

University of Groningen

A Functional Genomics Approach to Understand Variation in Cytokine Production in Humans

Li, Yang; Oosting, Marije; Smeekens, Sanne P.; Jaeger, Martin; Aguirre-Gamboa, Raul; Le, Kieu T. T.; Deelen, Patrick; Ricano-Ponce, Isis; Schoffelen, Teske; Jansen, Anne F. M.

Published in:
Cell

DOI:
[10.1016/j.cell.2016.10.017](https://doi.org/10.1016/j.cell.2016.10.017)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2016

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Li, Y., Oosting, M., Smeekens, S. P., Jaeger, M., Aguirre-Gamboa, R., Le, K. T. T., ... Netea, M. G. (2016). A Functional Genomics Approach to Understand Variation in Cytokine Production in Humans. *Cell*, 167(4), 1099-1110. <https://doi.org/10.1016/j.cell.2016.10.017>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

A Functional Genomics Approach to Understand Variation in Cytokine Production in Humans

Yang Li,^{1,8,9,*} Marije Oosting,^{2,8} Sanne P. Smeekens,^{2,8} Martin Jaeger,^{2,8} Raul Aguirre-Gamboa,¹ Kieu T.T. Le,¹ Patrick Deelen,^{1,3} Isis Ricaño-Ponce,¹ Teske Schoffelen,² Anne F.M. Jansen,² Morris A. Swertz,^{1,3} Sebo Withoff,¹ Esther van de Vosse,⁴ Marcel van Deuren,² Frank van de Veerdonk,² Alexandra Zhernakova,¹ Jos W.M. van der Meer,² Ramnik J. Xavier,^{5,6} Lude Franke,¹ Leo A.B. Joosten,^{2,*} Cisca Wijmenga,^{1,7,*} Vinod Kumar,^{1,*} and Mihai G. Netea^{2,*}

¹Department of Genetics, University Medical Center Groningen, University of Groningen, 9700 RB Groningen, the Netherlands

²Department of Internal Medicine and Radboud Center for Infectious Diseases, Radboud University Medical Center, 6525 HP Nijmegen, the Netherlands

³Genomics Coordination Center, University Medical Center Groningen, University of Groningen, 9700 RB Groningen, the Netherlands

⁴Department of Infectious Diseases, Leiden University Medical Center, 2333 ZA Leiden, the Netherlands

⁵Center for Computational and Integrative Biology and Gastrointestinal Unit, Massachusetts General Hospital, Harvard School of Medicine, Boston, MA 02114, USA

⁶Broad Institute of MIT and Harvard University, Cambridge, MA 02142, USA

⁷Department of Immunology, University of Oslo, Oslo University Hospital, Rikshospitalet, 0372 Oslo, Norway

⁸Co-first author

⁹Lead Contact

*Correspondence: y.li01@umcg.nl (Y.L.), leo.joosten@radboudumc.nl (L.A.B.J.), c.wijmenga@umcg.nl (C.W.), v.kumar@umcg.nl (V.K.), mihai.netea@radboudumc.nl (M.G.N.)

<http://dx.doi.org/10.1016/j.cell.2016.10.017>

SUMMARY

As part of the Human Functional Genomics Project, which aims to understand the factors that determine the variability of immune responses, we investigated genetic variants affecting cytokine production in response to *ex vivo* stimulation in two independent cohorts of 500 and 200 healthy individuals. We demonstrate a strong impact of genetic heritability on cytokine production capacity after challenge with bacterial, fungal, viral, and non-microbial stimuli. In addition to 17 novel genome-wide significant cytokine QTLs (cQTLs), our study provides a comprehensive picture of the genetic variants that influence six different cytokines in whole blood, blood mononuclear cells, and macrophages. Important biological pathways that contain cytokine QTLs map to pattern recognition receptors (*TLR1-6-10* cluster), cytokine and complement inhibitors, and the kallikrein system. The cytokine QTLs show enrichment for monocyte-specific enhancers, are more often located in regions under positive selection, and are significantly enriched among SNPs associated with infections and immune-mediated diseases.

INTRODUCTION

The Human Functional Genomics Project (HFGP) is an initiative that aims to identify the factors responsible for the variability of immune responses in health and disease (<http://www.humanfunctionalgenomics.org>).

Within the HFGP, the 500-Human Functional Genomics (500FG) cohort focuses on gaining a broader understanding of the variability in human cytokine responses. In a first study reported in this issue of *Cell*, we investigated the role of environmental and non-genetic host factors for cytokine responses (ter Horst et al., 2016). In the present study, we investigate the role of genetic variation for individuals human cytokine responses, while a third complementary study assessed the impact of microbiome factors (Schirmer et al., 2016).

Many targeted candidate gene studies have demonstrated the impact of specific genetic variants on immune responses, while a recent study that assessed the genetics of lipopolysaccharide (LPS)-induced cytokine responses by dendritic cells identified several candidate genes (Lee et al., 2014). Furthermore, genome-wide genetic studies have found genetic variants that impact transcript abundance for immune genes (so-called eQTLs [expression quantitative trait loci]) (Kumar et al., 2014a; Fairfax et al., 2014; Lee et al., 2014), while genome-wide association studies (GWASs) have identified hundreds of genetic variants predisposing to the susceptibility to immune-mediated diseases and/or their severity (Welter et al., 2014). However, there have been no comprehensive genome-wide association studies to investigate variation in cytokine production in humans so far. As a proof of concept, we assessed the genetics of three monocyte-derived cytokines (tumor necrosis factor α [TNF- α], interleukin [IL]-1 β , and IL-6) after stimulation with a few microbial stimuli: we identified four genome-wide loci that influence cytokine release (Li et al., 2016). This clearly demonstrated the importance of genetic variation for cytokine production in humans, and we decided to pursue a more comprehensive approach to reveal the most important genetic factors that influence cytokine responses.



Figure 1. Study Overview

We collected blood samples from 500 healthy individuals in the 500FG cohort and isolated their DNA. This was hybridized on the HumanCoreExome SNP Chip to provide genotype information on approximately 8 million SNPs. The blood was also used to perform a series of stimulation experiments with major human pathogens and to profile the cytokines released in the serum (see STAR Methods). See also Figures S1 and S2.

Here, we describe the stimulation of three different cellular systems (whole blood, peripheral blood mononuclear cells [PBMCs], and macrophages) with a broad panel of bacterial, fungal, viral, and non-microbial stimuli to induce cytokine production, which was analyzed with approximately 8.0 million genetic variants (SNPs). The discovery was performed in the 500FG cohort, and validation was performed in the 200FG cohort. We were able to validate 17 new genome-wide significant loci that represent cytokine QTLs (cQTLs), and we describe new pathways for the modulation of cytokine responses in humans.

RESULTS

Overview of Cytokine Response Architecture

We assessed cytokine production capacity in the 500FG discovery cohort in three cellular systems: whole-blood stimulations, PBMC stimulations, and stimulation of monocyte-derived macrophages. We used a comprehensive range of seven bacterial, three fungal, one viral, four Toll-like receptor (TLR) ligands, and

two non-microbial metabolic stimuli to assess three monocyte-derived and three lymphocyte-derived cytokines (see Figure 1 for overview).

Significant increases in the levels of all cytokines were observed in all stimulation systems compared to steady-state levels (see also the related paper by Schirmer et al., 2016, for the main bacterial and fungal stimuli, as well as Figure S1 for a full description of all stimuli). Cytokines IL-6 and TNF- α from whole blood and PBMCs showed higher inter-individual variation than production by macrophages (Figure S2), suggesting that the in vitro differentiation of macrophages is a process that partly overrides individual variation. In general, IL-6 showed a much stronger inter-individual variation than any other cytokines ($p < 0.001$), suggesting a much stronger impact of cell types and/or genetic variation on IL-6 production than other cytokines. These results were consistent with those we obtained from the 200FG cohort (data not shown).

