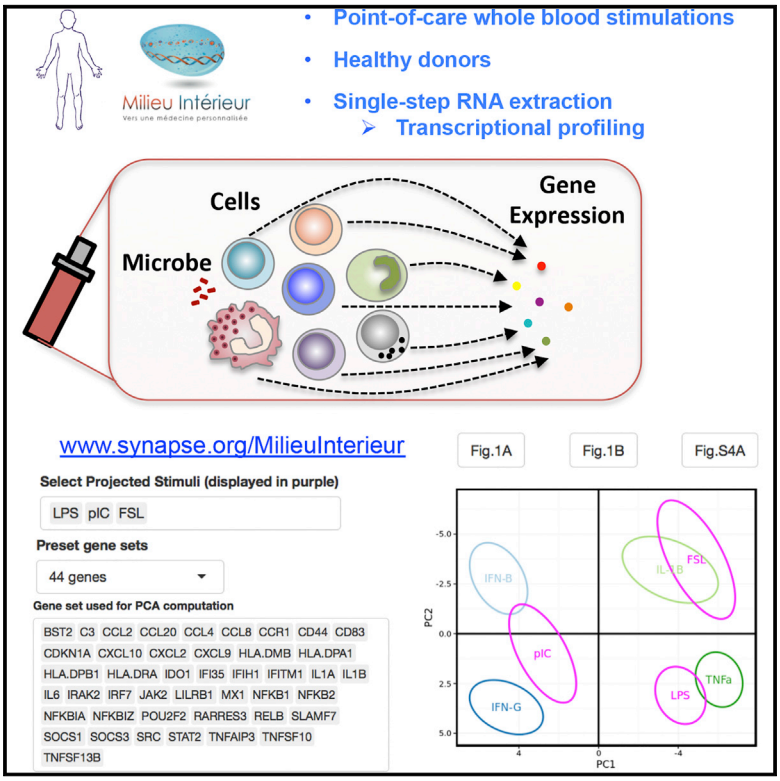
Resource

Cell Reports

Standardized Whole-Blood Transcriptional Profiling Enables the Deconvolution of Complex Induced Immune Responses

Graphical Abstract



Highlights

- Standardized whole-blood stimulation, single-step RNA extraction, and gene expression
- Response to TLR ligands and microbes can be captured by induced cytokine signatures
- A 44-gene signature allows deconvolution of complex responses
- Online R-Shiny application permits interactive data-mining

Authors

Alejandra Urrutia, Darragh Duffy, Vincent Rouilly, ..., Lluis Quintana-Murci, Matthew L. Albert, Milieu Intérieur Consortium

Correspondence

quintana@pasteur.fr (L.Q.-M.), albertm@pasteur.fr (M.L.A.)

In Brief

Urrutia et al. test the hypothesis that responses to TLR ligands or microbes can be captured by induced cytokine signatures. They identified 44 genes that improved segregation of complex stimuli. They also provide reference values that reflect natural variation of immune responses among humans and launched a companion online R-Shiny application for interactive data mining.

Accession Numbers

GSE85176

Urrutia et al., 2016, Cell Reports *16*, 2777–2791
September 6, 2016
http://dx.doi.org/10.1016/j.celrep.2016.08.011

Cell Reports
Resource

Standardized Whole-Blood Transcriptional Profiling Enables the Deconvolution of Complex Induced Immune Responses

Alejandra Urrutia,[1,2,10,11] Darragh Duffy,[1,2,3,11] Vincent Rouilly,[3] Céline Posseme,[3] Raouf Djebali,[3] Gabriel Illanes,[3,4,5] Valentina Libri,[3] Benoit Albaud,[6] David Gentien,[6] Barbara Piasecka,[3] Milena Hasan,[3] Magnus Fontes,[4,7] Lluis Quintana-Murci,[8,9,*] Matthew L. Albert,[1,2,3,10,12,*] and Milieu Intérieur Consortium

[1]Laboratory of Dendritic Cell Immunobiology, Department of Immunology, Institut Pasteur, Paris 75015, France
[2]INSERM U1223, Paris 75015, France
[3]Center for Translational Research, Institut Pasteur, Paris 75015, France
[4]IGDA, Institut Pasteur, Paris 75015, France
[5]Centro de Matemática, Facultad de Ciencias, Universidad de la República, 11200 Montevideo, Uruguay
[6]Institut Curie, Centre de Recherche, Département de recherche translationnelle, Plateforme de Génomique, Paris 75005, France
[7]Centre for Mathematical Sciences, Lund University, 221 00 Lund, Sweden
[8]Laboratory of Human Evolutionary Genetics, Department of Genomes and Genetics, Institut Pasteur, Paris 75015, France
[9]CNRS URA3012, Paris 75015, France
[10]Department of Cancer Immunology, Genentech Inc., San Francisco, CA 94080, USA
[11]Co-first author
[12]Lead Contact
*Correspondence: quintana@pasteur.fr (L.Q.-M.), albertm@pasteur.fr (M.L.A.)
http://dx.doi.org/10.1016/j.celrep.2016.08.011

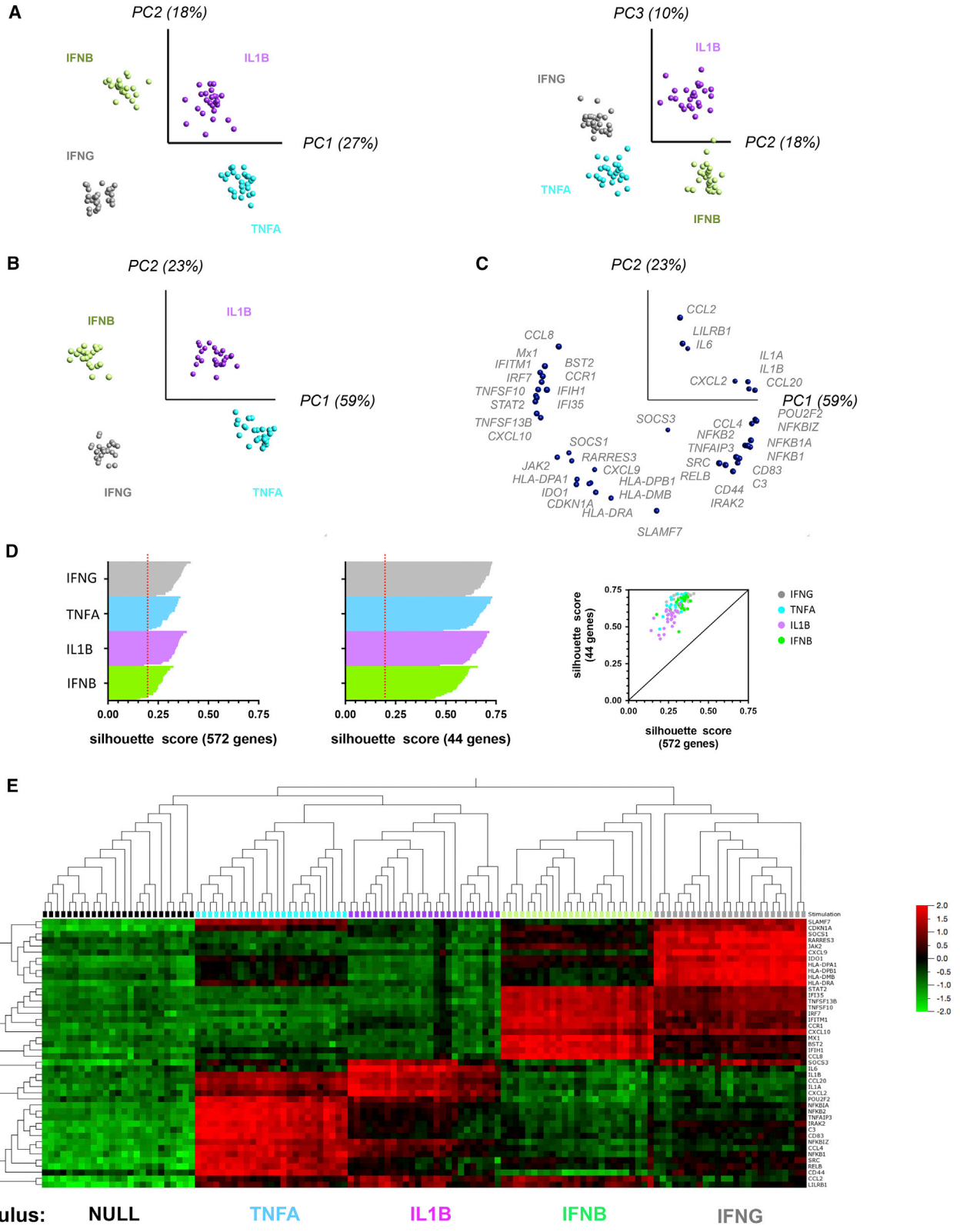
SUMMARY

Systems approaches for the study of immune signaling pathways have been traditionally based on purified cells or cultured lines. However, in vivo responses involve the coordinated action of multiple cell types, which interact to establish an inflammatory microenvironment. We employed standardized whole-blood stimulation systems to test the hypothesis that responses to Toll-like receptor ligands or whole microbes can be defined by the transcriptional signatures of key cytokines. We found 44 genes, identified using Support Vector Machine learning, that captured the diversity of complex innate immune responses with improved segregation between distinct stimuli. Furthermore, we used donor variability to identify shared inter-cellular pathways and trace cytokine loops involved in gene expression. This provides strategies for dimension reduction of large datasets and deconvolution of innate immune responses applicable for characterizing immunomodulatory molecules. Moreover, we provide an interactive R-Shiny application with healthy donor reference values for induced inflammatory genes.

INTRODUCTION

The initiation of inflammatory responses is typically triggered by a local event engaging sentinel cells, leading to the subsequent recruitment and accumulation of leukocytes. This process can result in the elimination of the initial cause of tissue disruption, the clearance of dying cells, and establishes a path toward tissue resolution. Cytokines mediate cell-to-cell communication, acting to recruit immune cells to inflammatory microenvironment and drive the required effector mechanisms. Despite the inherent complexity of these processes in natura, analyses of inflammation have typically focused on the decision-making circuits within cells, and, in most cases, have been restricted to single cell types (Amit et al., 2009; Jovanovic et al., 2015; Lee et al., 2014). Several other studies have assessed in vivo responses to vaccination, typically performing sampling over time to assess induced protein, mRNA expression, and seroconversion (Banchereau et al., 2014; Li et al., 2014; Tsang et al., 2014). While informative, these latter approaches permit the testing of only one stimulation condition per individual and are restricted to qualified or experimental vaccines. To properly account for inter-individual variability in the deconvolution of complex immune responses, both simple (synthetic or purified ligand) and complex (live or heat-killed microbe), stimulations must be performed in the same donor and at the same time, and standardized approaches for all steps from sample collection to analysis must be applied.

To test the hypothesis that responses to Toll-like receptor ligands or whole microbes can be captured by the transcriptional signature of key effector cytokines, we employed a standardized whole-blood stimulation approach with an automated single-step RNA extraction and hybridization gene array readout. Stimulations were performed at the point-of-care, using syringe-based medical devices (TruCulture tubes), in a pilot study that consisted of 25 well-characterized healthy individuals of European ancestry (Thomas et al., 2015). Previously, we reported the testing of protein signatures present in the culture supernatant (Duffy et al., 2014). Herein, we used the cell pellets extracted



(legend on next page)

from the TruCulture stimulation systems to define the transcriptional response to clinically relevant cytokines; interferon-alpha 2A (IFN-α), interferon-beta 1 (IFN-β1), interferon-gamma (IFN-γ), tumor necrosis factor alpha (TNF-α), and interleukin 1-beta (IL-1β). By defining unique and distinct gene expression signatures of cytokine-induced transcription, it was possible to test the clustering and classification of responses to Toll-like receptor (TLR) agonists or whole microbes (including heat killed [HK] gram-negative bacteria, HK gram-positive bacteria, HK fungi, live mycobacteria and viruses). Our results demonstrate the ability to define complex stimuli in terms of the underlying cytokine loops. Moreover, we provide reference values that reflect the degree of naturally occurring variation of immune responses among healthy individuals originating from a homogeneous European background. These data have been made available as a reference for the community, accessible through an online R-Shiny application that permits data-mining using the analytical methods presented.

RESULTS

Distinct Transcriptional Signatures Induced by the IFN-β, IFN-γ, IL-1β, and TNF-α Cytokines

To perform ex vivo stimulation while preserving physiological cellular interactions, we utilized syringe-based medical devices for activating immune cells present in whole blood. Based on initial dose-finding studies, quality assurance, solubility, and stability testing[8], we prioritized stimuli for development in TruCulture whole-blood collection and culture devices (Myriad RBM). After 22 hr stimulation, insertion of a valve separator yielded a cell pellet that was stabilized in Trizol LS and stored at −80°C for subsequent mRNA expression analysis utilizing the NanoString nCounter technology (Figure S1A). Due to the Trizol content in our samples and to minimize pre-analytical biases, we established an automated mRNA single-step chloroform-free extraction protocol (Tecan script provided on-line, see http://www.pasteur.fr/labex/milieu-interieur). Direct comparison with conventional RNA extraction protocols indicated excellent correlation in gene expression counts between the two extraction methods (Spearman's rank-order correlation, $r_s > 0.99$, Figure S1B). Expression data were normalized with nSolver Analysis Software (NanoString), using four housekeeping genes: *RPL19*, *TBP*, *POLR2A*, and *HPRT* (Figures S1C–S1F). These four housekeeping genes were selected following the application of the geNorm method (Vandesompele et al., 2002), an established algorithm for identifying stable housekeeping genes. The selection of these genes is supported by their strong correlations pre- and post-stimulation ($r_s > 0.9$) across the 25 donors, in contrast with those housekeeping genes that were discarded ($r_s < 0.7$) (Figure S1D and data not shown). The overall rationale for the selection of the NanoString platform, as compared to other transcriptional profiling strategies, is presented in Table S1. This choice was validated by the high reproducibility of the data obtained when experiments were performed at different times or at separate institutional core facilities ($r_s > 0.98$, Figure S1B).

To assess the signatures induced by cytokine stimulation, we analyzed the expression data of a total of 572 genes in the 25 donors, using unsupervised principal component analysis (PCA) (Figure 1A). The PCA revealed strong clustering of stimuli-specific responses, with the first three principal components (PCs) explaining 55% of the total variance; PC1 separated IL-1β and TNF-α from IFN-β and IFN-γ, and PC2 distinguished TNF-α from IL-1β and IFN-β from IFN-γ. Of note, the response to IFN-α was also tested and found to be similar to that of the IFN-β response (t test with $q < 0.05$ reported no variables as significantly different between the two stimuli) (Figure S3), and therefore, IFN-α was excluded from further analyses.

To reduce the dimensionality of the data and exclude genes that did not contribute to unique cytokine-induced signatures, we next defined the differential gene expression for each stimulus with respect to the null control using linear support vector machine (SVM) approaches (Burges, 1998). This enabled us the selection of predictive cytokine gene signatures from gene lists ranked according to a paired t test (individual stimulus versus null condition). Bootstrapping of data in the SVM training phase ensured robust results (details provided in the Experimental Procedures). The union of the selected cytokine gene signatures yielded a set of 44 genes that separated the four cytokine stimuli (Table 1). The resulting PCA projection revealed that the four stimulation conditions could be separated into four clearly distinct clusters based on the expression levels of these 44 genes, with PC1 and PC2 capturing 82% of the total variance (Figure 1B). The 44 genes are represented on a biplot—a synchronized dual projection of the variables that drive the loading of the PC vectors (Figure 1C). To quantify the improved clustering provided by this approach, we calculated silhouette scores, i.e., a measure of the distance between the respective k-means clusters, reported for each sample based on the likelihood to localize into one cluster as compared to any of the three other defined clusters. Comparison between the scores that

Figure 1. Distinct Gene Expression Signature Induced by Cytokine Stimulation

Whole-blood stimulation was performed on 25 healthy donors using TruCulture systems pre-loaded with IFN-β (pale green), IFN-γ (gray), IL-1β (purple), and TNF-α (turquoise). Principle component analysis (PCA) was used to project mRNA expression data from 572 genes employing Qlucore Omics Explorer v3.1. Prior to applying PCA, values for each of the 572 mRNA were log transformed, centered to a mean value of zero across each donor, and scaled to unit variance. The four cytokine stimuli are indicated by the colored circles and the vector position of each of the 25 donors is represented.

