marrow derived macrophages), from the Rhapsody scRNAseq platform.
- B) Single cell distributions for example genes for stimulated iMPDMs vs. BMDMs at the 8hr timepoint.
- C) Scatterplot of pseudobulk values across 5 stimuli at the 3hr timepoint for 2 replicates from the Rhapsody scRNAseq platform. Rhapsody replicates of the single-cell data for five stimuli were concordant in means (avg. Pearson's r = 0.86).
- D) Single-cell distributions for example genes from each replicate across stimuli at the 3hr timepoint, indicating concordant stimulus-specific patterns in response distributions.



## Figure S4. Machine learning test probabilities and top gene combinations. Related to Figure 2.

- A) Distribution of prediction probabilities across cells for the test data. Based on its transcriptome, each cell was given a probability of having encountered a particular ligand, and the highest probability ligand was assigned as the prediction. Prediction probabilities helped distinguish weak versus strong assignments: for instance, correct IFNβ predictions consistently had prediction probabilities close to 1, whereas correct CpG predictions had prediction probabilities of ~0.6-0.7 with the second-best choice being P3C.
- B) Comparison of overall accuracy for three different model types (neural net, random forest, k-nearest neighbors) show similar overall accuracies, resulting in a similar prediction accuracy from ensemble-based majority voting from all three models.
- C) Gene expression values for the top gene combinations providing the greatest ligand discrimination, as quantified by max MI, for M0 macrophages using 1, 2, or 3 genes. Units are log<sub>2</sub>(normalized counts + 1).



## Figure S5. Gene regulatory strategy assignments and relationship between NFκB signaling dynamics and gene expression specificity. Related to Figure 4.

- A) Graphical summary of Table S4, including mathematically modeled gene-assignments (first row), motif analyses, assignments made in prior literature, and assignments used in the present paper (last row). GRS assignments are not definitive but remain working hypotheses. Wang et al, 2021 made multiple possible assignments per gene, and gray bars in the "Wang 2021 BMDMs" row indicate additional assignments.
- B) Motif enrichment for genes in each of the assigned clusters. Only motifs significantly enriched at α < 0.001 in at least one cluster are shown.
- C) Correlation between max MI provided by features of NFκB signaling dynamics and max MI retained in gene expression at 3hrs after stimulation. Stimulus information in *Cxcl10* expression best correlates with information provided by pairwise differences in NFκB Total activity, *Tnf* with Oscillations, and *Clec4e* with Speed.
- D) As counterexamples, Cxcl10 does not correlate with pairwise differences in NFkB Oscillations or Duration.



## Figure S6. Stimulus-specific genes and changes to their Response Specificity in polarized macrophages. Related to Figure 5.

- A) PCA and UMAP of M0, M1, and M2 macrophages at baseline, 0hrs. M1 and M2 macrophages have been treated with IFNγ or IL4 for 24hrs.
- B) Macrophage marker gene Adgre1 is expressed in all three conditions.
- C) M1 marker Cd86 at 0hrs, and M1 marker Nos2 at 0.25hour, combining cells from all stimuli.
- D) M2 marker genes Arg1, Retnla, Chil3, at Ohrs.
- E) Left: UMAP of M0, M1, M2 cells together, using the top 20 principal components. Right: PCA loadings from PCA on M0, M1, M2 cells together.
- F) Correlation of the change in max MI for M1 vs M2 macrophages, each compared to M0, for all genes. Colors indicate categories of genes that increase or decrease in Specificity in M1 or M2 vs. M0.
- G) Biological processes that lose specificity in M1 or M2 macrophages compared to M0 naïve macrophages.



## Figure S7. Mouse peritoneal macrophage subpopulations at steady-state. Related to Figure 7.

- A) tSNE of scRNAseq data of all peritoneal cells from young and old mice.
- B) Overlay of marker genes for macrophages, T-cells, and B-cells, to identify macrophages.
- C) Overlay of marker genes for potential macrophage subpopulations on the tSNE plot of all cells.
- D) Top: Clustering of all young and old peritoneal cells. Bottom: Zoom-in of identified macrophage clusters.
- E) Clusters 2,7,8,10,11 are positive for macrophage marker *Adgre1* or are Ccr2-dependent. Other marker genes clarify additional subpopulations.
- F) Subpopulation proportions are similar in young and old mice, with old mice having proportionally more Ccr2+ (clusters 8, 10) macrophages, as well as proportionally more cluster 2 macrophages.
- G) Cluster marker genes on the tSNE of the Rhapsody stimulus-response scRNAseq data of macrophages from LFD, OLD, HFD mice. Violin plots show cluster marker gene expression within each of the mouse types.