Unsupervised clustering of the cytokine responses showed a clear distinction between stimulations with bacteria, fungi, or

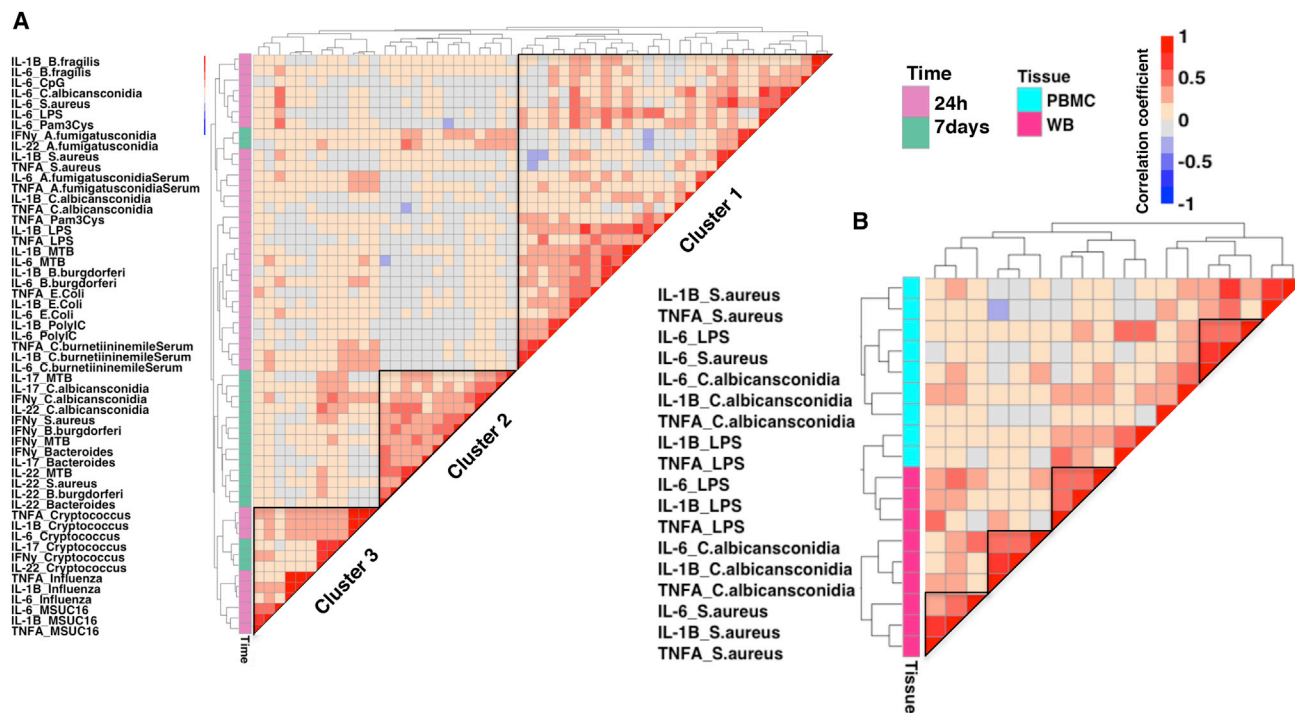


Figure 2. The Cytokine Responses Are Organized around the Physiological Response toward Specific Pathogens

(A) The results from unsupervised hierarchical clustering of the cytokine responses in PBMCs induced by various pathogens and microbial ligands are shown. Clustering was performed using Spearman's correlation as the measure of similarity. Red indicates a strong positive correlation, whereas blue indicates a strong negative correlation. Cluster 1 depicts the positive correlation between monocyte-induced cytokines (IL-6, IL-1 β , and TNF- α) on stimulation of PBMCs for 24 hr. Cluster 2 depicts the positive correlation among cytokines derived from T-helper cells (IL-17, IL-22, and IFN- γ) on stimulation of PBMCs for 7 days. Cluster 3 depicts the strong correlation between influenza- and *Cryptococcus*-induced cytokines for both T cell- and monocyte-derived cytokines.

(B) The results from unsupervised hierarchical clustering of the cytokine responses in blood were compared with responses in PBMCs. The stimulation-cytokine pairs that were available for both cell systems were chosen to perform unsupervised hierarchical clustering. Four different clusters indicate the pathogen-specific clustering of cytokines. WB, whole blood.

viruses (Figure 2A). Correlations between the productions of various cytokines were found mainly for stimulation with a certain microbe rather than between cytokine productions induced by different microbes, which suggests that immune responses are organized to respond to a specific pathogen rather than through a specific immune pathway. The clustering also revealed a poor correlation between monocyte-derived- and T-helper-derived cytokine responses (Figure 2A). This is surprising, as the differentiation of naive T cells into Th1- or Th17-effector lymphocytes is controlled by monocyte-derived cytokines. However, this conclusion is also supported by our clustering analyses of whole-blood stimulations (Figure 2B). An exception to these patterns was the fungal *Cryptococcus*-induced cytokine responses, in which the distinction between monocyte-derived- and T cell-derived cytokines was weak. In addition, the *Cryptococcus*-induced cytokines were more similar to cytokine responses induced by influenza virus than to other fungi (Figure 2A).

To assess whether cell-based factors are the only factor determining variation in cytokine responses, or whether plasma-derived factors can qualitatively modulate the responses, we correlated specific responses in purified PBMCs versus whole-blood stimulations (Figure 2B). Unsupervised clustering demonstrated stronger correlations of responses in the two stimulation

systems, but we also found positive correlations between them (Figure 2B). These findings suggest that, although intrinsic factors in the mononuclear cells mainly determine the cytokine response, additional variation in cytokine production may also be induced by other whole-blood components, such as neutrophils or plasma factors.

Contribution of Genetic Variation to How Cytokines Respond to Pathogens

We observed that cytokines show higher inter-individual variation upon stimulation (Figure S2). Since a difference in cell-count proportions can be an important factor influencing the amount of cytokines produced, we tested whether cell-count differences determine inter-individual variation in cytokine levels. For this, we obtained immune-cell-count data measured by fluorescence-activated cell sorting (FACS) for total lymphocytes, T cells, B cells, monocytes, and natural killer (NK) cells from all 500FG individuals (Aguirre-Gamboa et al., 2016). We observed weak correlations between cell counts and cytokine levels (Figure S3A), suggesting a minor effect of cell-count differences on cytokine production capacity. We then estimated the proportion of cytokine variance explained by genome-wide SNPs for all cytokine measurements before and after correcting