(A) Left: PC1 versus PC2. Right: PC2 versus PC3. The percentage of variance captured by each PC is indicated.

(B) PCA on filtered gene expression data; first for differential gene expression (paired t test comparing each cytokine with null and a q value cut-off of 10^{-3}); followed by the classification of samples using linear support vector machine (SVM) approaches, and genes ranked according to a paired t test, yielding a union gene set of 44 genes.

(C) A bi-plot of the 44 gene set variable PCA is depicted.

(D) Silhouette scores for each cytokine IFN-β (green), IFN-γ (gray), IL-1β (purple), and TNF-α (turquoise) based on the complete 572-gene set and the selected 44-gene set.

(E) Hierarchical clustering of the donors based on the filtered gene list and four cytokine stimuli and Null condition showing the unique and overlapping expression.

Table 1. Cytokine Gene Signature that Defines Transcriptional Response to IFN-β, IFN-γ, IL-1β, and TNF-α

Gene Name	Associated Cytokine	q Value (Stim versus Null)	q Value (ANOVA on Four Cytokine Stimuli)
BST2	IFN-β	4.3×10^{-43}	4.6×10^{-43}
C3	TNF-α	1.6×10^{-64}	2.3×10^{-64}
CCL2	IL-1β	3.1×10^{-21}	2.7×10^{-21}
CCL20	IL-1β	6.3×10^{-62}	1.6×10^{-61}
CCL4	TNF-α	4.1×10^{-57}	7.9×10^{-57}
CCL8	IFN-β	8.8×10^{-53}	9.6×10^{-53}
CCR1	IFN-β	8.8×10^{-33}	9.0×10^{-33}
CD44	TNF-α	3.2×10^{-58}	5.8×10^{-58}
CD83	TNF-α	1.4×10^{-59}	2.4×10^{-59}
CDKN1A	IFN-γ	1.2×10^{-41}	1.3×10^{-41}
CXCL10	IFN-β	5.3×10^{-51}	5.7×10^{-51}
CXCL2	IL-1β	8.4×10^{-39}	7.5×10^{-39}
CXCL9	IFN-γ	4.0×10^{-43}	4.0×10^{-43}
HLA-DMB	IFN-γ	3.8×10^{-61}	2.5×10^{-61}
HLA-DPA1	IFN-γ	4.2×10^{-51}	3.5×10^{-51}
HLA-DPB1	IFN-γ	3.5×10^{-51}	2.7×10^{-51}
HLA-DRA	IFN-γ	4.0×10^{-45}	3.9×10^{-45}
IDO1	IFN-γ	2.2×10^{-61}	1.2×10^{-61}
IFI35	IFN-β	3.4×10^{-55}	2.7×10^{-55}
IFIH1	IFN-β	2.3×10^{-54}	2.2×10^{-54}
IFITM1	IFN-β	5.7×10^{-49}	6.1×10^{-49}
IL1A	IL-1β	1.1×10^{-59}	2.0×10^{-59}
IL1B	IL-1β	7.8×10^{-83}	2.9×10^{-82}
IL6	IL-1β	1.9×10^{-67}	6.9×10^{-67}
IRAK2	TNF-α	4.8×10^{-62}	9.8×10^{-62}
IRF7	IFN-β	3.4×10^{-56}	2.3×10^{-56}
JAK2	IFN-γ	5.8×10^{-51}	5.0×10^{-51}
LILRB1	IL-1β	1.6×10^{-37}	1.5×10^{-37}
MX1	IFN-β	1.4×10^{-61}	6.3×10^{-62}
NFKB1	IL-1β	8.7×10^{-52}	1.0×10^{-51}
NFKB2	TNF-α	2.0×10^{-64}	3.6×10^{-64}
NFKBIA	TNF-α	2.6×10^{-67}	3.2×10^{-67}
NFKBIZ	IL-1β	2.1×10^{-61}	3.5×10^{-61}
POU2F2	IL-1β	1.8×10^{-70}	6.6×10^{-70}
RARRES3	IFN-γ	2.2×10^{-49}	2.1×10^{-49}
RELB	TNF-α	1.8×10^{-40}	1.9×10^{-40}
SLAMF7	IFN-γ	9.0×10^{-62}	2.5×10^{-62}
SOCS1	IFN-γ	1.6×10^{-42}	1.6×10^{-42}
SOCS3	TNF-α	9.0×10^{-62}	3.6×10^{-55}
SRC	TNF-α	2.3×10^{-57}	4.3×10^{-57}
STAT2	IFN-β	4.3×10^{-55}	3.8×10^{-55}
TNFAIP3	TNF-α	2.7×10^{-59}	4.9×10^{-59}
TNFSF10	IFN-β	2.3×10^{-58}	9.9×10^{-59}
TNFSF13B	IFN-β	1.7×10^{-57}	1.0×10^{-57}

The union set of 44 genes as selected for each cytokine stimulus using linear support vector machine (SVM) approaches and paired t tests with respect to the null control. The q values for each cytokine as compared to the Null (paired t tests) and within the four cytokines (multi-group ANOVA) are shown.

were based on the complete 572 gene set versus the selected 44 gene set revealed a higher score with reduced dimensionality of the feature list and a focus on those most highly discriminating genes (Figure 1D). While our analyses revealed specific cytokine gene signatures, there was modest overlap in the induced gene lists when the stimulation conditions were compared to the null (Figure 1E). Hierarchical clustering of the filtered gene list displayed the unique and overlapping gene expression for the four cytokine groups (Figure 1E).

To examine the intersection among cytokine-induced genes, we first analyzed the induction of *IFN-β*, *IFN-γ*, *IL-1β* and *TNF-α* gene expression. While none of the four cytokines triggered high levels of type I or type II IFN expression (Figure 2A), IL-1β and TNF-α both induced high expression of *IL-1β* mRNA, and all four cytokine stimuli induced modest expression of *TNF-α* (Figure 2A). These data suggest potential cross-talk among the pathways and highlight a strong feed-forward intercellular spread of IL-1β signaling. While this has been previously shown (Dinarello et al., 1987), there is no mechanistic understanding of how IL-1β activates the inflammasome and triggers caspase-1 activation. Unexpectedly, this analysis revealed two outlier individuals who showed high expression levels of IL-1β-induced *IFN-γ* (marked by red and blue dots, Figure 2A). To establish if the observed high levels of *IFN-γ* expression resulted in higher protein secretion, we re-analyzed our previously published protein dataset (Duffy et al., 2014) generated using samples from the same donors and indeed, the two individuals showed the highest levels of IFN-γ protein in the culture supernatants (Figure 2B). The presence of recombinant protein that was used as the stimulus restricted the interpretation of potential positive feedback loops for the given protein (these data points are masked by a gray box, Figure 2B). In addition to the induction of IFN-γ by the two outlier individuals, we also observed higher expression of several IFN-γ-induced genes, as compared to the other donors studied (Figures 2C–2E). Together, these data support the concept that the induced innate responses include the spreading of signals through cytokine feedback loops and potential cross-talk among the intercellular pathways.

Variable Responses to TLR and Microbe Stimulation Are Captured by Induced Cytokine Response

During vaccination or acute infection, the immune system is exposed to agonists that stimulate Toll-like receptor (TLRs) signaling. In such conditions, small numbers of cells are engaged, triggering in turn the production of cytokines that spread the inflammatory response. To test this concept, we evaluated whether the induced transcriptional responses to the four effector cytokines are capable of capturing the diversity of seven well-defined TLR agonists (Duffy et al., 2014): FSL-1 (FSL, also known as Pam2C) that engages the TLR2-TLR6 heterodimer; poly IC (pIC) that engages TLR3; lipopolysaccharide (LPS) that engages TLR4; flagellin (FLA) that engages TLR5; gardiquimod (GARD) that engages TLR7; R848 that engages both TLR7 and TLR8; and CpG-2216 oligonucleotide (ODN) that engages TLR9. Limiting doses of the respective agonists were selected to more closely reflect in vivo responses and to ensure that we were working within the linear range of physiological responses

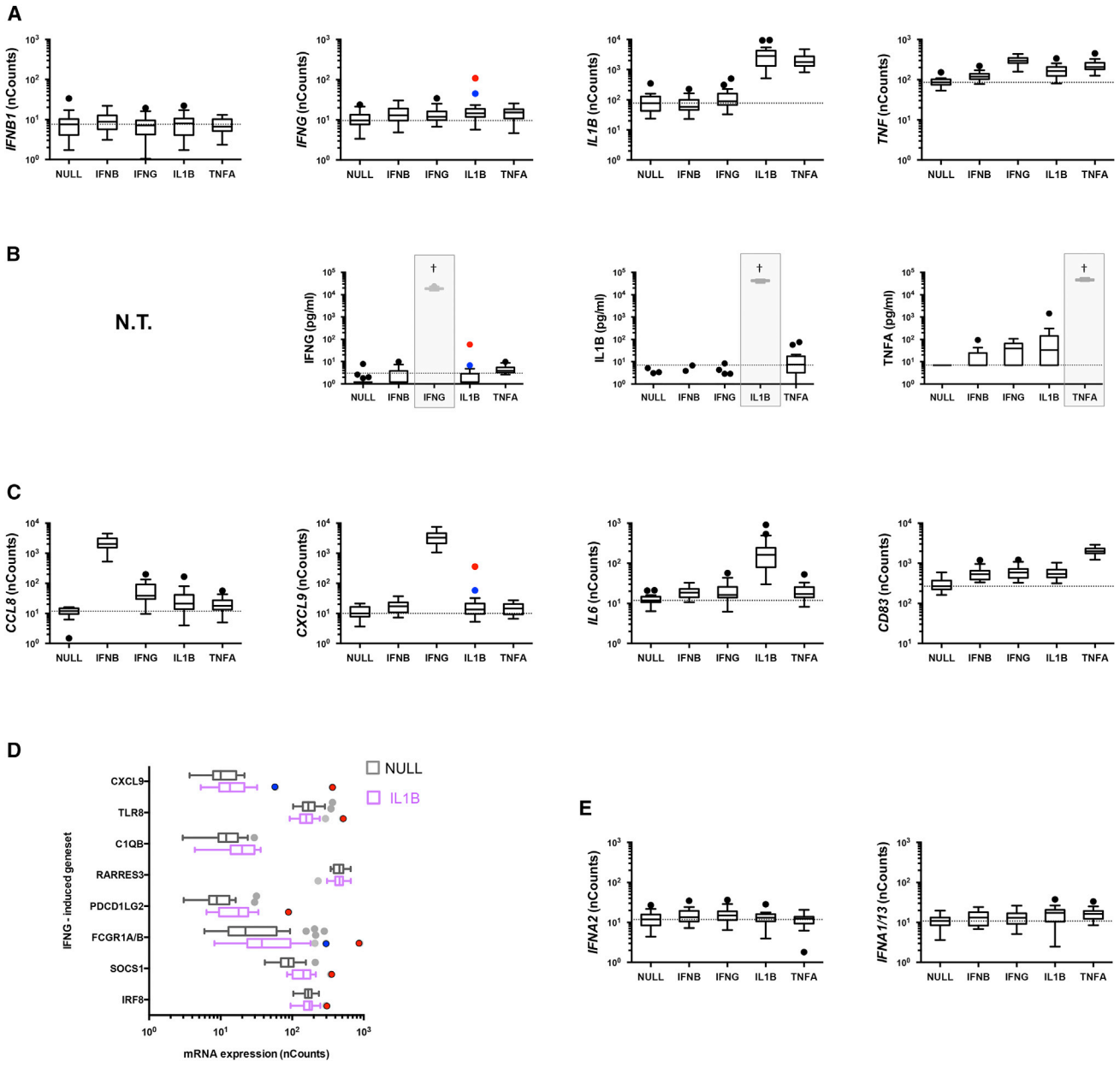


Figure 2. Interactions and Outlier Responses among the Cytokine-Induced Gene Expression Signatures

(A–D) Whole blood from 25 healthy donors was stimulated using the Null, IFN-β, IFN-γ, IL-1β, and TNF-α stimulation conditions. mRNA gene expression (A), absolute nCounts, or induced protein expression (B) are plotted for each of the four genes or gene products: *IFN-β1*, *IFN-γ*, *IL-1β*, and *TNF-α* (N.T. signifies not tested for IFN-β protein; gray shaded boxes mask those protein assays that are detecting the input stimulus in the TruCulture tube). mRNA expression for the most differentially expressed gene is shown (C), one per cytokine stimulus as reported in Table 1 gene list. Top IFN-γ-induced gene expression is shown for the Null (gray) and IL-1β (purple) stimuli (D). Data are represented as box-whisker Tukey plots. Dotted lines indicate the median value for the Null stimulation. Two individual outliers (identified by their induction of *IFN-γ* expression in response to IL-1β stimulation) are indicated using blue and red circles, respectively.
(E) Cytokine stimulation does not induce expression of IFN-α genes. Box-whisker Tukey plots of IFN-α2 and IFN-α1/13 mRNA expression following stimulation with NULL, IFN-β, IFN-γ, IL-1β, and TNF-α. Dotted line indicates median null value.

(please refer to Duffy et al., 2014 or http://www.milieuinterieur.fr/en for details on the dose and source of these reagents). To assess potential similarity in gene expression, we projected the data from each of the seven TLR stimuli onto a fixed PCA coordinate, which was defined by the eigenvectors and eigenvalues of the optimized PCA of the four cytokine-induced mRNA expression data (44 genes defined in Figure 1C). Strikingly, two of the TLR stimuli clustered with a defined cytokine—FLA and FSL vectors both projected onto the IL-1B cluster (Figures S4A and S4B). ODN eigenvectors projected into the IFN-β quadrant,

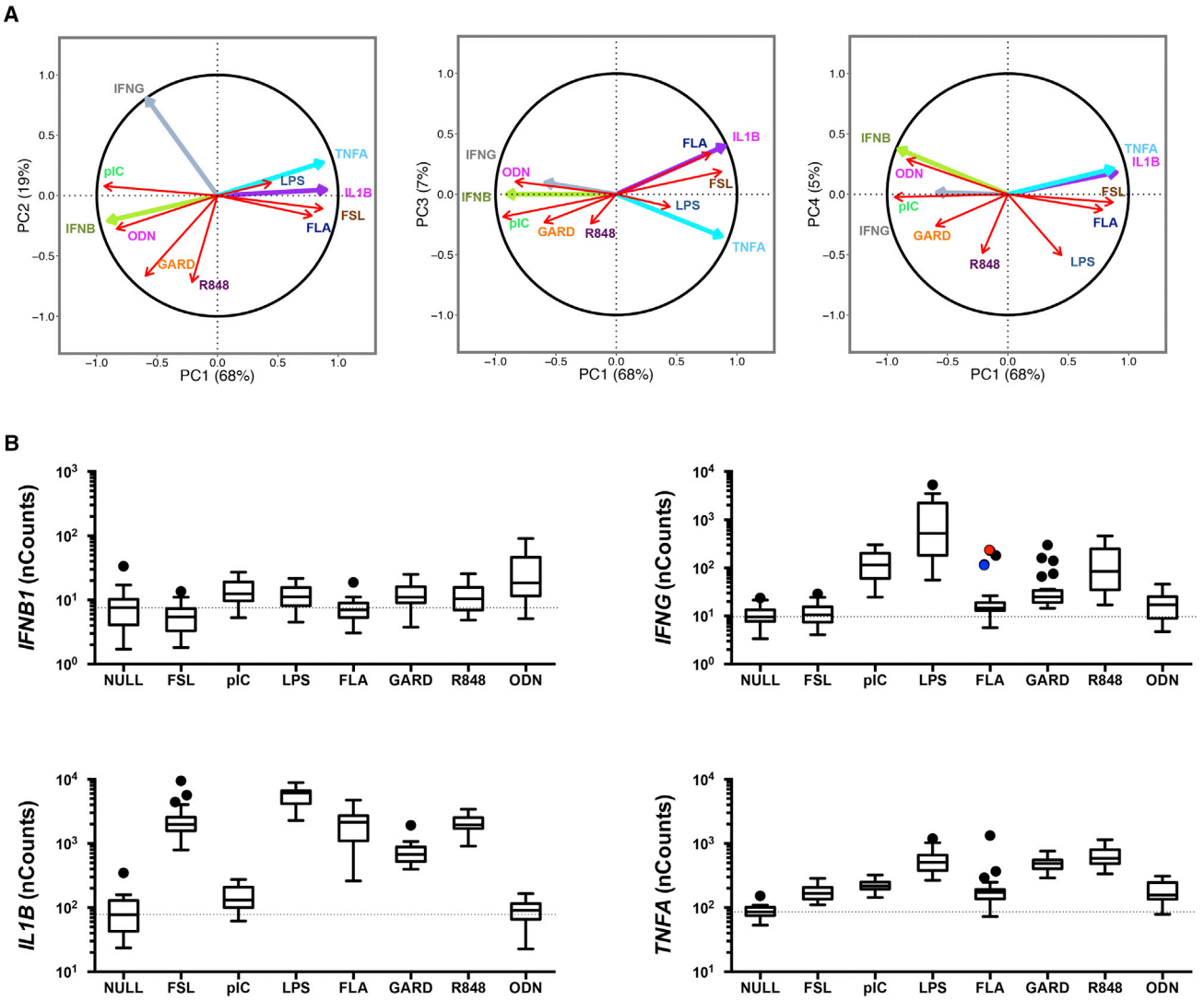


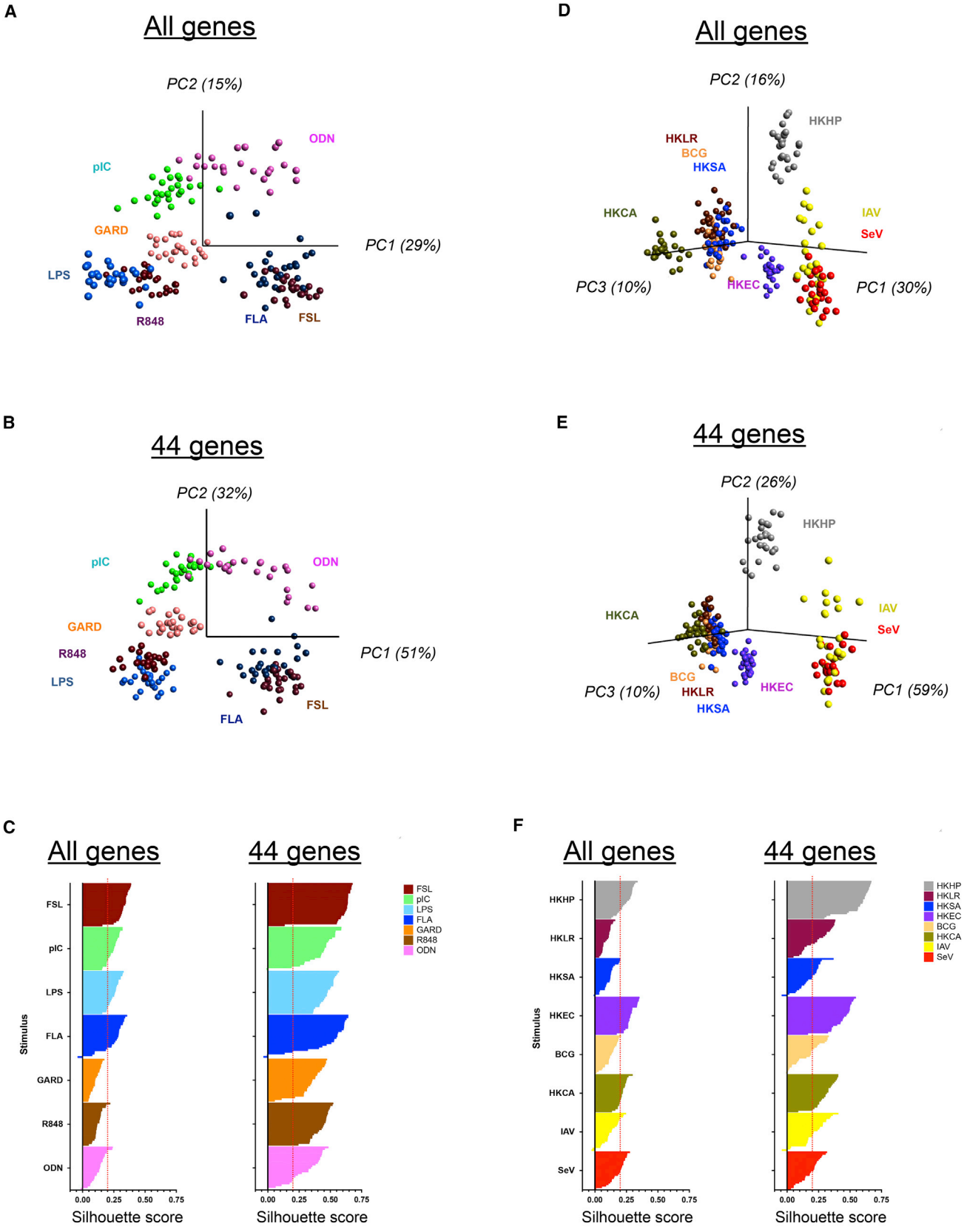
Figure 3. TLR-Induced Gene Expression Can Be Represented as a Function of Cytokine-Induced Gene Signatures

(A) Correlation circles of unit length were constructed using the 44 gene set and PCA loadings were obtained using the gene expression dataset from the four cytokine stimuli (as defined in Table 1). The vectors for TLR-induced gene expression signatures were generated from the median value for the 25 donors, projected onto the correlation circles across the four PC.

(B) *IFN-β1*, *IFN-γ*, *TNF-α*, and *IL-1β* gene expression nCounts are shown for the Null and TLR stimulation conditions. Data are represented as box-whisker Tukey plots. Dotted lines indicate the median value for the Null stimulation. Two individual outliers (identified by their induction of *IFN-γ* expression in response to IL-1β stimulation, Figure 2A) are indicated using blue and red circles, respectively.

with an inter-donor variance in the intensity of gene expression (Figure S4A), which was consistent with our previous study of induced proteins. This analytical approach can be further explored using the online user interface (http://www.synapse.org/MilieuInterieur, http://dx.doi.org/10.7303/syn7059574).

We next represented the data on a correlation circle, as an alternative for visualizing the relationships among stimuli (Figure 3A), allowing us the projection of all TLR stimulation conditions across the four PC axes. When two stimulation vectors are close to the unit circle and are proximal to each other, then they are positively correlated (e.g., FLA and FSL). By contrast, if they are orthogonal to each other, they are not correlated (e.g., FLA and R848). Alternatively, when a stimulation vector is close to the center (e.g., LPS in PC1 versus PC2), it means that information is carried in the other axes (e.g., in the case of LPS almost all variance is carried by PC3 and PC4). Collectively, these data suggest that FLA- and FSL-induced transcriptional signatures are highly correlated to the IL-1β stimulation response; pIC, GARD, R848, and ODN are correlated with type I or type II IFN stimulation; and LPS is intermediate between the two. These results were consistent with the TLR induced expression of *IFN-β1*, *IFN-γ*, *IL-1β*, and *TNF-α* (Figure 3B). One unanticipated result was the similarity between FLA and FSL and the IL-1β gene expression signature. In the case of FLA,



(legend on next page)

we suggest this may be occurring due to the engagement of the intracellular sensor NLRC4, in turn activating caspase-1 (Gay et al., 2014); however, the mechanisms underlying FSL activation of the inflammasome also remains uncharacterized. Notably, these analyses also identified the two outlier individuals discussed above, who showed high expression levels of FLA-induced *IFN-γ* (blue and red dots, Figure 3B).

We applied the same approach to characterize several less well-studied agonists. These included whole β-glucan particles (WGP) derived from *Saccharomyces cerevisiae*, known to engage Dectin-1 and lacking TLR-stimulating activity (Li et al., 2007); lipoarabamanin (LAM), a component of mycobacterial cell walls and an inducer of TLR2; and calcium pyrophosphate dihydrate crystals (CPPD), the etiological agent of pseudogout (Martinon et al., 2006), and a stimulator of NLRP3. Consistent with inflammasome activation, CPPD mapped to the IL-1β cluster, and similar to FSL1, we demonstrate that the LAM-induced gene expression overlaid the IL-1β gene set (Figure S4B). By contrast, WGP induced an mRNA expression signature that projected between IL-1β and TNF-α. Extension of this method may support the classification of unknown adjuvants or innate stimuli.

Next, we performed unsupervised PCA on the TLR-stimulated gene expression data using the entire 572-gene set (Figure 4A). The first two PCs, capturing 44% of the total variance, segregated all TLR stimuli with the exception of FLA and FSL (shown to have similar gene expression patterns), and to a lesser extent LPS and R848. The clustering achieved with the entire dataset was then compared to a PCA plot built using the 44-gene signature, selected for the four effector cytokines (Table 1). Strikingly, the vectors built from the cytokine-gene set fully captured the diversity of responses among the TLR stimuli (Figure 4B). Moreover, the cytokine-optimized gene set provided improved definition of the clusters, as indicated by a higher silhouette scores (Figure 4C). This is most evident for the improved discrimination of LPS from R848 (Figure 4B, see PC2; and an increase in the median silhouette score from 0.26 to 0.46 for LPS and from 0.11 to 0.35 for R848 samples, Figure 4C). These observations support the hypothesis that, in situations of limited agonist concentration and heterogeneous cell types, the characteristic TLR gene signatures can be identified by a limited set of cytokine-induced genes. From the perspective of population-based studies, this introduces the concept that a handful of highly discriminatory gene expression responses are sufficient to distinguish the transcriptional landscape activated by TLR pathways.

To test the robustness of this prediction, we subsequently evaluated the gene expression patterns induced by whole microbes, first using the entire 572-gene set (Figure 4D). The microbes included heat-killed *Escherichia coli O111:B4* (HKEC), *Staphylococcus aureus* (HKSA), *Lactobacillus rhamnosus* (HKLR), *Helicobacter pylori* (HKHP), *Candida albicans* (HKCA), a clinical preparation of live bacillus Calmette-Guerin (BCG), H1N1 attenuated influenza A/PR8 (IAV), and Sendai virus (SeV). The first three principal components, capturing 56% of the total variance, segregated samples from the viral stimuli and HKEC from the other microbes in PC1; HKHP was separated by PC2; and the remaining microbes falling along PC3 with HKCA being distinguishable from HKLR, HKSA, and BCG. Again, we demonstrated improved clustering when using the 44-gene set, as defined by the response to the four effector cytokines (Figure 1B,C). Strikingly, when using the 44-gene set, the variance captured by the first three principle components reached 95% (Figure 4E). Indeed, even with whole microbe stimulation—representing a higher level of biological complexity due to the activation of multiple signaling pathways—we obtained improved silhouette scores for k-means clustering across all stimuli when the PCA was based on the 44-gene set (Figure 4F). For example, the clustering of HKHP samples improved from a median silhouette score of 0.27 to 0.52, when applying the selected 44-gene set in place of the complete 572 genes. Notably, HKLR, HKSA, and BCG were less distinguishable, likely a result of common agonist activity and similar levels of induced cytokines. IAV and SeV also co-segregated for similar reasons. Nonetheless, a doubling of the median silhouette score indicated that, here too, a focused feature list improved clustering of the data. In light of these results, we conclude that a standardized sample collection combined with precise measurement of induced gene expression supports a massive reduction in the dimensionality of the data space, while preserving the ability to discriminate the inflammatory trigger as well as the variability among human donors.

Inter-individual Variable Gene Expression Supports Tracing of Cytokine Loops

We next extended the concept of correlation among the stimulation conditions to shed light onto possible cytokine loops

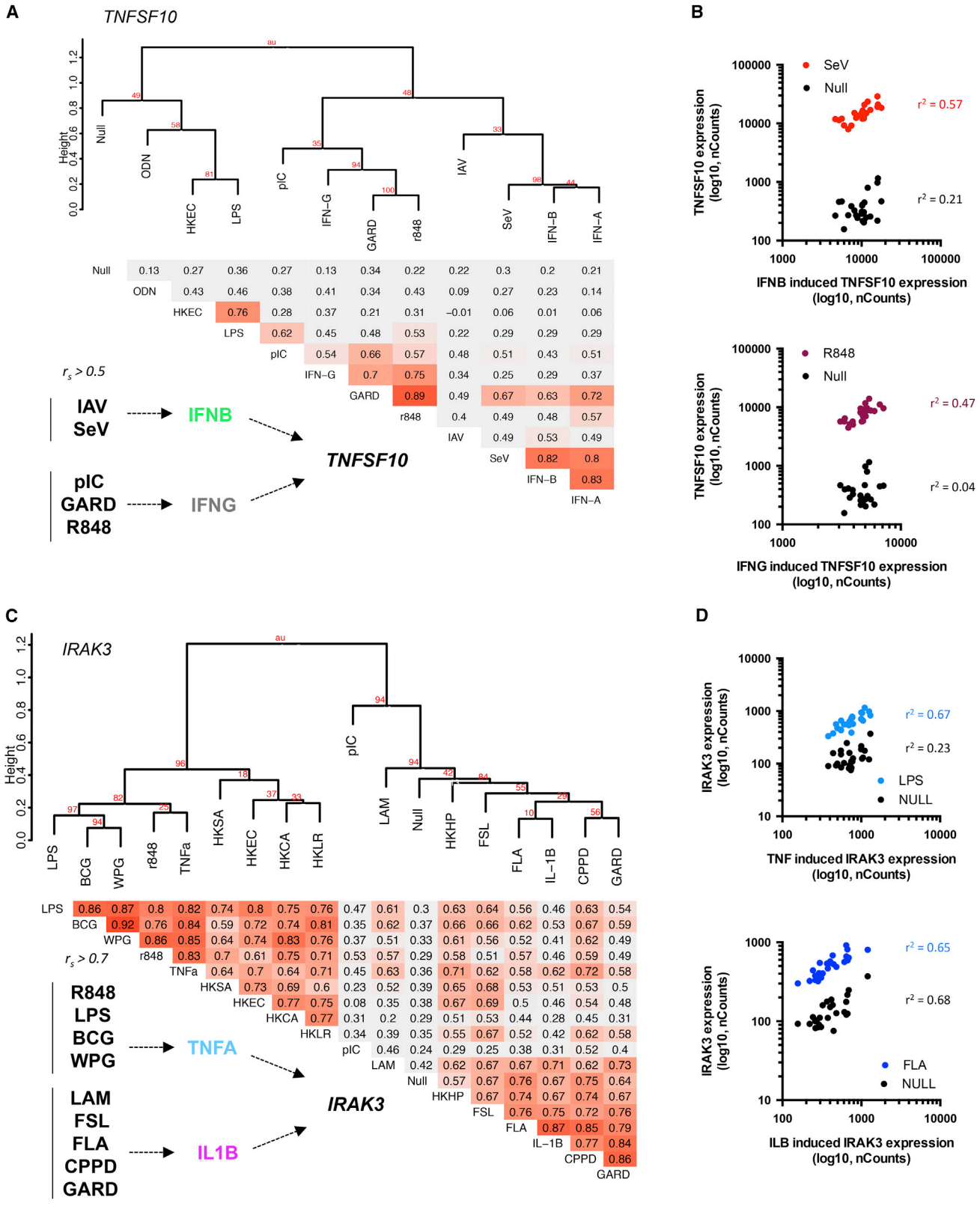
Figure 4. Distinct and Variable Response to TLR Agonist and Microbial Stimulation Can Be Captured Using the Cytokine-Induced 44-Gene Signature

(A and B) Whole-blood stimulation was performed on 25 healthy donors using TruCulture systems pre-loaded with FSL (maroon), pIC (green), LPS (light blue), FLA (dark blue), GARD (orange), R848 (brown), and ODN (pink). Principle component analysis (PCA) was used to project mRNA expression data from 572 genes (A), PC1 versus PC2 (the percentage of variance captured by each PC is indicated). A parallel PCA was constructed using the mRNA expression data from the filtered set of 44 cytokine-induced genes (from gene lists reported in Table 1) (B), PC1 versus PC2 (the percentage of variance captured by each PC is indicated).