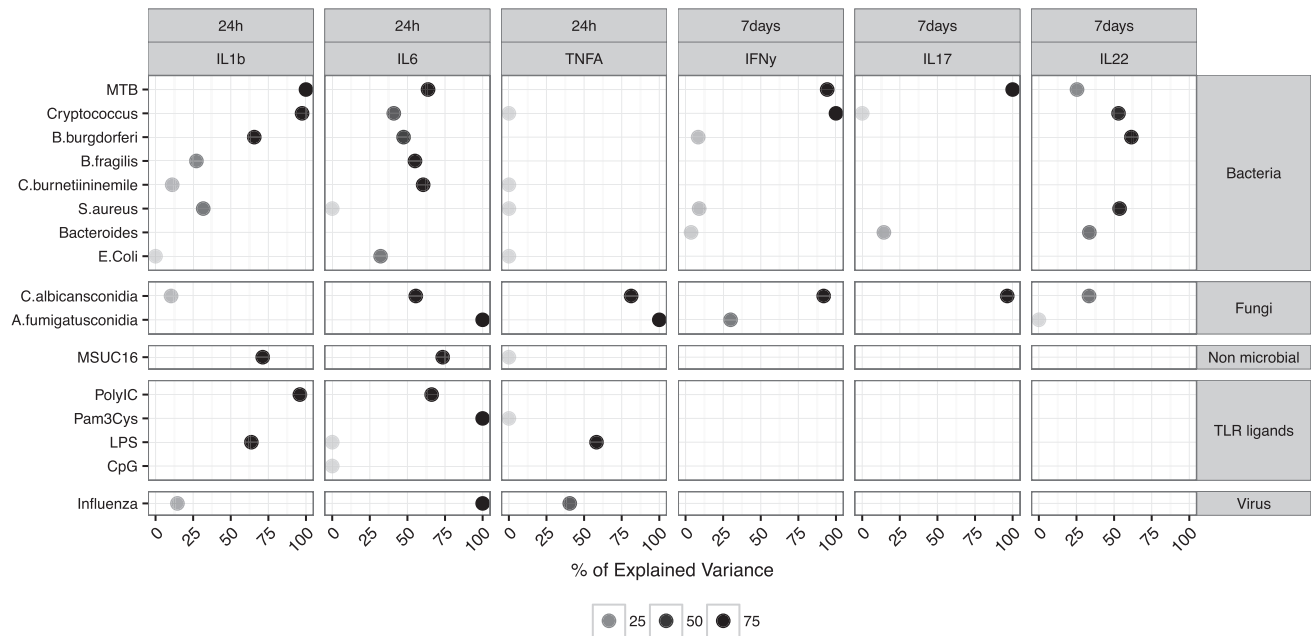


Figure 3. Proportion of the Estimated Cytokine Variance Explained by Genetic Factors

A summary of all the estimates of cytokine variance explained by genome-wide SNP data after age, gender, and cell-count correction is shown. The estimates <25% are shown in gray, and the estimates >50% are shown in black.

See also [Figures S6 and S7](#) and [Table S1](#).

for age, gender, and cell counts ([Figure 3](#); [Figure S3B](#); [Table S1](#)) using the GREML method ([Yang et al., 2010](#)).

In total, for around 70% of all the cytokine responses in PBMCs, the genetic influence was considerably larger than previously reported (>25% of explained variance) ([Brodin et al., 2015](#)). We found similar results when we estimated heritability without correcting age, gender, and cell counts ([Table S1](#)). In general, we found a higher explained variance for monocyte-derived cytokines from genetic factors (>50% of explained variance especially for IL-6 and IL1- β) than for T cell-derived cytokines ([Figure 3](#)). Finding the strongest inter-individual variation in IL-6 levels upon stimulation, in addition to the highest heritability for IL-6 levels, indicates that there may be many genome-wide significant QTLs for IL-6 in the context of infectious pressure. In T cell-derived cytokines, we found a higher explained variance for IL-17 from genetic factors. Although it may be expected that the cytokine production capacity is affected by genetic factors, we observed that the estimated explained variance due to genetic factors differed for the stimulation by the various microorganisms and for the individual cytokines studied. This finding indicates there are genetic variations that may be strongly regulating the cytokine production in response to certain pathogens.

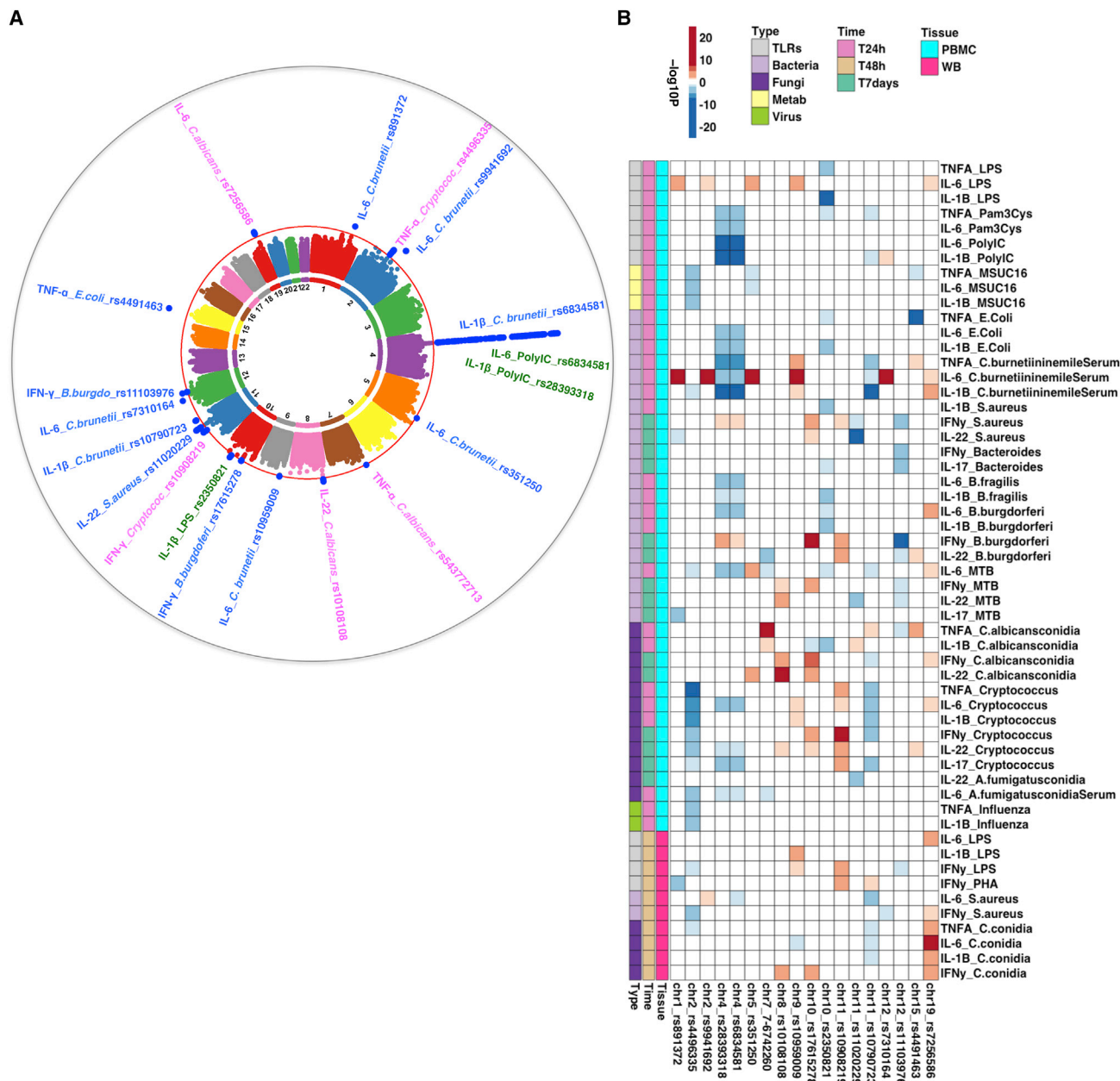
Identifying Genome-wide Genetic Variations Affecting Cytokine Production in Response to Pathogens

To identify the most significant genetic loci that determine cytokine levels upon stimulation, we mapped cQTLs using genome-wide SNP genotypes. After correcting for age, gender, and cell counts, we identified 18 genome-wide ($p < 5 \times 10^{-8}$) significant lead SNPs in 17 independent loci ([Figures 4A and 4B](#)). These

include seven independent QTLs for IL-6, three independent QTLs for IL-1 β , and three independent QTL for TNF- α levels ([Table 1](#)). Of the 17 loci, all but one were identified for cytokines measured after PBMC stimulations, while one locus on chromosome 19 came from the whole-blood stimulation system ([Table 1](#)). We identified cQTLs for both monocyte- and T cell-derived cytokines upon bacterial and fungal stimulations, whereas stimulation with purified TLR ligands only yielded cQTLs for monocyte-derived cytokines ([Figure S3B](#)). The validity of the 17 loci was further corroborated for the 12 cytokine-microbial stimulations that were performed in the 200FG cohort. Of the cQTLs, 9/12 (75%) were replicated ($p < 0.05$), and, in all cases, the effects were in the same direction ([Table S2](#)). We could not replicate five of the cQTLs, as these stimulus-cytokine measurements were not tested in the 200FG cohort. No genome-wide significant cQTLs were identified for the macrophage production of cytokines, which agrees with the conclusion that the process of in vitro differentiation erases many of the differences between individuals. This conclusion has important consequences, as it suggests that in-vitro-differentiated cell systems (such as monocyte-derived macrophages or dendritic cells) may not be suitable for studying the genetics of cytokine responses.