(C) Silhouette scores were determined for each sample based on k means clustering (k = 7). Samples are plotted according to TLR stimulus. The red-line indicates a silhouette score of 0.2 (considered a strong fit). The median silhouette score for 572-gene set was 0.19; and for the 44-gene set it was 0.45.

(D and E) Whole-blood stimulation was also preformed using HKHP (gray), HKLR (brown), HKSA (blue), HKEC (purple), HKCA (gray-green), BCG (orange), IAV (yellow), and SeV (red). PCA was used to project mRNA expression data from 572 genes (C); and the parallel PCA was constructed using the mRNA expression data from 44 genes.

(F) Silhouette scores were determined for each sample based on k means clustering (k = 8). Samples are plotted according to microbial stimulus. Note that IAV and SeV were mixed among two clusters (not depicted); and two samples were misclustered using the 572-gene set versus five samples misclustered using the 44-gene set (not depicted). The red-line indicates a silhouette score of 0.2 (considered a strong fit). The median silhouette score for 572-gene set was 0.18; and for the 44-gene set it was 0.26.



(legend on next page)

involved in individual gene expression. This approach provides an exploratory analysis of possible cell-to-cell interactions that can be tested in future experimental studies. Spearman correlation matrices and hierarchical clustering, based on a connected correlation dissimilarity metric, were performed for each gene, and results were bootstrapped to ensure the identified correlations were robust. Using these outputs, we identified cases where the variable responses to TLR or microbe stimulations could be explained by the inter-individual gene expression variance observed when using one of the four cytokine stimuli. To illustrate this observation, the dendrogram depicting the clusters of Spearman correlations and a table indicating the respective r_s coefficients are shown for *TNFSF10* (Figure 5A). A cut-off value of 2-fold expression change greater than the null condition was utilized for inclusion of stimuli in the cluster. Interestingly, the viral stimuli clearly clustered with type I IFN stimulation, with SeV showing a high correlation with IFN-β-induced *TNFSF10* (r_s = 0.82); whereas GARD and R848 clustered with IFN-γ (r_s = 0.7 and 0.75, respectively) (Figures 5A and 5B). As a second example, *IRAK3* is shown, illustrating distinct clustering of bacterial/TLR stimuli with TNF or IL-1β (Figures 5C and 5D). Schematic depictions of the putative stimulus-induced cytokine-mediated expression of *TNFSF10* or *IRAK3* are shown with dotted line arrows provided for illustrative purposes. This analytical approach allows us to predict the distinct cytokine loops that drive common gene expression following stimulation by TLR agonists or microbes. While this modeling approach to population-based data must be experimentally validated, we highlight the possibility that inter-individual variance can be utilized as a means to identify causal pathways driving gene expression, which will support future experimental inquiry.

Microbial Gene Expression Is Defined by Lymphocyte-Derived Cytokines

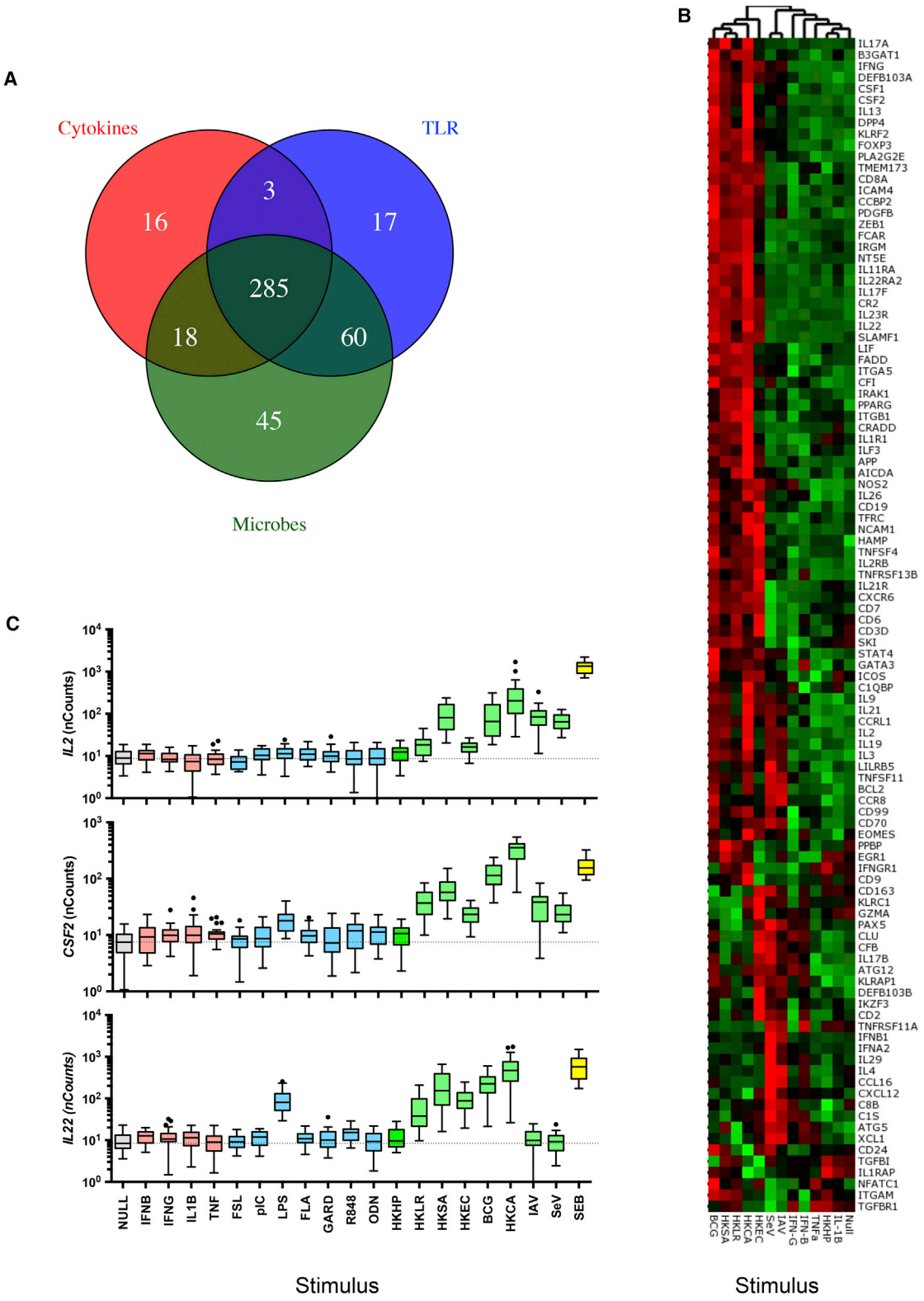
Although the four cytokines studied herein represent major effector pathways in host response and disease pathogenesis, we were cognizant of additional upstream factors that help to specify the inflammatory reaction. To identify other potential effector cytokines, we generated a list of genes upregulated by each stimulus as compared to the null condition (stimulus > null, paired t test $q < 10^{-3}$) and then merged the resulting gene lists for the four cytokines, the seven TLR, and the eight microbial stimuli. A Venn diagram depicts the overlap and intersections in gene expression for these three groups, respectively (Figure 6A). Additionally, we calculated the median gene expression for each stimulus and generated heat maps, clustering by both genes and samples, using either the set of genes that were expressed after microbial but not cytokine stimulation (Figure 6B); TLR but not cytokine stimulation (Figure S6A); and microbial but not TLR stimulation (Figure S6B). Strikingly, the complex stimuli induced a subset of genes indicative of lymphocyte activation. This subset of genes included: (1) transcription factors such as *FoxP3* (highly induced after bacterial stimulation), *EOMES* (induced by HKCA) and *GATA3* (induced by BCG); (2) cytolytic effectors such as *GZMA* (highly induced by HKEC); and (3) anti-microbial genes such as *NOS2* (induced after bacterial stimulation), *DEFB103A* (induced by BCG) and HAMP (highly induced by HKEC) (Figure 6B). Additionally, we detected the differential induction of 18 cytokines, which included *IL2* (induced by HKSA, BCG, HKCA, IAV, and SeV), *CSF2* (highly induced by HKCA), and *IL22* (induced after bacterial and HKCA stimulation) (Figure 6C). As indicated by the comparison with Staphylococcal enterotoxin B (SEB) stimulation and consistent with the presence of microbial antigen-specific T cells within the repertoire of healthy donors (Becattini et al., 2015; Geiger et al., 2009), these cytokine genes likely reflect the activation of lymphocyte subsets (Figure 6C). The characterization of these lymphocyte-derived cytokines may further establish the role of feed-forward cytokine loops in the deconvolution of microbial-induced gene signatures.

DISCUSSION

In this study, we aimed at testing if standardized whole-blood stimulation systems can support the identification of a handful of genes that are capable of deconvoluting complex responses to immune stimulation. We utilized medically relevant stimuli to determine their inflammatory signatures and, in doing so, established the degree of naturally occurring variation present in a population of well-defined healthy donors of European descent. The definition of host immune responses to adjuvants and microbial agents, and subsequent characterization of inter-individual variability in the human population, is of major fundamental interest and provides the necessary foundation for understanding human health and disease pathogenesis. Although functional tests are routinely used in laboratory investigation (Folds and Schmitz, 2003), the standardization of such assays has been challenging. While whole-blood assays are more biologically relevant and introduce less experimental bias than, for example, PBMC stimulation, they are not without technical challenges in particular due to the high levels of globin RNA and enzyme-inhibiting compounds (e.g., heparin interference of reverse transcriptase) (Chaussabel et al., 2010). Previous efforts have focused on removing the globin RNA before downstream analysis, however, these processes can introduce, in turn, higher levels of technical variance as compared to what was achieved with our data generation pipeline (Shin et al., 2014). Specifically, the innovation brought forward in this study is an automated single-step RNA extraction method from whole blood, which minimized pre-analytical bias and generated highly reproducible results when

Figure 5. Correlation among Variable Stimulus-Induced Gene Expression Helps to Trace Cytokine Loops

Gene expression data from all 23 stimulation conditions were used to generate Spearman correlation matrices and hierarchical cluster analysis followed by bootstrapping. The dendrograms shown depict clustering of stimuli based on Spearman correlations for *TNFSF10* (A) or *IRAK3* (C) and the associated triangular matrix indicates the respective pairwise r_s coefficients. Scatter plots for indicated stimulation pairs are shown. Each dot represents the absolute nCount for a single individual of the 25 healthy donors tested for *TNFSF10* (B) or *IRAK3* (D). Red numbers at the intersection of the dendrogram branches indicate approximately unbiased (au) p values, reported as percentage for 1,000 sampled dendrograms. Color scale on tables indicates strength of correlation. Proposed schematics for stimulus-driven cytokine-induced gene expression is proposed using indicated cut-off for r_s.



(legend on next page)

using a gene hybridization read-out. These solutions are essential for multicenter population-based studies, as well as for assays with ambitions for clinical deployment.

Using the reference data presented herein, we tested the hypothesis that responses to TLR ligands or whole microbes can be captured by the transcriptional signature of key effector cytokines. We tested a total of 23 stimulation systems, all built into whole-blood syringes for point of care sampling. Using linear SVM learning, it was possible to identify a 44-gene set, selected based on their ability to differentially cluster cytokine-induced genes. Strikingly, these same genes, when applied to the stratification of responses to TLR ligands or microbes, resulted in improved discrimination among the stimuli as indicated by a marked improvement in silhouette scores. In the era of an increased use of whole-genome transcriptional profiles, our results suggest that limiting the pre-analytical bias introduced by cell separation and non-standardized stimulation protocols may be more important than obtaining greater numbers of measured genes. In addition to sample collection and data analysis standardization, we minimized intrinsic variability by the recruitment of donors of Western European ancestry (third generation born in Metropolitan France). Furthermore, we minimized pre-analytic or environmental sources of variability by applying highly precise inclusion and exclusion criteria (Thomas et al., 2015). To restrict other sources of variability, in addition to the standardization of the assay systems, all donors were sampled at the same time of day (09:00–11:00), during the same week, and in the same location. Such a reliable monitoring of induced immune gene expression responses permitted the classification of inflammatory and host immune responses based on the variance observed in healthy donors.

In addition to defining detailed healthy reference ranges to be considered in future clinical studies, this work permitted the identification of a number of outlier responses. This included two individuals that responded to FLA or IL-1β by producing IFN-γ and in turn the induction of IFN-γ-stimulated genes. Following from this observation, we extended the approach of tracing cytokine loops and gene expression pathways, using inter-individual variance and correlation among the stimulation signatures as a means to deconvolute complex transcriptional responses. This approach may also support the future classification of unknown adjuvants, innate stimuli, new pathogenic agents or the stratification of disease and treatment response. If extended to the study of disease states, it may be possible to classify, for example, subsets of rheumatoid arthritis patients that are responsive to IL-1β versus TNF-α blockade (Gibbons and Hyrich, 2009; McInnes and Schett, 2007).

This reference dataset and the applied analytical approach offers a useful resource to the community, nevertheless, several specific limitations should be highlighted. First, some of the employed TLR stimuli may engage secondary pathways in addition to their commonly ascribed receptors. Notably, the observation that FLA is highly correlated with the IL-1β-induced gene signature suggests that it may also trigger NLRC4 within the whole-blood stimulation systems. This may occur within neutrophils, which express high levels of the NLRC4 inflammasome and release IL-1β (Chen et al., 2014). If correct, it would also help to explain why, despite the high prevalence of dominant-negative forms of TLR5 in Europeans (Barreiro et al., 2009; Hawn et al., 2003), all 25 donors showed an induced response after FLA stimulation (Barreiro et al., 2009). Alternatively, TLR sensor pathways on platelets and neutrophils may be unique in their ability to engage caspase-1 (Hayashi et al., 2003). We also observed that IAV and SeV were highly correlated with pIC, suggesting that the latter is engaging RIG-I like receptors (RLRs) in addition to TLR3. We also acknowledge that, in the natural setting, human immune responses typically occur in mucosal tissues and, as such, stromal cells and tissue resident immune populations such as macrophages and ILCs may need to be considered to fully apply our dataset to physiologic and pathologic responses. Lastly, our analyses consider a single analytical time point only, thus capturing a snapshot of the complexity inherent in dynamic immune responses.

Finally, it is our aim with this resource paper to highlight the growing need to make data more accessible and easier to explore. In line with recent efforts (Gorenshteyn et al., 2015; Speake et al., 2015), we have thus developed an online R-Shiny application software that will allow readers to fully query the dataset based on their specific questions. This application software was built as a direct companion to the presented analyses with publically available R-scripts and downloading options for gene expression data. In sum, the data resource presented here and the available online tools provide a foundation for association studies, kinetic analyses, and in vivo mechanistic experimentation. For example, it remains to be established how the inter-individual variation in gene expression that we identified here is accounted for by host genetic variants (i.e., expression quantitative trait loci [eQTLs]), specifically in cases where gene expression variation is altered upon activation with certain immune stimuli (i.e., response/interaction eQTLs). Conceptually, the strategy to trace inter-cellular cytokine driven gene expression may support such future eQTL association studies, especially in cases where inter-cellular *trans*-eQTL are identified. From a practical viewpoint, the tools will support a path toward more targeted immune monitoring from whole blood, enabling

Figure 6. Microbial-Induced Lymphokines Are Absent from TLR and Cytokine Gene Expression Signatures

(A) Gene expression data from all 23 stimulation conditions were used to generate stimulus-induced signatures (stimulus > null, paired t test with q value cut-off of 10^{-3}). The union sets of cytokine (IFN-β, IFN-γ, IL-1β, TNF-α); TLR (FSL, pIC, LPS, FLA, GARD, R848, ODN); and microbes (HKHP, HKLR, HKSA, HKEC, HKCA, BCG, IAV, SeV) were generated. The Venn diagram indicates the number of shared and unique genes among the three groups of stimuli.

(B) Hierarchical clustering of the donors and genes based on the 105 genes present in the union set of microbes but not cytokines was performed. The median gene expression value was used for each stimulus, with variables log-transformed, mean-centered per donor, and scaled to unit variance. NB, the dendrogram for clustering of genes not shown.

(C) Representative gene expression data are shown for *IL2*, *CSF2*, and *IL22* for each stimuli, as well as Staphylococcal enterotoxin B (SEB) stimulation for reference. Data are represented as box-whisker Tukey plots. Dotted lines indicate the median value for the Null stimulation.

the use of standardized approaches that capture the common variation within the human population.

EXPERIMENTAL PROCEDURES

Donors

Samples were obtained as part of the Milieur Intérieur Healthy Donor Cohort (https://www.clinicaltrials.gov/; NCT01699893). The study protocol was designed and conducted in accordance with the ethical principles of the Declaration of Helsinki and Good Clinical Practices as outlined in the ICH Guideline for Good Clinical Practices. The data were collected under pseudo-anonymized conditions: the identity of the subject is coded in a way that does not allow third-party persons to detect the identity of the person. All subjects (12 male, 13 female) aged 30–39 years old, gave informed consent and were considered as healthy based on medical history, clinical examination, laboratory results, and electrocardiography (ECG). More specific details on criteria to define healthy can be found in previously published work (Thomas et al., 2015).

TruCulture Stimulation

TruCulture tubes were prepared in batch with the indicated stimulus, resuspended in a volume of 2 ml buffered media and maintained at −20°C until use. Blood was obtained from the antecubital vein using a 60 ml syringe containing sodium-heparin (50 IU/ml final concentration). Within 15 min of collection, 1 ml of whole blood was distributed into pre-warmed TruCulture tubes, inserted into a dry block incubator, and maintained at 37°C (+/− 1°C), room air for 22 hr (+/− 15 min). After incubation, a valve was inserted to separate cells from the supernatant and to stop the stimulation reaction. Upon removal of the liquid supernatant, cell pellets were resuspended in 2 ml Trizol LS (Sigma), vortexed for 2 min, and rested for 10 min at room temperature (RT) before −80°C storage.

High-Throughput Standardized RNA Extraction

Samples were randomized and extracted in groups of 95. Cell pellets in Trizol LS were thawed on ice 60 min prior to processing. To complete thawing and RNA release, tubes were vortexed twice for 5 min at 2,000 rpm. Before processing, a centrifugation (3,000 × *g* for 5 min at 4°C) of the thawed samples was performed to pellet the cellular debris generated during the Trizol lysis. The barcoded tubes were loaded in the rack module of the Freedom EVO platform (TECAN) and scanned for sample traceability. For extraction, a modified protocol of the NucleoSpin 96 RNA tissue kit (Macherey-Nagel) was developed and adapted to the Freedom EVO integrated vacuum system. The detailed script for the operation of the TECAN system is provided online (http://www.milieuinterieur.fr/en). In brief, 600 μl of clarified phase of the Trizol lysate was transferred to a deep well plate preloaded with 900 μl of 100% ethanol. The binding mixture was transferred into the silica membrane plate. The columns were washed with buffers MW1 and MW2 (×2) and RNA eluted into 0.5 ml 2D barcoded tubes (ThermoScientific) using 60 μl RNase-free water. As an internal control of the extraction process, a tube containing a defined quantity of spiked RNA was included in each run. To avoid unnecessary freeze and thaw of the RNA, distinct aliquots for quality control and gene expression analysis were prepared, and all aliquots were frozen at −80°C until use.

RNA Quality Controls

RNA concentration was estimated using Qubit RNA HS Assay Kit (Life Technologies) according to the protocol provided by the manufacturer. An automated RNA integrity assessment was performed using the Standard RNA Reagent Kit on a LabChipGX (Perkin Elmer). The RNA quality score (RQS) was calculated using the LabChip System software, and all samples with a RQS greater than four were processed for gene expression analysis.

Selection Criteria for Gene Expression Analysis

NanoString nCounter, a hybridization-based multiplex assay, was selected after comparison with multiple gene expression technologies (microarray, qPCR-based methods) (Table S1). All assays were performed at the genomic platform (Institut Curie), with the exception of the cross platform control comparison performed at Institut Pasteur, Paris. The Human Immunology v2 gene code set was selected as it covers 25 immunology-related gene networks as illustrated by the use of KEGG charts (Figure S2). The code set contains a total of 594 probes (15 correspond to housekeeping genes), of which 572 probes were included in downstream analysis after removing probes mapping to multiple genes and probes aligning to polymorphic regions with greater than two SNPs (Table S2). To this end, the probes were mapped against the human genomic sequence (GRCh37/hg19) with GSNAP (Wu and Nacu, 2010), a splice-aware aligner. A total of 573 out of 594 probes were mapped with 100% identity to the genome. Twelve probes mapped with one to two mismatches in the middle of the sequence, eight probes were misaligned in the first/last 1–9 bp, and one probe did not map at all (PECAM1 located on HG183_PATCH). The misaligned probes were realigned manually using BLASTN against Ab-initio cDNAs database. Of the 594 probes, 15 mapped to more than one genomic location (see Table S2). We removed from further analysis KIR_Activating_Subgroup_1 probe, which mapped to three different genomic locations, as well as three other KIR probes that mapped to multiple locations: KIR_Activating_Subgroup_2, KIR_Inhibiting_Subgroup_1, and KIR_Inhibiting_Subgroup_2. Bioconductor biomaRt package (Durinck et al., 2005) version 2.24.0 was used to query Ensembl (release 75) (Flicek et al., 2014) and retrieve exonic variants that mapped to the same regions as the NanoString probes. We considered only SNPs with minor allelic frequency >0.05 (1000 Genomes Project). Forty-eight probes showed the presence of one to two SNPs in their sequence. HLA-DRB1, HLA-DQA1, and HLA-DQB1 probes contained 4, 9, and 13 SNPs, respectively, and were therefore removed from further analysis.

Gene Expression Analysis

Total mRNA were diluted with RNase-free water at 20 ng/μl in the 12-strip provided by NanoString. We analyzed 100 ng (5 μl) of total RNA from each sample using the Human Immunology kit v2 according to manufacturer's instructions. Each sample was analyzed in a separate multiplexed reaction including in each, eight negative probes and six serial concentrations of positive control probes. Negative control analysis was performed to determine the background for each sample. Of note, we observed variable expression of two negative control probes (NEG B, NEG F), which cross-reacted with bacterial nucleic acid present in two of the TruCulture systems (HKSA and BCG, respectively, Figures S1D and S1E), and thus these probes were not used for data normalization. Data was imported into nSolver analysis software (version 2.5) for quality checking and normalization of data. A first step of normalization using the internal positive controls permitted correction of potential sources of variation associated with the technical platform. To do so, we calculated for each sample the geometric mean of the positive probe counts. A scaling factor for a sample was a ratio of the average across all geometric means and the geometric mean of the sample. For each sample, we multiplied all gene counts by the corresponding scaling factor. Next, for each sample we calculated the background level as the median +2 SD across the six negative probe counts. For each gene in a sample, we subtracted the background level. Finally, to normalize for differences in RNA input we used the same method as in the positive control normalization, except that geometric means were calculated over four housekeeping genes (RPL19, TBP, POLR2A, and HPRT1). These genes were selected using geNorm method (Vandesompele et al., 2002), an established approach for identification of stable housekeeping genes, from the 15 candidate genes provided by NanoString.

Statistical Analysis, Data Visualization, and Software

Principal component analysis (PCA) or singular value decomposition (SVD) was used to decompose the data matrix in a way that is amenable for dimension reduction (Alter et al., 2000). The decomposition was used to orthogonally project both the rows and the columns of the data matrix into lower dimensional space in an optimal way—optimal signifying the retention of as much of the original variance in the dataset as possible. For a comprehensive overview of PCA and the exploratory analysis using dual PCA and the accompanying PCA biplots, we refer to Fontes (2012). Before applying PCA, the variables (mRNA expression levels) were log-transformed, mean-centered per donor, to avoid inter-donor variability obscuring inter-stimuli responses, and finally the variables were scaled to unit variance. The mean-centering per donor is in accordance with the paired structure in the data and paired t tests or

ANOVA were performed throughout. Scaling to unit variance prevents large variances in the data from obscuring the correlation structure in the data. Q values, which are defined as false discovery rate (FDR)-adjusted p values (Benjamini and Hochberg, 1995), were used to define statistical significance.

Correlation circles were generated by computing the median value across the 25 donors, for each of the considered 44 genes; we then transposed the data matrix to consider the four stimulation conditions as the four PCA dimensions; finally, the vectors representing the TLR stimuli were projected onto the four-dimensional PCA. The respective 2D PCA projection plots were made with the R package "FactoMineR" (version 1.28) to compute PCA scores and projected coordinates. Silhouette analysis was used to study the separation distance among the TLR and microbial stimuli. K means clustering was performed using the Open CV library (Bradski and Kaehler, 2008); with settings equal to 100 iterations and 500 attempts and the silhouette scores were computed (Bradski and Kaehler, 2008; Steinhaus, 1956). Cluster number was selected based on the number of stimuli represented in the PCA (k = 7 for TLR, k = 8 for microbes). Note, silhouette coefficients near +1 indicate that the sample is far away from the neighboring clusters; a value of 0 indicates that the sample is on or very close to the decision boundary between two neighboring clusters, and negative values indicate that those samples might have been assigned to the wrong cluster. Bootstrapped hierarchical clustering analysis was performed using the "pvclust" R package (version 1.3-2) using a Spearman-based dissimilarity metric. One thousand trees were sampled to evaluate the robustness of each cluster. Correlation matrices were plotted using the R graphics package ggplot2 (version 1.0.0). Plots were exported from the Qlucore Omics Explorer 3.1 or created using the ggplot2 package (version 1.0.0) on the R platform (version 3.1.1).

Stimulus signatures, consisting of gene lists specific for each of the four cytokines, were created by training a support vector machine (SVM) for each individual stimulus versus null. This approach was used to define stimulus signatures set by a discrete number of variables. In order to discover reasonably complex gene interaction networks among the four stimuli, SVMs were optimized from 12–57 gene subsets (~2%–10% of the total gene number). For all four cytokine stimuli, the optimal classifier determined by the SVM cross validation scheme corresponded to the smallest gene set size. The identified gene lists had perfect accuracy upon ten repeated complete SVM test runs. This shows that a small number of selected variables can predict the specific stimulus used. The small overlap between the established stimulus signatures indicates that the selection strikes a reasonable balance between capturing the complexity in the data and at the same time identifying those important individual genes. The open source C++ software library OpenCV was used to build and evaluate the SVMs (Burges, 1998; Chang and Lin, 2011). For comparison, a kNN classifier was also tested, using the implementation in the OpenCV library with default parameter settings, which gave exactly the same stimulus signatures.

R Shiny (Interactive Web Application) Development

To complement this manuscript, we provide an interactive web application that allows exploration of the dataset presented in this study. The application presents four different types of analytical visualizations: PCA, boxplots, hierarchical clustering, and a searchable reference table. For each visualization, we provide default settings that match figures presented in the manuscript. Visualization controls enable the user to navigate the entire dataset following their own scientific interests. The interactive table provides reference values, based on the 25 healthy donors, which can be directly browsed using a selected method (median expression values, coefficient of variations or q values from paired t tests as compared to the Null condition). The application was implemented using the Open Source R platform, Shiny package (version 0.12.2), ggplot2 package (version 1.0.0), dplyr package (version 0.4.3), and tinyr package (version 0.3.1). All visualization and analysis methods are accessible through a web browser, without the need to install any additional software or possess knowledge of a programming language and is available at https://www.synapse.org/MilieuInterieur (http://dx.doi.org/10.7303/syn7059574).

ACCESSION NUMBERS

The accession number for the expression profiling by array data reported in this paper is GEO: GSE85176.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2016.08.011.

CONSORTIUM

The members of the Milieu Intérieur Consortium are Laurent Abel, Andres Alcover, Kalla Astrom, Philippe Bousso, Pierre Bruhns, Ana Cumano, Caroline Demangel, Ludovic Deriano, James Di Santo, Françoise Dromer, Gérard Eberl, Jost Enninga, Jacques Fellay, Antonio Freitas, Odile Gelpi, Ivo Gomperts-Boneca, Serge Hercberg, Olivier Lantz, Claude Leclerc, Hugo Mouquet, Sandra Pellegrini, Stanislas Pol, Lars Rogge, Anavaj Sakuntabhai, Olivier Schwartz, Benno Schwikowski, Spencer Shorte, Vassili Soumelis, Frédéric Tangy, Eric Tartour, Antoine Toubert, Marie-Noëlle Ungeheuer, Lluis Quintana-Murci, and Matthew L. Albert.

AUTHOR CONTRIBUTIONS

A.U. performed experiments, analyzed data, and wrote the paper. V.R., G.I., and B.P. analyzed data. C.P., R.D., V.L., B.A., D.G., and M.H. performed experiments. D.D., M.F., L.Q.-M., and M.L.A. designed the study, analyzed data, and wrote the paper. M.F., L.Q.-M., and M.L.A. contributed equally.

ACKNOWLEDGMENTS

This work benefited from support of the French government's Invest in the Future program managed by the Agence Nationale de la Recherche (ANR, reference 10-LABX-69-01). We thank Stéphanie Thomas for management of the Milieu Intérieur Consortium, and Haiyin Chen, Sebastian Amigorena, and Shannon Turley for critical review of the manuscript.

Received: January 19, 2016
Revised: May 31, 2016
Accepted: August 2, 2016
Published: August 25, 2016

REFERENCES

Alter, O., Brown, P.O., and Botstein, D. (2000). Singular value decomposition for genome-wide expression data processing and modeling. Proc. Natl. Acad. Sci. USA *97*, 10101–10106.

Amit, I., Garber, M., Chevrier, N., Leite, A.P., Donner, Y., Eisenhaure, T., Guttman, M., Grenier, J.K., Li, W., Zuk, O., et al. (2009). Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. Science *326*, 257–263.

Banchereau, R., Baldwin, N., Cepika, A.-M., Athale, S., Xue, Y., Yu, C.I., Metang, P., Cheruku, A., Berthier, I., Gayet, I., et al. (2014). Transcriptional specialization of human dendritic cell subsets in response to microbial vaccines. Nat. Commun. *5*, 5283.

Barreiro, L.B., Ben-Ali, M., Quach, H., Laval, G., Patin, E., Pickrell, J.K., Bouchier, C., Tichit, M., Neyrolles, O., Gicquel, B., et al. (2009). Evolutionary dynamics of human Toll-like receptors and their different contributions to host defense. PLoS Genet. *5*, e1000562.

Becattini, S., Latorre, D., Mele, F., Foglierini, M., De Gregorio, C., Cassotta, A., Fernandez, B., Kelderman, S., Schumacher, T.N., Corti, D., et al. (2015). T cell immunity. Functional heterogeneity of human memory $CD4^+$ T cell clones primed by pathogens or vaccines. Science *347*, 400–406.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J.R. Stat. Soc. Series B Stat. Methodol. *57*, 289–300.

Bradski, G., and Kaehler, A. (2008). Learning OpenCV: Computer Vision with the OpenCV Library (O'Reilly).

Burges, C.J.C. (1998). A tutorial on support vector machines for pattern recognition. Data Min. Knowl. Discov. *2*, 121–167.

Chang, C.-C., and Lin, C.-J. (2011). LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. TIST *2*.

Chaussabel, D., Pascual, V., and Banchereau, J. (2010). Assessing the human immune system through blood transcriptomics. BMC Biol. *8*, 84.

Chen, K.W., Groß, C.J., Sotomayor, F.V., Stacey, K.J., Tschopp, J., Sweet, M.J., and Schroder, K. (2014). The neutrophil NLRC4 inflammasome selectively promotes IL-1β maturation without pyroptosis during acute Salmonella challenge. Cell Rep. *8*, 570–582.

Dinarello, C.A., Ikejima, T., Warner, S.J., Orencole, S.F., Lonnemann, G., Cannon, J.G., and Libby, P. (1987). Interleukin 1 induces interleukin 1. I. Induction of circulating interleukin 1 in rabbits in vivo and in human mononuclear cells in vitro. J. Immunol. *139*, 1902–1910.

Duffy, D., Rouilly, V., Libri, V., Hasan, M., Beitz, B., David, M., Urrutia, A., Bisiaux, A., Labrie, S.T., Dubois, A., et al.; Milieu Intérieur Consortium (2014). Functional analysis via standardized whole-blood stimulation systems defines the boundaries of a healthy immune response to complex stimuli. Immunity *40*, 436–450.

Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics *21*, 3439–3440.

Flicek, P., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., et al. (2014). Ensembl 2014. Nucleic Acids Res. *42*, D749–D755.

Folds, J.D., and Schmitz, J.L. (2003). 24. Clinical and laboratory assessment of immunity. J. Allergy Clin. Immunol. *111* (*2, Suppl*), S702–S711.

Fontes, M. (2012). Statistical and knowledge supported visualization of multivariate data. In Analysis for Science, Engineering and Beyond, K. Åström, L.-E. Persson, and S.D. Silvestrov, eds. (Springer Berlin Heidelberg), pp. 143–173.

Gay, N.J., Symmons, M.F., Gangloff, M., and Bryant, C.E. (2014). Assembly and localization of Toll-like receptor signalling complexes. Nat. Rev. Immunol. *14*, 546–558.

Geiger, R., Duhen, T., Lanzavecchia, A., and Sallusto, F. (2009). Human naive and memory CD4+ T cell repertoires specific for naturally processed antigens analyzed using libraries of amplified T cells. J. Exp. Med. *206*, 1525–1534.

Gibbons, L.J., and Hyrich, K.L. (2009). Biologic therapy for rheumatoid arthritis: clinical efficacy and predictors of response. BioDrugs *23*, 111–124.

Gorenshteyn, D., Zaslavsky, E., Fribourg, M., Park, C.Y., Wong, A.K., Tadych, A., Hartmann, B.M., Albrecht, R.A., García-Sastre, A., Kleinstein, S.H., et al. (2015). Interactive Big Data Resource to Elucidate Human Immune Pathways and Diseases. Immunity *43*, 605–614.

Hawn, T.R., Verbon, A., Lettinga, K.D., Zhao, L.P., Li, S.S., Laws, R.J., Skerrett, S.J., Beutler, B., Schroeder, L., Nachman, A., et al. (2003). A common dominant TLR5 stop codon polymorphism abolishes flagellin signaling and is associated with susceptibility to legionnaires' disease. J. Exp. Med. *198*, 1563–1572.

Hayashi, F., Means, T.K., and Luster, A.D. (2003). Toll-like receptors stimulate human neutrophil function. Blood *102*, 2660–2669.

Jovanovic, M., Rooney, M.S., Mertins, P., Przybylski, D., Chevrier, N., Satija, R., Rodriguez, E.H., Fields, A.P., Schwartz, S., Raychowdhury, R., et al. (2015). Immunogenetics. Dynamic profiling of the protein life cycle in response to pathogens. Science *347*, 1259038.

Lee, M.N., Ye, C., Villani, A.-C., Raj, T., Li, W., Eisenhaure, T.M., Imboywa, S.H., Chipendo, P.I., Ran, F.A., Slowikowski, K., et al. (2014). Common genetic variants modulate pathogen-sensing responses in human dendritic cells. Science *343*, 1246980.

Li, B., Cramer, D., Wagner, S., Hansen, R., King, C., Kakar, S., Ding, C., and Yan, J. (2007). Yeast glucan particles activate murine resident macrophages to secrete proinflammatory cytokines via MyD88- and Syk kinase-dependent pathways. Clin. Immunol. *124*, 170–181.

Li, S., Rouphael, N., Duraisingham, S., Romero-Steiner, S., Presnell, S., Davis, C., Schmidt, D.S., Johnson, S.E., Milton, A., Rajam, G., et al. (2014). Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. Nat. Immunol. *15*, 195–204.

Martinon, F., Pétrilli, V., Mayor, A., Tardivel, A., and Tschopp, J. (2006). Gout-associated uric acid crystals activate the NALP3 inflammasome. Nature *440*, 237–241.

McInnes, I.B., and Schett, G. (2007). Cytokines in the pathogenesis of rheumatoid arthritis. Nat. Rev. Immunol. *7*, 429–442.

Shin, H., Shannon, C.P., Fishbane, N., Ruan, J., Zhou, M., Balshaw, R., Wilson-McManus, J.E., Ng, R.T., McManus, B.M., and Tebbutt, S.J.; PROOF Centre of Excellence Team (2014). Variation in RNA-Seq transcriptome profiles of peripheral whole blood from healthy individuals with and without globin depletion. PLoS ONE *9*, e91041.

Speake, C., Presnell, S., Domico, K., Zeitner, B., Bjork, A., Anderson, D., Mason, M.J., Whalen, E., Vargas, O., Popov, D., et al. (2015). An interactive web application for the dissemination of human systems immunology data. J. Transl. Med. *13*, 196.

Steinhaus, H. (1956). Sur la division des corps matériels en parties. Bull. Acad. Pol. Sci. Fr. *4*, 801–804.

Thomas, S., Rouilly, V., Patin, E., Alanio, C., Dubois, A., Delval, C., Marquier, L.-G., Fauchoux, N., Sayegrih, S., Vray, M., et al. (2015). The Milieu Intérieur study - an integrative approach for study of human immunological variance. Clin. Immunol. *157*, 277–293.

Tsang, J.S., Schwartzberg, P.L., Kotliarov, Y., Biancotto, A., Xie, Z., Germain, R.N., Wang, E., Olnes, M.J., Narayanan, M., Golding, H., et al.; Baylor HIPC Center; CHI Consortium (2014). Global analyses of human immune variation reveal baseline predictors of postvaccination responses. Cell *157*, 499–513.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. *3*, RESEARCH0034.

Wu, T.D., and Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics *26*, 873–881.

Cell Reports, Volume *16*

Supplemental Information

Standardized Whole-Blood Transcriptional Profiling Enables the Deconvolution of Complex Induced Immune Responses

Alejandra Urrutia, Darragh Duffy, Vincent Rouilly, Céline Posseme, Raouf Djebali, Gabriel Illanes, Valentina Libri, Benoit Albaud, David Gentien, Barbara Piasecka, Milena Hasan, Magnus Fontes, Lluis Quintana-Murci, Matthew L. Albert, and Milieu Intérieur Consortium

Supplementary Figure Legends

Table S1. Comparison of different gene expression profiling technologies for whole blood analysis.

Table S2. Gene expression probes. List of genes analyzed by Nanostring nCounter technology including chromosome number, probe map position, SNPs in probe, Ensemble gene IDs, and probe sequence

Figure S1. Quality control measures for gene expression analysis. (**A**) Schematic overview of workflow from blood draw to gene expression analysis. (**B**) Comparison between mRNA counts for single step extraction protocol and standard extraction protocol utilizing chloroform step, for nCounter analysis at 2 separate time point (75 days apart), and at 2 different locations (Institut Curie, Paris and Institut Pasteur, Paris) (Representative examples are shown and r_s^2 is reported, based on a Spearman correlation). (**C**) Mean of mRNA counts (log scale) for the 4 selected house keeping genes (HPRT1, POLR2A, RPL19, TBP) across the different stimulation conditions for the 25 donors included in the study. (**D**) Comparison of mRNA counts (linear scale) for two geNorm selected genes (left plots) versus two candidate house keeping genes (right plots) upon TNFA and SeV stimulation **(E)** Box-whisker Tukey plots for the negative control probe counts. (**F**) Example of Neg B and Neg_F probes from TruCulture stimuli LPS, HKSA, and BCG.

Figure S2. Gene expression pathways used to select NanoString Immunology panel. KEGG database pathway analysis of (A) NF-κB, (B) TNFA, (C) Cytokine-Cytokine Receptor, and (D) TLR signaling pathways, with genes included in NanoString analysis colored green, and effector cytokines (IFNB, IFNG, IL1B, TNFA) studied herein colored yellow. Genes in white were not represented on the NanoString codeset.

Figure S3. IFNA and IFNB show overlapping gene expression profiles. (**A**) Whole-blood stimulation was performed on 25 healthy donors using TruCulture systems pre-loaded with IFNA (red), IFNB (pale green), IFNG (grey), IL1B (purple), and TNFA (turquoise). Principle component analysis (PCA) was used to project mRNA expression data from 572 genes (the percentage of variance captured by each PC is indicated). (**B-C**) Hierarchical cluster analysis of the donors and gene expression following stimulation with IFNA, IFNB, and NULL control (black) identified 58 genes commonly down regulated (**B**) and 212 genes commonly

upregulated (**C**) (ANOVA test, q value < 10^{-3}). Each donor is color-coded revealing that in most instances, individual donors clustered for IFNA / IFNB responses.

Figure S4. Projection of TLR stimuli onto PCA analysis as defined by 4 effector cytokines.

(**A**) PCA defined by the eigenvectors and eigenvalues as based on the four-cytokine induced mRNA expression data of the 44 genes defined in Table 1. Ellipses representing 95% confidence interval (CI) were constructed and replaced the individual samples. Projected sample vectors of TLR stimuli (shown in red) for each of the 25 donors (FSL, pIC, LPS, FLA, GARD, R848, ODN), individually projected onto the first 3 PC vectors, using the 44 selected genes (**B**) Projection of different synthetic ligands (WGP, LAM, CPPD) onto the PCA as defined by four-cytokine induced mRNA expression.

Figure S5. Projection of microbial stimuli onto fixed PCA analysis defined by 4 effector cytokines.

PCA defined by the eigenvectors and eigenvalues and optimized based on the four-cytokine induced mRNA expression data (44 genes defined in Table 1). Ellipses representing 95% confidence interval (CI) were constructed and replaced the individual samples. Projected sample vectors (shown in red) for microbial stimuli for each of the 25 donors (HKHP, HKSA, HKLR, HKEC, BCG, HKCA), individually, projected onto the first 3 PC vectors, using the 44 selected genes.

Figure S6. Gene expression patterns not captured by four effector cytokine induced changes. Hierarchical cluster analysis of donors and gene expression based on genes expressed after microbial stimulation but not cytokine stimulation (**A**), and TLR but not cytokine stimulation (**B**) as defined by firstly by a paired T test for all stimuli versus null (q < 10^{-3}) and merging the gene lists.

Table S1

Technology	Method	Gene set	Adaptable for use with whole blood
Fluidigm Biomark	Microfluidic-based PCR	Up to 90 genes	• cDNA preparation source of possible pre-analytical bias • PCR reaction inhibited by presence of heparin (and other serum factors) in TruCulture systems
Nanostring nCounter	RNA Hybridization	Up to 780 genes	• **Direct quantification of mRNA expression (i.e., no cDNA preparation, no amplification)** • **suitable for single-step extracted RNA**
Affymetrix microarray	cDNA Hybridization	**Whole genome**	• requirement for high quality RNA • cDNA preparation source of possible pre-analytical bias
Ion Torrent / Illumina	NGS	**Whole genome**	• requirement for high quality RNA • globin RNA present in whole blood dominate read count (~70% of total RNA)

Figure S1

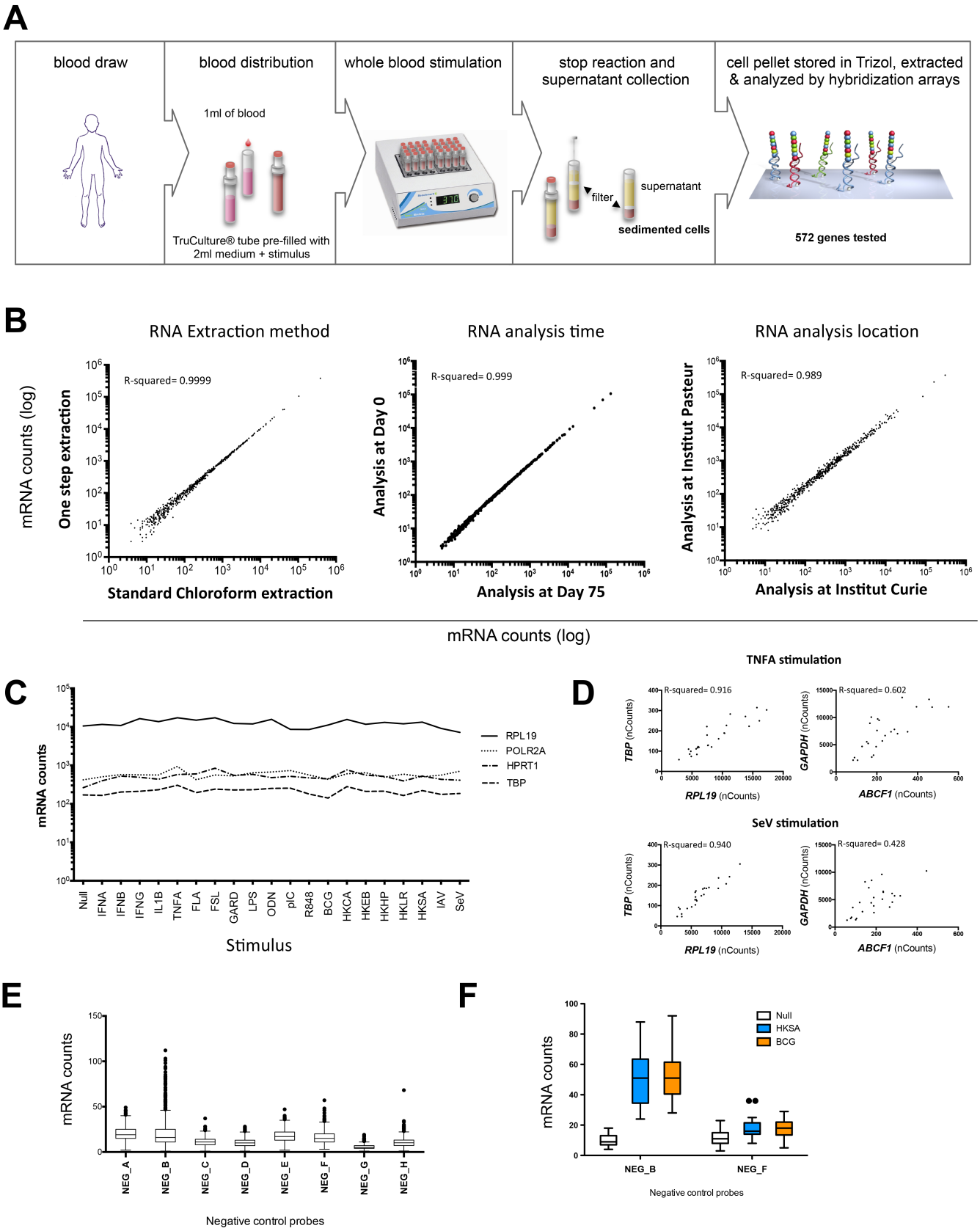
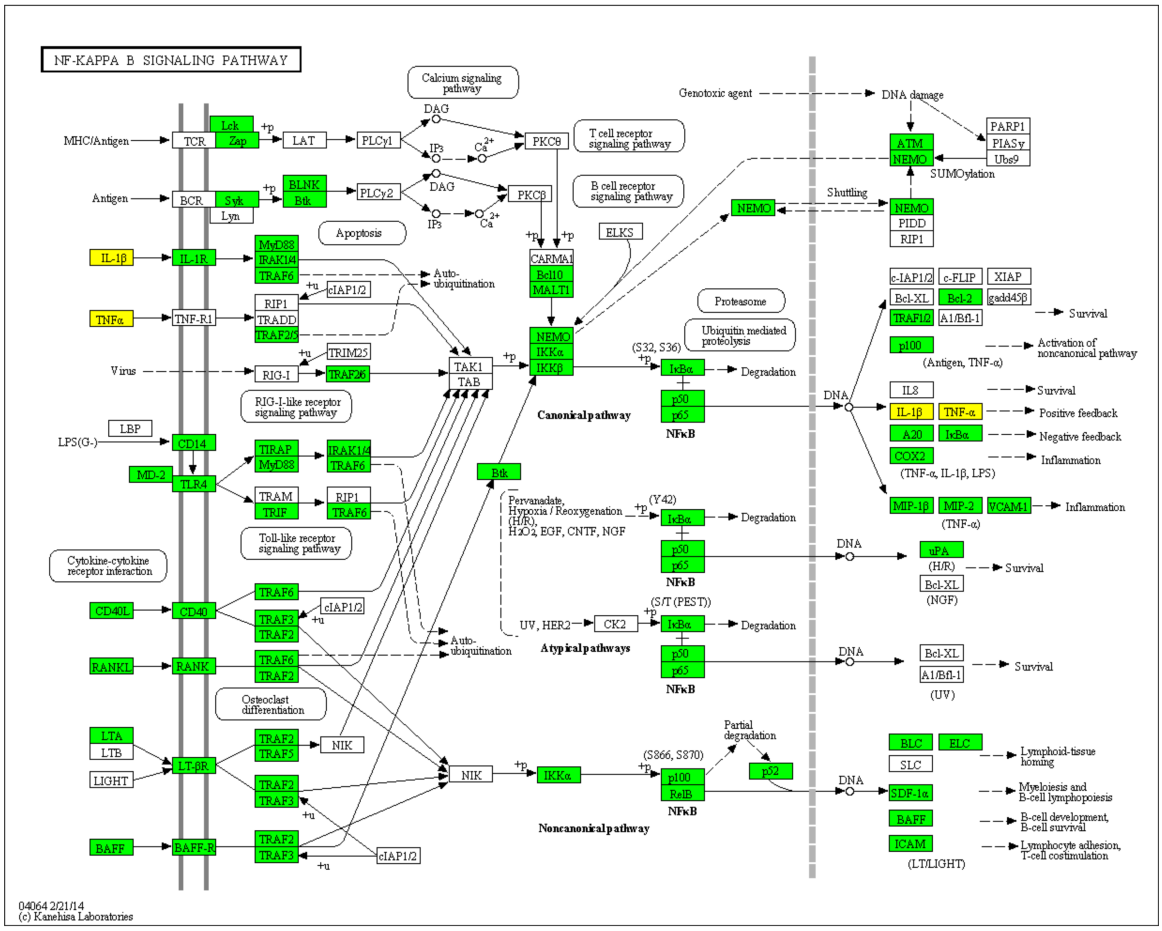