Prioritizing cQTL-Affected Genes Indicates Those Involved in Microbial Sensing and Processing Molecules as Putative Causal Genes

We used three approaches to identify the causal genes at the 17 significant cQTL loci. First, we tested whether the cQTLs were strongly correlated with other SNPs that alter the protein structure of any genes. Using the HaploReg SNP annotation tool



(Ward and Kellis, 2012), we extracted all SNPs in linkage disequilibrium (LD) ($R^2 > 0.8$; using the CEU population as a reference) with the cQTLs. We found two loci that were in strong LD with missense variants (Table S3): SNPs rs28393318 and rs6834581 on chromosome 4 were in strong LD ($R^2 = 0.97$,

D prime = 0.99) with a missense variant, rs4833095, on the *TLR1* gene. SNP rs7256586 on chromosome 19 was in strong LD with a missense variant rs198977 ($R^2 = 0.82$, D prime = 0.92) on the *KLK2* gene. These observations suggest that *TLR1* and *KLK2* could be causal genes at these cQTL loci. The

Table 1. Genome-Wide Significant Cytokine QTL Loci

Locus	SNPs	Chromosome	Base Pair	Stimulation	Cytokine	Cell System	Time	p Value ^a	Causal Genes
1	rs891372	1	207414046	<i>C. burnetii</i>	IL-6	PBMC	24 hr	3.2×10^{-9}	<i>CD55</i> ^{b,c} and <i>CR2</i> ^{b,c}
2	rs4496335	2	113844475	<i>Cryptococcus</i>	TNF- α	PBMC	24 hr	4.2×10^{-9}	<i>IL1RN</i> ^b
3	rs9941692	2	153183071	<i>C. burnetii</i>	IL-6	PBMC	24 hr	3.7×10^{-10}	<i>FMNL2</i> ^b and <i>STAM2</i> ^c
4	rs28393318	4	38784267	Poly(I:C)	IL-1 β	PBMC	24 hr	2.8×10^{-11}	<i>TLR1</i> ^{b,c,d} , <i>TLR6</i> ^{b,c} , <i>TLR10</i> ^{b,c} , and <i>FAM114A1</i> ^b
	rs6834581	4	38788234	<i>C. burnetii</i>	IL-1 β	PBMC	24 hr	4.6×10^{-13}	
	rs6834581	4	38788234	Poly(I:C)	IL-6	PBMC	24 hr	3.9×10^{-25}	
5	rs351250	5	141286682	<i>C. burnetii</i>	IL-6	PBMC	24 hr	3.3×10^{-8}	<i>SLC25A2</i> ^b , <i>PCDH12</i> ^b , and <i>KIAA0141</i> ^c
6	rs543772713	7	6742260	<i>C. albicans</i>	TNF- α	PBMC	24 hr	4.2×10^{-8}	<i>ZNF12</i> ^e
7	rs10108108	8	5056750	<i>C. albicans</i>	IL-22	T cell	7 days	1.5×10^{-8}	<i>CSMD1</i> ^e
8	rs10959009	9	10278031	<i>C. burnetii</i>	IL-6	PBMC	24 hr	3.9×10^{-8}	<i>PTPRD</i> ^c
9	rs17615278	10	36600213	<i>Borrelia</i>	IFN- γ	T cell	7 days	3.9×10^{-8}	<i>RP11-92J19.3</i> ^e and <i>RP11-810B23.1</i> ^e
10	rs2350821	10	86927851	LPS	IL-1 β	PBMC	24 hr	1.3×10^{-8}	<i>RP11-181F12.1</i> ^e
11	rs10908219	11	69602333	<i>Cryptococcus</i>	IFN- γ	T cell	7 days	1.0×10^{-8}	<i>FGF19</i> ^e , <i>FGF4</i> ^e , <i>FGF3</i> ^e , and <i>MIR3164</i> ^b
12	rs11020229	11	93000542	<i>S. aureus</i>	IL-22	T cell	7 days	2.4×10^{-9}	<i>SLC36A4</i> ^{b,c}
13	rs10790723	11	124961985	<i>C. burnetii</i>	IL-1 β	PBMC	24 hr	1.0×10^{-8}	<i>SLC37A2</i> ^{b,c} and <i>AP001007.1</i> ^b
14	rs7310164	12	56185207	<i>C. burnetii</i>	IL-6	PBMC	24 hr	4.7×10^{-9}	<i>ITGA7</i> ^{b,c} , <i>CD63</i> ^c , and <i>MMP19</i> ^b
15	rs11103976	12	86908614	<i>Borrelia</i>	IFN- γ	T cell	7 days	9.2×10^{-9}	<i>MGAT4C</i> ^e
16	rs4491463	15	36645770	<i>E. coli</i>	TNF- α	PBMC	24 hr	1.2×10^{-10}	<i>RP11-47513.1</i> ^e and <i>RP1147513.2</i> ^e
17	rs7256586	19	51390809	<i>C. albicans</i>	IL-6	Blood	48 hr	8.5×10^{-9}	<i>KLK2</i> ^d and <i>KLK4</i> ^e

^aCytokine QTL p values are derived after correcting for age, gender, and cell-count levels.

^bExpression QTL results in blood show a correlation between cytokine QTL SNP and the expression of that gene.

^cThe gene is differentially expressed in response to microbial stimulation in PBMCs.

^dCytokine QTL SNP is in linkage disequilibrium with a missense variant within that gene.

^eThe closest gene to the cytokine QTL is shown.

other 15 cQTLs and their proxies are all located in non-coding regions of the genome, suggesting a possible regulatory function of cQTL loci.

As a second approach, we performed *cis*-eQTL mapping using RNA-sequencing (RNA-seq) data and genotype data from 629 healthy-donor blood samples (LifeLines-Deep cohort) (Ricaño-Ponce et al., 2016; Tigchelaar et al., 2015) as well as eQTL results obtained from publicly available datasets provided by HaploReg (Ward and Kellis, 2012).

As a third approach, we hypothesized that the genes that are differentially regulated in response to different microbial stimuli are the potential causal genes at our cQTL loci. To test this, we extracted all the genes, including the non-coding genes, located in a 500-kb *cis*-window of the 17 cQTLs and tested their expression in PBMCs stimulated with different microbial antigens (Figure S3C). By combining these three approaches, we identified 21 putative causal genes in 12 of the loci (Table 1). In the remaining five loci, the gene nearest to the cQTL SNP is shown. Intriguingly, the genes we identified by these three approaches were all regulatory genes modulating cytokine production, rather than eQTLs directly modulating the transcription of cytokine genes and the production of cytokines themselves. The identified genes were mainly involved in microbial sensing (*TLR1*, *TLR6*,

TLR10, *CSMD1*, *CD63*, *CR2*, and *CD55*), processing molecules (*SLC36A4* and *SLC37A2*), endoplasmic reticulum organization, and cytokine signaling (*IL1F10*, *IL1RN*, and *STAM2*).

TLR1-TLR6-TLR10 Locus Is Associated with Cytokine Production Capacity for Diverse Stimulations

The strongest association among the 17 cQTL loci was at the *TLR1-TLR6-TLR10* locus (Figures 5A and 5B) on chromosome 4, influencing poly(I:C)-induced IL-6 ($p = 3.93 \times 10^{-25}$) and IL-1 β levels ($p = 2.47 \times 10^{-10}$) in PBMCs. This locus also showed significant association with *Coxiella burnetii*-induced IL-1 β ($p = 4.62 \times 10^{-13}$) and TNF- α ($p = 7.59 \times 10^{-8}$) levels, as well as moderate association with 20 different cytokine levels (false-discovery-rate [FDR]-corrected $p < 0.05$) in response to multiple other microbial stimulations (Figure 4B). This locus was also found to be under strong evolutionary selection (discussed later). Since we also had full transcriptomics data obtained by RNA-seq on PBMCs from 70 healthy individuals in the LifeLines-Deep cohort (Tigchelaar et al., 2015) stimulated with *Candida albicans*, we were able to construct co-expression networks for the various alleles of the *TLR1-TLR6-TLR10* locus. Our pathway analysis showed an interesting differential induction of genes that are important for cytokine regulation and

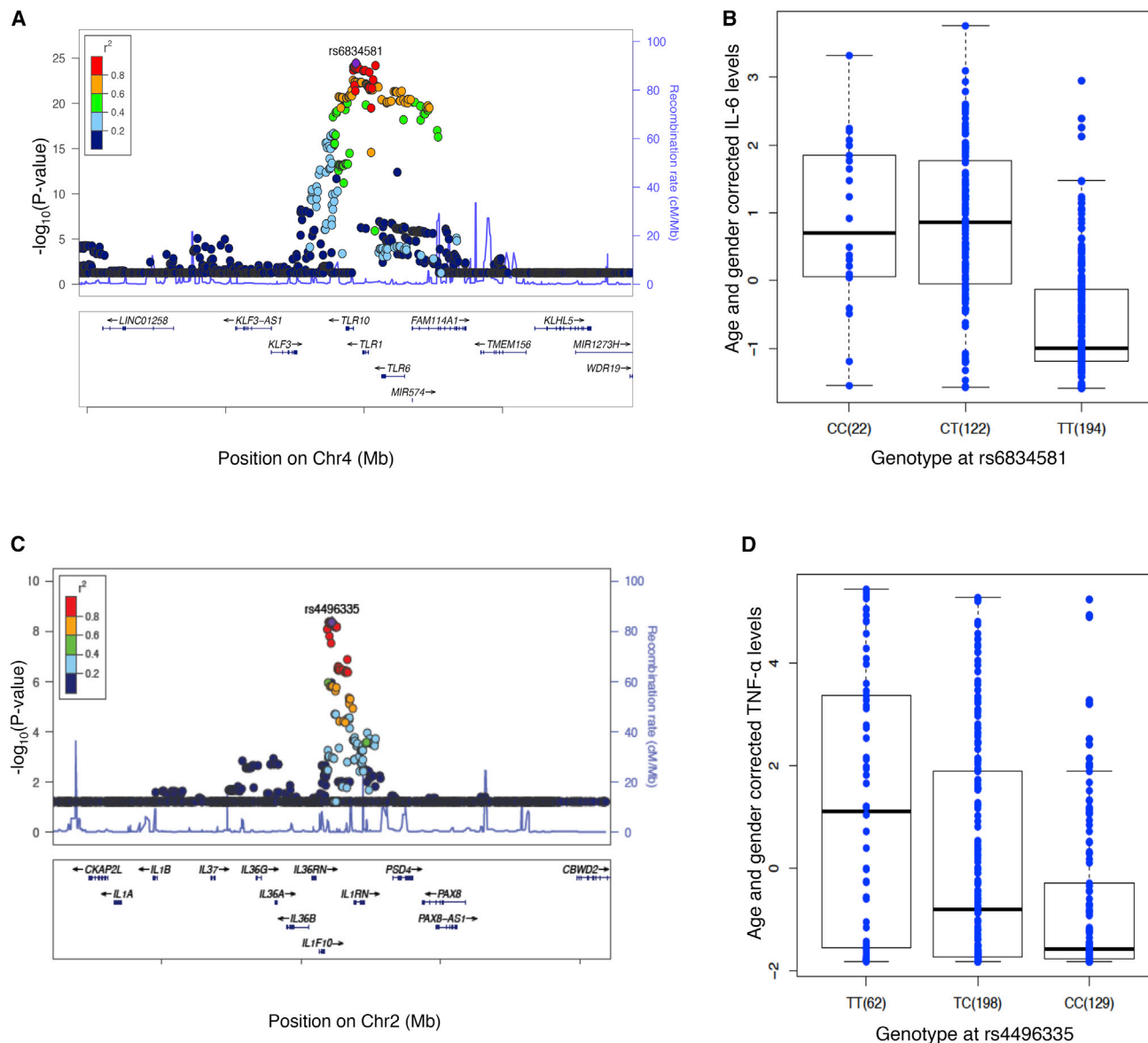


Figure 5. Regional Association Plots and Boxplots for Cytokine QTLs

(A–D) Regional association plots at (A) the TLR10-TLR1-TLR6 locus associated with poly(I:C)-induced IL-6 levels and (C) the IL1F10-IL1RN locus associated with *Cryptococcus*-induced TNF- α levels. The corresponding p values (as $-\log_{10}$ values) of all SNPs in the region were plotted against their chromosomal position. Estimated recombination rates are shown in blue to reflect the local LD structure (based on the CEU population) around the associated top SNP and its correlated proxies (with bright red indicating highly correlated and pale red indicating weakly correlated). Boxplots in (B) and (D) show the genotype-stratified cytokine levels for the TLR and IL1RN loci, respectively. See also Figures S4 and S5.

dependent on the cytokine-inducing allele (rs6834581*C) or the alternative allele (rs6834581*T) (Figure S4). Many of these differentially regulated genes have been shown to be important for cytokine regulation; for example, *TREML4* encodes for a protein crucial for TLR7 signaling and antiviral defense (Ramirez-Ortiz et al., 2015), and *SCGB3A1* encodes for the cytokine-like secretoglobulin family 3A member 1 (or HIN-1), which plays an important role in lung inflammation (Yamada et al., 2009).

IL1F10-IL1RN Locus Is a Shared QTL for *Cryptococcus*- and Influenza-Induced Cytokines

We found a significant cQTL for TNF- α levels in response to *Cryptococcus* (Figures 5C and 5D) on chromosome 2 ($p = 4.22 \times 10^{-9}$). The same SNP (rs4496335) also has an effect on the expression of genes in the *IL-1F10-IL1RN* locus that encodes the IL-1 receptor antagonist. This locus also showed moderate association with IL-6 (3.38×10^{-6}) and IL-1 β levels

(5.51×10^{-6}) in response to *Cryptococcus*. IL-1Ra is a known natural antagonist of the IL-1 receptor pathway, but it was not known whether this also influences *Cryptococcus*-induced cytokine production. In order to validate this finding, we performed a series of experiments in which we show that pre-incubation of PBMCs with IL-1Ra significantly inhibits the induction of cytokines by *Cryptococcus* (Figure S5A). This is the only locus in our 17 cQTL loci that was also moderately associated with influenza-induced TNF- α , IL-1 β , and IL-6 levels (Figure 4A). Unsupervised clustering analysis of cytokines (Figure 2A) showed much stronger similarities between influenza- and *Cryptococcus*-induced responses than between *Cryptococcus* and other fungi-induced responses, suggesting that these two pathogens activate similar inflammatory pathways. This association warrants further study.