A
blood draw
blood distribution
1ml of blood
TruCulture® tube pre-filled with 2ml medium + stimulus
whole blood stimulation
stop reaction and supernatant collection
filter
supernatant
sedimented cells
cell pellet stored in Trizol, extracted & analyzed by hybridization arrays
572 genes tested
B
RNA Extraction method
RNA analysis time
RNA analysis location
mRNA counts (log)
One step extraction
Standard Chloroform extraction
R-squared= 0.9999
Analysis at Day 0
Analysis at Day 75
R-squared= 0.999
Analysis at Institut Pasteur
Analysis at Institut Curie
R-squared= 0.989
C
mRNA counts
RPL19
POLR2A
HPRT1
TBP
Null IFNA IFNB IFNG IL1B TNFA FLA FSL GARD LPS ODN pIC R848 BCG HKCA HKEB HKHP HKLR HKSA IAV SeV
Stimulus
D
TNFA stimulation
R-squared= 0.916
R-squared= 0.602
TBP (nCounts)
GAPDH (nCounts)
RPL19 (nCounts)
ABCF1 (nCounts)
SeV stimulation
R-squared= 0.940
R-squared= 0.428
E
mRNA counts
NEG_A NEG_B NEG_C NEG_D NEG_E NEG_F NEG_G NEG_H
Negative control probes
F
mRNA counts
Null
HKSA
BCG
NEG_B
NEG_F
Negative control probes

Figure S2

A


NF-KAPPA B SIGNALING PATHWAY
Calcium signaling pathway
T cell receptor signaling pathway
B cell receptor signaling pathway
Apoptosis
RIG-I-like receptor signaling pathway
Toll-like receptor signaling pathway
Cytokine-cytokine receptor interaction
Osteoclast differentiation
Proteasome
Ubiquitin mediated proteolysis
Canonical pathway
Atypical pathways
Noncanonical pathway
NFkB
Genotoxic agent
DNA damage
SUMOylation
Shuttling
Survival
Activation of noncanonical pathway
Positive feedback
Negative feedback
Inflammation
Degradation
Lymphoid-tissue homing
Myeloiesis and B-cell lymphopoiesis
B-cell development, B-cell survival
Lymphocyte adhesion, T-cell costimulation
04064 2/21/14
(c) Kanehisa Laboratories

B

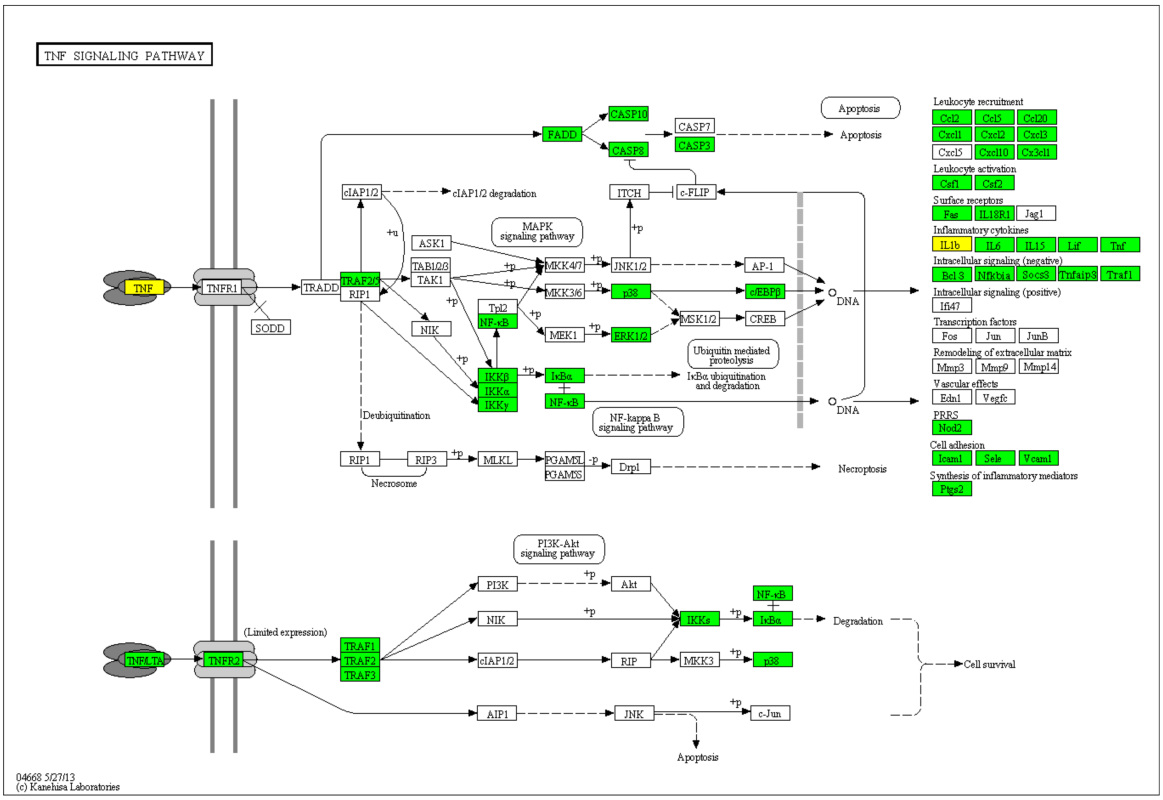
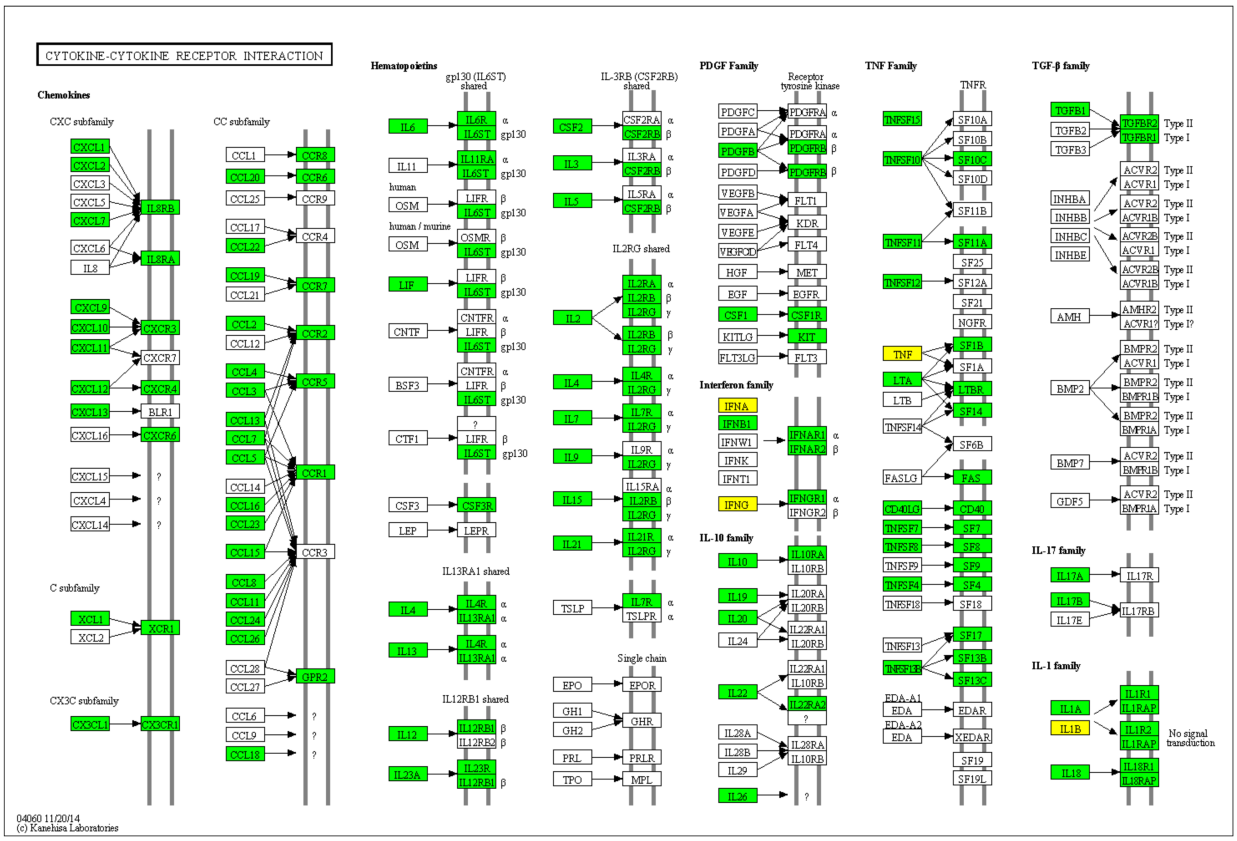

TNF SIGNALING PATHWAY
Apoptosis
MAPK signaling pathway
NF-kappa B signaling pathway
Ubiquitin mediated proteolysis
Necrosome
Necroptosis
PI3K-Akt signaling pathway
Leukocyte recruitment
Leukocyte activation
Surface receptors
Inflammatory cytokines
Intracellular signaling (negative)
Intracellular signaling (positive)
Transcription factors
Remodeling of extracellular matrix
Vascular effects
PRRS
Cell adhesion
Synthesis of inflammatory mediators
04668 5/27/13
(c) Kanehisa Laboratories

Figure S2

C


CYTOKINE-CYTOKINE RECEPTOR INTERACTION
Chemokines
CXC subfamily
CC subfamily
C subfamily
CX3C subfamily
Hematopoietins
gp130 (IL6ST) shared
IL-3RB (CSF2RB) shared
IL2RG shared
IL13RA1 shared
IL12RB1 shared
Single chain
PDGF Family
Receptor tyrosine kinase
Interferon family
IL-10 family
TNF Family
TNFR
TGF-β family
IL-17 family
IL-1 family
04060 11/20/14
(c) Kanehisa Laboratories

D

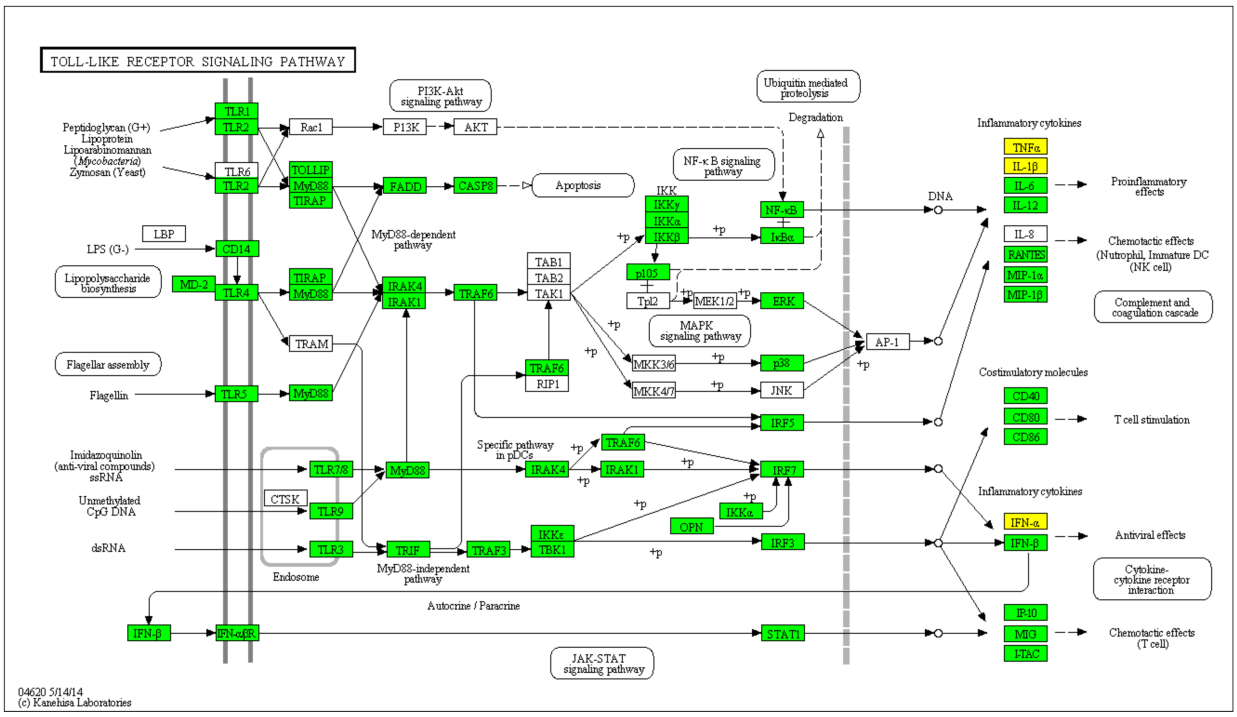

TOLL-LIKE RECEPTOR SIGNALING PATHWAY
PI3K-Akt signaling pathway
Ubiquitin mediated proteolysis
Degradation
Peptidoglycan (G+)
Lipoprotein
Lipoarabinomannan (Mycobacteria)
Zymosan (Yeast)
LPS (G-)
LBP
Lipopolysaccharide biosynthesis
Flagellar assembly
Flagellin
Imidazoquinolin (anti-viral compounds) ssRNA
Unmethylated CpG DNA
dsRNA
Endosome
MyD88-dependent pathway
MyD88-independent pathway
Specific pathway in pDCs
NF-κB signaling pathway
MAPK signaling pathway
JAK-STAT signaling pathway
Autocrine / Paracrine
Apoptosis
DNA
Inflammatory cytokines
Proinflammatory effects
Chemotactic effects (Nutrophil, Immature DC (NK cell)
Complement and coagulation cascade
Costimulatory molecules
T cell stimulation
Antiviral effects
Cytokine-cytokine receptor interaction
Chemotactic effects (T cell)
04620 5/14/14
(c) Kanehisa Laboratories