CSMD1 and SLC36A4 Loci Specifically Regulate T Cell-Derived Cytokines

We quantified three T cell-derived cytokines (IL-22, IL-17, and interferon [IFN]- γ) after stimulating PBMCs for 7 days with various stimuli. We found three genome-wide significant loci for IFN- γ and two for IL-22 levels (Table 1). The strongest association in these five loci was at the *SLC36A4* locus on chromosome 11 (Figures S5B and S5C), with *Staphylococcus aureus*-induced IL-22 levels ($p = 2.42 \times 10^{-9}$). *SLC36A4* encodes for an amino-acid transporter with a high affinity for glutamine, tryptophan, and proline (Pillai and Meredith, 2011). Amino-acid metabolism (especially tryptophan and glutamine) has been reported to modulate cytokine production (Bosco et al., 2000; Coëffier et al., 2001; Harden et al., 2015). We validated this pathway in our study by showing that blocking glutaminolysis with BPTES (bis-2-(5-phenylacetamido-1,3,4-thiadiazol-2-yl)ethyl sulfide) significantly inhibits *S. aureus*-induced IL-22 production (Figure S5D). The other cQTL for *C. albicans*-induced IL-22 levels was found on chromosome 8 ($p = 2.42 \times 10^{-9}$) on the *CSMD1* gene. *CSMD1* encodes a protein that functions as a complement inhibitor (Escudero-Esparza et al., 2013), and recent studies have shown that IL-22 and the complement pathway influence each other and synergize during host defense against pathogens (Yamamoto and Kemper, 2014). The importance of the complement pathway for modulating cytokine production is underscored by the presence of *CR2* and *CD55* among the genes whose genetic variation regulates cytokine production (Table 1).

Finally, we tested whether the five independent T cell cytokine QTLs were also associated with other cytokines. Although they were moderately associated with cytokines produced in response to other types of stimuli (e.g., *Mycobacterium tuberculosis*, *Cryptococcus*; see Table 1), *CSMD1* and *SLC36A4* loci were specifically associated with IL-22 and IFN- γ production capacity.

KLK2-KLK4 Locus Is Significantly Associated with IL-6 Levels in Whole-Blood Stimulation

We quantified IL-6, IL-1 β , TNF- α , and IFN- γ after stimulating whole blood for 48 hr with various stimuli. We identified one locus on chromosome 19 as significantly associated ($p = 8.50 \times 10^{-9}$) with *C. albicans*-induced IL-6 levels. This locus encodes for kal-

licrein-related peptidases 2 and 4 (KLK2 and KLK4); the latter has been described as inducing IL-6 production through activation of protease-activated receptor 1 (PAR-1), a well-known immune-activated receptor (Wang et al., 2010). This locus is also moderately associated with both monocyte- and T cell-derived cytokines in response to multiple microbial stimulations (Figure 4A).

cQTLs Are Enriched for Regions under Positive Selection and Monocyte-Specific Enhancers and Are Associated with Complex Human Diseases

Some of the cQTL loci were common to several stimuli, but many were stimulus specific (Figure 4B). This finding supports the conclusion that the cytokine response is evolutionarily built around the response to specific pathogens, a process most likely shaped by the selective evolutionary processes exerted by local infections in certain geographical locations. This is also supported by our observation that the cQTL genes are under strong selective pressure. We intersected our 17 cQTLs with the regions in the human genome catalogued as “loci under positive selection” in 230 ancient Eurasian genomes (Mathieson et al., 2015). Our cQTLs were significantly enriched (Kolmogorov-Smirnov test, p value < 0.01) for “genes under positive selection” in the Eurasian genomes (Figure 6A), including the well-known *TLR1-TLR6-TLR10* and *LCT* loci.

The fact that the majority of the cQTL loci were in LD with SNPs in non-coding regions means that they could have regulatory functions. Therefore, we intersected these 17 top cQTLs and their proxies ($r^2 \leq 0.8$) with the ENCODE-defined cell-type-specific enhancers. This showed a significant enrichment of these cQTLs in monocyte-specific enhancers (Figure 6B), suggesting that many of the cQTLs influence gene expression in monocytes and, thereby, alter cytokine production. We also tested whether the 17 cQTLs are associated with human diseases. We intersected the cQTLs with GWAS SNPs known to influence susceptibility to various immune-mediated diseases. Interestingly, SNPs that affect monocyte-derived cytokines are also enriched for SNPs associated with infectious diseases (Figure 6C). In contrast, cQTLs that affect T cell-derived cytokines are enriched for SNPs associated with autoimmune diseases (Figure 6D). In addition, we identified a trend for the association of cQTLs with SNPs associated with other complex human phenotypes, such as blood-related traits and cancer (Figure 6E). These results suggest that proinflammatory cytokines have an important role as underlying mediators in many complex human diseases.

DISCUSSION

In three complementary studies in this issue of *Cell*, we report on the impact of genetic (the present study), environmental (ter Horst et al., 2016), and microbiome (Schirmer et al., 2016) factors on the cytokine production capacity in a large cohort of healthy individuals of Western European background. The present study is broad both at the genomic level (8 million SNPs) and at a functional level: we assessed three types of cellular stimulation models (whole blood, PBMCs, and monocyte-derived macrophages) challenged with a comprehensive panel of bacterial,

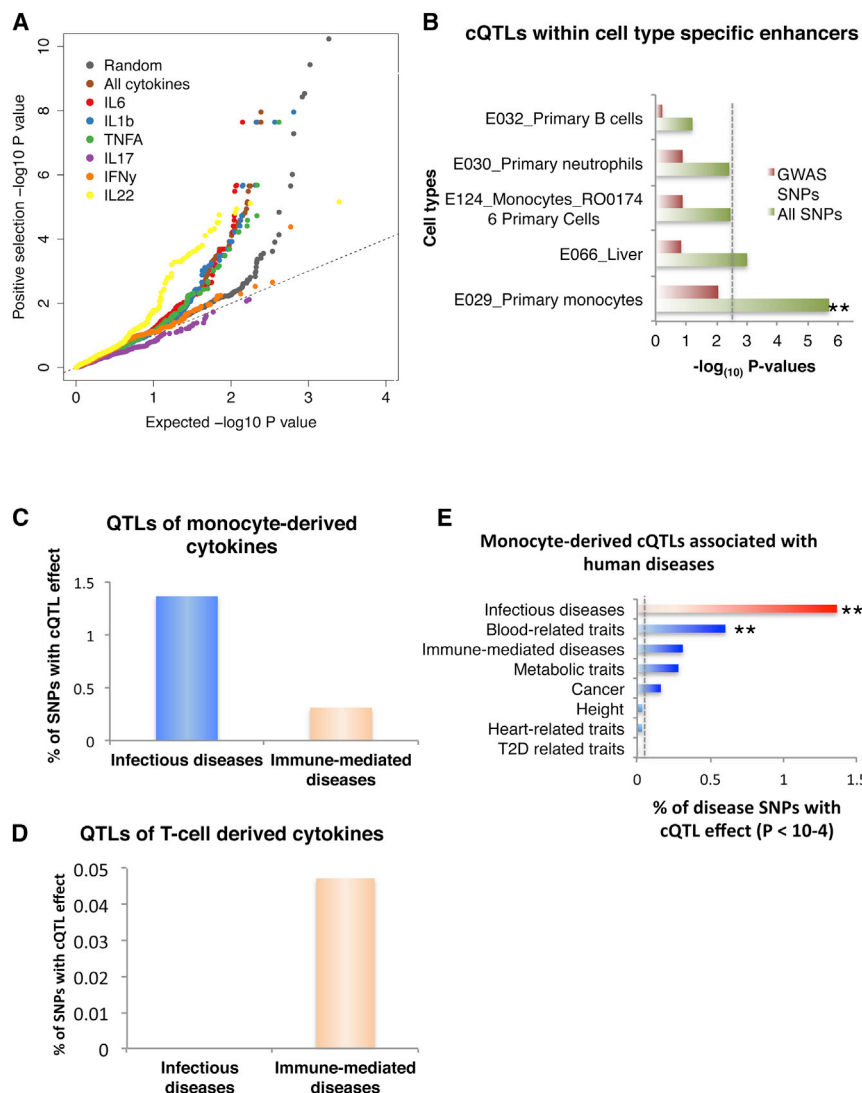


Figure 6. Cytokine QTLs Are Enriched for Human Diseases

(A) A Q-Q plot showing the enrichment of cytokine QTLs is under positive selection. The cytokine QTLs were intersected with loci under positive selection and tested for their inflation compared to a randomly selected set of SNPs.

(B) Impact of genome-wide significant cytokine QTLs on human diseases. All 17 cQTL loci and their proxies ($R^2 > 0.8$) were intersected with cell-type-specific enhancers from the ENCODE project. The x axis depicts the $-\log_{10}$ binomial uncorrected p values; the y axis shows the different cell types. The dotted gray line indicates the significance threshold after Bonferroni correction for the number of cell types tested. Different colors indicate the two sets of background SNPs included for testing enrichment of cQTLs located in any cell-type-specific enhancers.

(C and D) The percentages of SNPs associated with (C) infectious disease and (D) immune-mediated diseases that affect either monocyte-derived or T cell-derived cytokines levels are shown.

(E) The percentage of disease-associated SNPs showing suggestive cytokine QTLs. GWAS SNPs and their proxies from each disease were compared to SNPs associated with “height,” a trait serving as a reference (null) set. The p values of enrichment analysis from Fisher exact tests are indicated by asterisks ($**p < 10^{-4}$). T2D, type 2 diabetes.

fungal, viral, and non-microbial metabolic stimuli. The data presented here support the hypothesis that genetic variation is one of the main factors influencing cytokine responses, with variation of several cytokines, especially the IL-1 β /IL-6 pathway, being mainly regulated by genetic factors (Figures S6 and S7). In line with this, we identify 17 new genome-wide significant loci that influence cytokine production, and we provide genetic and functional validation for their biological importance.

The conclusion that there is a high genetic heritability of cytokine production capacity in the context of microbial stimulation is supported by a recent study showing a strong genetic component in the regulation of multiple immune traits in twins (Roederer et al., 2015). The genes we identify as cytokine regulators can be grouped into two processes: (a) innate immune genes, such as the pattern recognition receptors (e.g., the TLR1/6/10 cluster), and complement modulators; and (b) genes important for antigen processing in the endoplasmic reticulum. Among this last groups of genes, one of them (SLC36A4) is also an amino acid transporter, particularly for tryptophan, proline, and glutamine

cQTLs are associated to a susceptibility to infections, while T cell-derived cQTLs overlap with loci associated to autoimmune diseases. This result has important implications, as it provides a model to gain insight into the mechanistic basis of disease associations in a cell-type-specific manner. This important finding is underpinned by another HFGP study by Aguirre-Gamboa et al. (2016), which describes cell-count QTLs that influence lymphocyte numbers being associated with susceptibility to autoimmune diseases. We were also able to further dissect the impact of the various cQTLs for different pathologies and have, for example, gained important new insight into the preferential impact of monocyte-derived cQTLs for blood-related (hematological) diseases, while these are seen to be less strongly involved in autoimmune diseases or cancer.

We also extracted additional important biological properties characterizing cytokine responses. First, we observed a surprisingly small impact of cell numbers on cytokine production capacity, even in the case of whole-blood stimulation, with only a few exceptions showing a moderate impact. This conclusion is

supported by the results of a previous study published by our group (Li et al., 2016) and demonstrates that the most important impact on cytokine secretion is determined by the intrinsic cellular characteristics. Second, we identified an important pattern in the architecture of cytokine responses: the production capacity of various monocyte-derived or lymphocyte-derived cytokines correlated strongly when cells were stimulated with a specific pathogen, whereas the correlation was poor when comparing bacterial stimuli with fungal and viral stimuli. This makes sense from an evolutionary point of view, as immune responses mainly need to have plasticity to respond to specific infectious pressures in any given geographic area (Netea et al., 2012). This conclusion is supported by the enrichment of cQTL genes among the genes recently reported to be under selection in Eurasian populations (Mathieson et al., 2015). Finally, nearly all the cQTLs described here are *trans*-QTLs, and they were significantly enriched in monocyte-specific enhancers. This suggests that many cQTLs influence the expression of target genes in monocytes to alter cytokine production indirectly.

There are also a number of limitations to the present study. It is possible that we may have underestimated the heritability of cytokine responses by performing a SNP-based analysis. To address this aspect, it will be interesting in the future to assess the heritability of cytokine responses using longitudinal data, as this will allow us to take other factors (e.g., seasonal variation) into account. Moreover, future twin-based studies could take the stimulation aspect into account when performing the heritability estimation of immune parameters. In addition, although this is the most comprehensive study on cytokine production in humans so far, the number of cytokines measured is still relatively limited. We have chosen to study the most important proinflammatory cytokines produced by monocytes, Th1 and Th17 cells; however, future studies should also include other classes of cytokines, such as anti-inflammatory cytokines, IFNs, or chemokines.

In conclusion, we present a comprehensive analysis of how genetic variation affects cytokine production capacity in humans. Our study should be considered in the broader framework of the HFGP, in which environmental, non-genetic host factors and the microbiome have also been shown to influence immune responses (see the accompanying studies by ter Horst et al., 2016, and Schirmer et al., 2016, in this issue). In a first study, we show that age is an important factor, with a specific defect of IFN- γ and IL-22 production in the elderly (ter Horst et al., 2016): future studies in a population of elderly individuals should assess whether the genetic factors identified here also play a role in this process. Similarly, gender and seasonality also influence immune responses. In addition, the accompanying study by Schirmer et al. demonstrates the impact of microbiome variability on cytokine responses. Although it is important to note that the microbiome has an important role, it does seem to have a smaller impact on cytokine production than host genetic factors: while we observe here a 25% to 75% genetic heritability for most of the cytokines mentioned, the microbiome variation explains up to 10% of the cytokine production capacity (Schirmer et al., 2016).

However, the genomic variation may act either directly or indirectly through impacting on other factors, e.g., the microbiome. In

this respect, recent studies have reported host genetic factors that influence the microbiome (Bonder et al., 2016; Davenport et al., 2015; Goodrich et al., 2016; Knights et al., 2014): future studies in the HFGP plan to integrate the complex patient-related and omics databases into comprehensive models to explain cytokine production and other complex immune traits. These complementary studies will be important in understanding the influences on human cytokine responses and their variation and can be used to pinpoint factors that can be modified and targeted for the personalized treatment of immune-mediated diseases. Finally, the HFGP in general, and these three studies reported in this issue of *Cell* in particular, offer an important methodology beyond the analyses of cytokine responses, because they provide a framework for future functional genomics studies looking to assess immune and non-immune biological processes.

STAR★METHODS

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

- **KEY RESOURCES TABLE**
- **CONTACT FOR REAGENT AND RESOURCE SHARING**
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
 - Ethics statement
 - Population cohorts
 - Stimuli
- **METHOD DETAILS**
 - PBMC collection and stimulation experiments
 - Macrophage differentiation and stimulation
 - Whole-blood stimulation experiments
 - Cytokine measurements
 - Cytokine clustering and variance analysis
 - Genotyping, quality control and imputation
 - Cytokine QTL mapping
 - Cytokine heritability estimation
 - Expression QTL analysis to prioritize causal genes
 - Genotype-dependent gene expression analysis at rs6834581
 - Intersection of ENCODE enhancers and regions under positive selection
 - Extraction of infectious disease associated SNPs
 - GWAS SNP extraction and enrichment analysis
 - Enrichment for positive selection
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
- **DATA AND SOFTWARE AVAILABILITY**
 - Online database

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and three tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2016.10.017>.