Figure S3

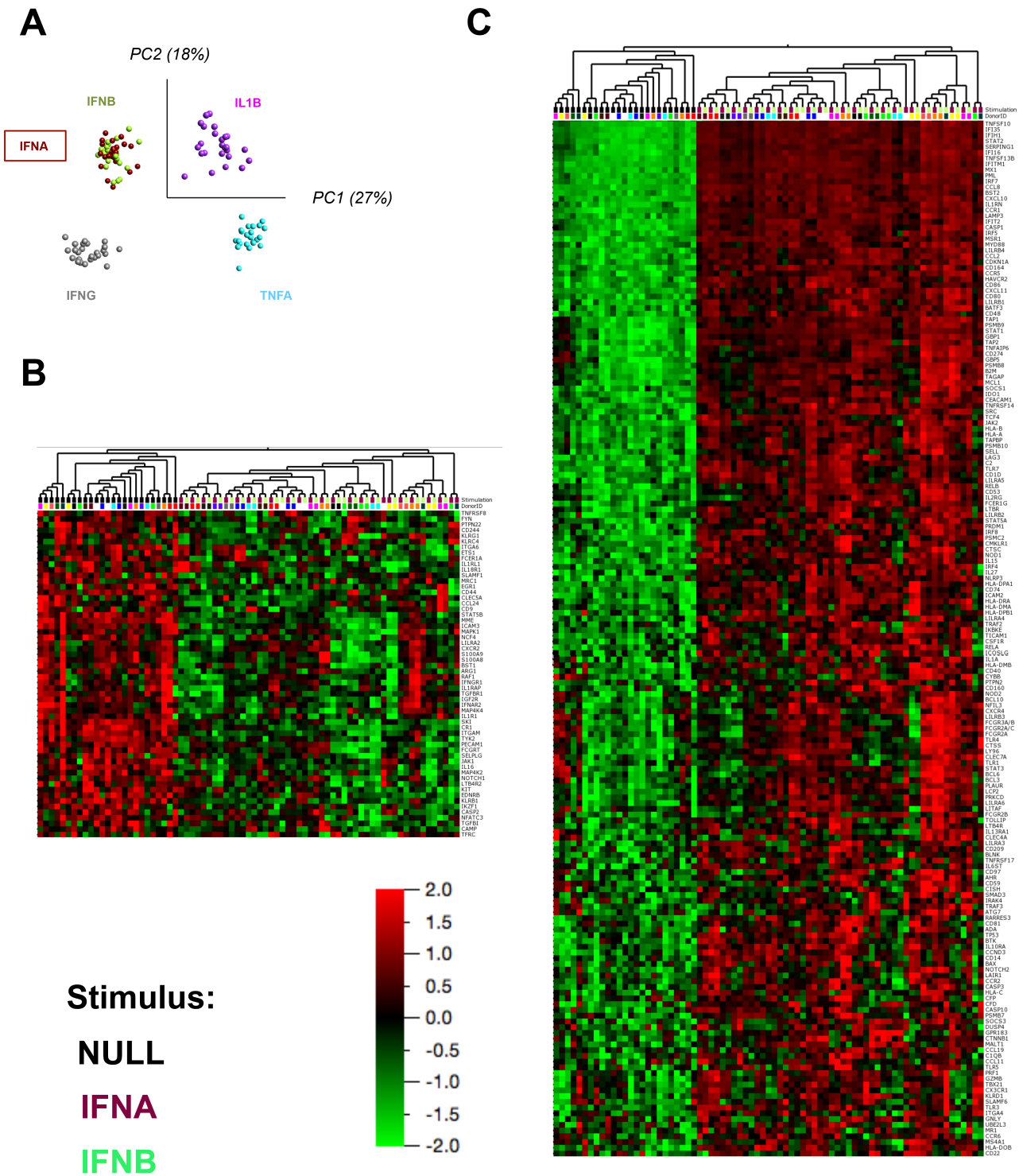

A
PC2 (18%)
IFNB
IL1B
IFNA
PC1 (27%)
IFNG
TNFA
B
C
Stimulation
DonorID
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
Stimulus:
NULL
IFNA
IFNB

Figure S4

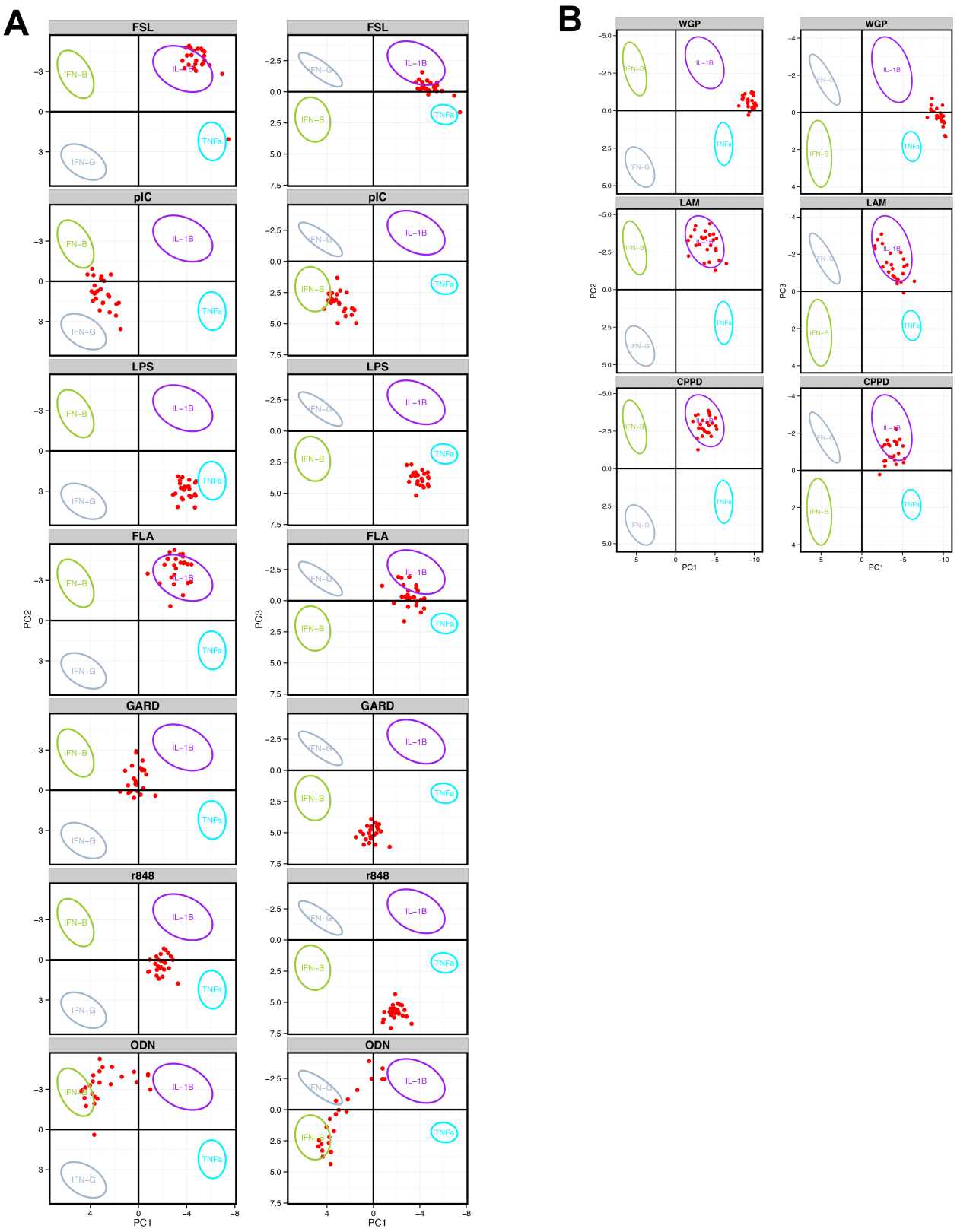

A
FSL
pIC
LPS
FLA
GARD
r848
ODN
IFN−B
IFN−G
IL−1B
TNFa
PC1
PC2
PC3
B
WGP
LAM
CPPD
PC1
PC2
PC3

Figure S5

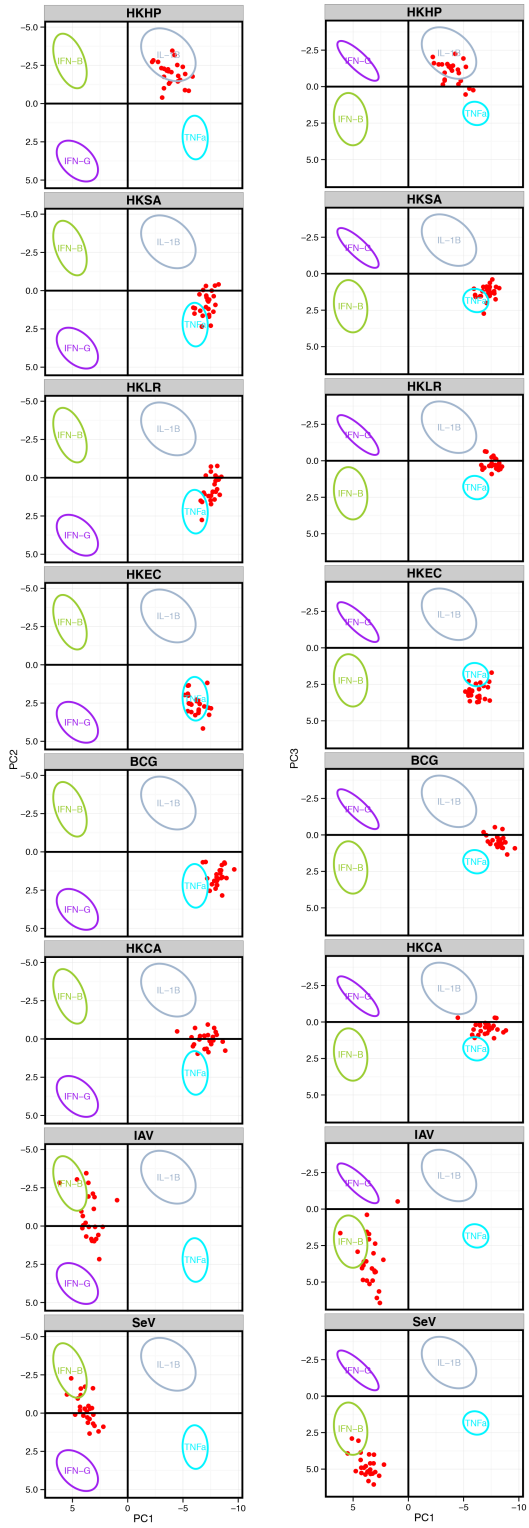
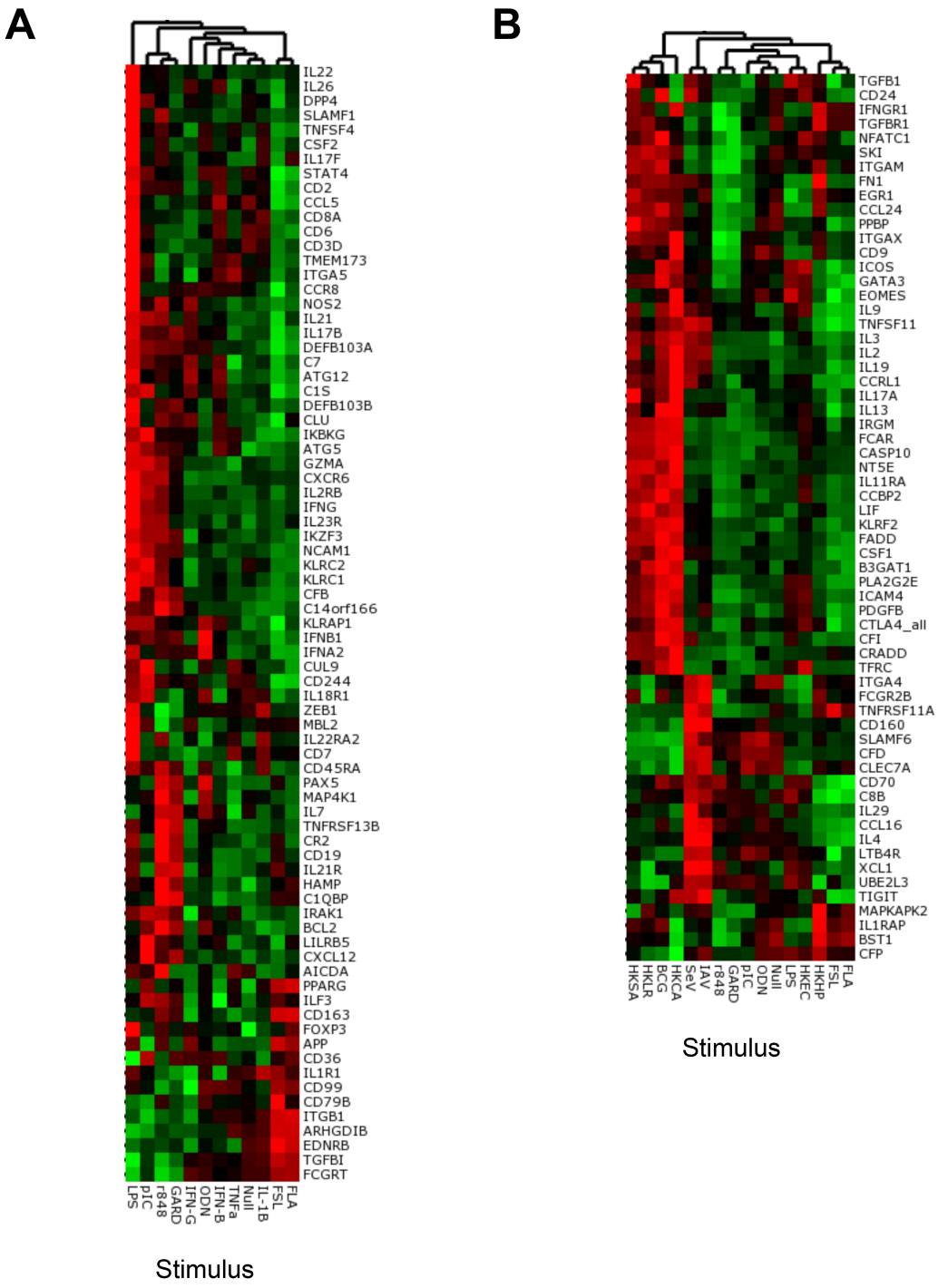

HKHP
HKSA
HKLR
HKEC
BCG
HKCA
IAV
SeV
IFN-B
IFN-G
IL-1B
TNFa
PC2
PC3
PC1

Figure S6


A
IL22
IL26
DPP4
SLAMF1
TNFSF4
CSF2
IL17F
STAT4
CD2
CCL5
CD8A
CD6
CD3D
TMEM173
ITGA5
CCR8
NOS2
IL21
IL17B
DEFB103A
C7
ATG12
C1S
DEFB103B
CLU
IKBKG
ATG5
GZMA
CXCR6
IL2RB
IFNG
IL23R
IKZF3
NCAM1
KLRC2
KLRC1
CFB
C14orf166
KLRAP1
IFNB1
IFNA2
CUL9
CD244
IL18R1
ZEB1
MBL2
IL22RA2
CD7
CD45RA
PAX5
MAP4K1
IL7
TNFRSF13B
CR2
CD19
IL21R
HAMP
C1QBP
IRAK1
BCL2
LILRB5
CXCL12
AICDA
PPARG
ILF3
CD163
FOXP3
APP
CD36
IL1R1
CD99
CD79B
ITGB1
ARHGDIB
EDNRB
TGFBI
FCGRT
LPS
pIC
r848
GARD
IFN-G
ODN
IFN-B
TNFa
Null
IL-1B
FSL
FLA
Stimulus
B
TGFB1
CD24
IFNGR1
TGFBR1
NFATC1
SKI
ITGAM
FN1
EGR1
CCL24
PPBP
ITGAX
CD9
ICOS
GATA3
EOMES
IL9
TNFSF11
IL3
IL2
IL19
CCRL1
IL17A
IL13
IRGM
FCAR
CASP10
NT5E
IL11RA
CCBP2
LIF
KLRF2
FADD
CSF1
B3GAT1
PLA2G2E
ICAM4
PDGFB
CTLA4_all
CFI
CRADD
TFRC
ITGA4
FCGR2B
TNFRSF11A
CD160
SLAMF6
CFD
CLEC7A
CD70
C8B
IL29
CCL16
IL4
LTB4R
XCL1
UBE2L3
TIGIT
MAPKAPK2
IL1RAP
BST1
CFP
HKSA
HKLR
BCG
HKCA
SeV
IAV
r848
GARD
pIC
ODN
Null
LPS
HKEC
HKHP
FSL
FLA
Stimulus