An audio PaperClip is available at <http://dx.doi.org/10.1016/j.cell.2016.10.017#mmc4>.

AUTHOR CONTRIBUTIONS

M.G.N. and C.W. coordinated the recruitment of cohorts and data generation. M.G.N., V.K., L.A.B.J., and C.W. conceived and directed the study, with input

from all authors. Y.L. analyzed and interpreted the data. M.O., S.S., and M.J. conducted the stimulation experiments and cytokine quantification. R.A.-G., K.T.T.L., P.D., I.R.-P., and V.K. performed genotyping and imputation. R.A.-G. performed the analysis of variance explained by genetics. T.S., A.F.M.J., E.v.d.V., M.v.D., F.v.d.V., and J.W.M.v.d.M. assisted the lab experiments. M.A.S., S.W., R.J.X., A.Z., and L.F. provided the computational framework for the study and critical inputs to the study design. M.G.N., V.K., C.W., Y.L., M.O., S.S., and M.J. wrote the manuscript with input from all authors.

ACKNOWLEDGMENTS

We thank all volunteers in the 500FG and 200FG cohorts of the Human Functional Genomics Project (HFGP) for their participation. We thank Kate McIntyre and Jackie Senior for editing the text. The HFGP is supported by an ERC Consolidator grant (3310372), an IN-CONTROL CVON grant (CVON2012-03), and a Spinoza prize (NWO SPI 94-212) to M.G.N.; an ERC advanced grant (FP/2007-2013/ERC grant 2012-322698) and a Spinoza prize (NWO SPI 92-266) to C.W.; a Dutch Digestive Diseases Foundation grant (MLDS WO11-30) to C.W. and V.K.; a European Union Seventh Framework Programme grant (EU FP7) TANDEM project (HEALTH-F3-2012-305279) to C.W. and V.K.; and a Netherlands Organization for Scientific Research (NWO) VENI grant (863.13.011) to Y.L. The creation and hosting of the online database has been supported by BBMRI-NL, a research infrastructure financed by the Netherlands Organization for Scientific Research (NWO), grant number 184.021.007. We thank Edith Adriaanse, Marije van der Geest, and Marieke Bijlsma for structuring the data into the online database and the MOLFENIS open source team (Dennis Hendriksen, Erwin Winder, Bart Charbon, Fleur Kelpin, Jonathan Jetten, Mark de Haan, Tommy de Boer, David van Enckevort, Chao Pang, and Joeri van der Velde) for designing and implementing the database software.

Received: May 13, 2016

Revised: August 4, 2016

Accepted: October 11, 2016

Published: November 3, 2016

REFERENCES

Aguirre-Gamboa, R., Joosten, I., P.C.M., U., van der Molen, R.G., van Rijssen, E., van Cranenbroek, B., Oosting, M., Smeekens, S.P., Jaeger, M., Zorro, M., et al. (2016). Differential effects of environmental and genetic factors on T and B cell immune traits. *Cell Rep.* Published online November 3, 2016. <http://dx.doi.org/10.1016/j.celrep.2016.10.053>.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* *31*, 166–169.

Bonder, M.J., Kurilshikov, A., Tigchelaar, E.F., Mujagic, Z., Imhann, F., Vila, A.V., Deelen, P., Vatanen, T., Schirmer, M., Smeekens, S.P., et al. (2016). The effect of host genetics on the gut microbiome. *Nat. Genet.* Published October 3, 2016. <http://dx.doi.org/10.1038/ng.3663>.

Bosco, M.C., Rapisarda, A., Massazza, S., Melillo, G., Young, H., and Varesio, L. (2000). The tryptophan catabolite picolinic acid selectively induces the chemokines macrophage inflammatory protein-1 alpha and -1 beta in macrophages. *J. Immunol.* *164*, 3283–3291.

Brodin, P., Jojic, V., Gao, T., Bhattacharya, S., Angel, C.J.L., Furman, D., Shen-Orr, S., Dekker, C.L., Swan, G.E., Butte, A.J., et al. (2015). Variation in the human immune system is largely driven by non-heritable influences. *Cell* *160*, 37–47.

Coëffier, M., Miralles-Barrachina, O., Le Pessot, F., Lalaude, O., Daveau, M., Lavoine, A., Lerebours, E., and Déchelotte, P. (2001). Influence of glutamine on cytokine production by human gut in vitro. *Cytokine* *13*, 148–154.

Davenport, E.R., Cusanovich, D.A., Michelini, K., Barreiro, L.B., Ober, C., and Gilad, Y. (2015). Genome-Wide Association Studies of the Human Gut Microbiota. *PLoS ONE* *10*, e0140301.

Deelen, P., Bonder, M.J., van der Velde, K.J., Westra, H.-J., Winder, E., Hendriksen, D., Franke, L., and Swertz, M.A. (2014a). Genotype harmonizer: auto-

matic strand alignment and format conversion for genotype data integration. *BMC Res. Notes* *7*, 901.

Deelen, P., Menelaou, A., van Leeuwen, E.M., Kanterakis, A., van Dijk, F., Medina-Gomez, C., Francioli, L.C., Hottenga, J.J., Karssen, L.C., Estrada, K., et al.; Genome of Netherlands Consortium (2014b). Improved imputation quality of low-frequency and rare variants in European samples using the 'Genome of The Netherlands'. *Eur. J. Hum. Genet.* *22*, 1321–1326.

Delaneau, O., Zagury, J.-F., and Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. *Nat. Methods* *10*, 5–6.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* *29*, 15–21.

Escudero-Esparza, A., Kalchishkova, N., Kurbasic, E., Jiang, W.G., and Blom, A.M. (2013). The novel complement inhibitor human CUB and Sushi multiple domains 1 (CSMD1) protein promotes factor I-mediated degradation of C4b and C3b and inhibits the membrane attack complex assembly. *FASEB J.* *27*, 5083–5093.

Fairfax, B.P., Humburg, P., Makino, S., Naranbhai, V., Wong, D., Lau, E., Jostins, L., Plant, K., Andrews, R., McGee, C., and Knight, J.C. (2014). Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science* *343*, 1246949.

Genome of the Netherlands Consortium (2014). Whole-genome sequence variation, population structure and demographic history of the Dutch population. *Nat. Genet.* *46*, 818–825.

Goodrich, J.K., Davenport, E.R., Beaumont, M., Jackson, M.A., Knight, R., Ober, C., Spector, T.D., Bell, J.T., Clark, A.G., and Ley, R.E. (2016). Genetic determinants of the gut microbiome in UK twins. *Cell Host Microbe* *19*, 731–743.

Harden, J.L., Lewis, S.M., Lish, S.R., Suárez-Fariñas, M., Gareau, D., Lentini, T., Johnson-Huang, L.M., Krueger, J.G., and Lowes, M.A. (2015). The tryptophan metabolism enzyme L-kynureninase is a novel inflammatory factor in psoriasis and other inflammatory diseases. *J. Allergy Clin. Immunol.* *137*, 1830–1840.

Hirschfeld, M., Ma, Y., Weis, J.H., Vogel, S.N., and Weis, J.J. (2000). Cutting edge: repurification of lipopolysaccharide eliminates signaling through both human and murine toll-like receptor 2. *J. Immunol.* *165*, 618–622.

Hitchcock, P.J., and Brown, T.M. (1983). Morphological heterogeneity among *Salmonella* lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. *J. Bacteriol.* *154*, 269–277.

Howie, B., Marchini, J., and Stephens, M. (2011). Genotype imputation with thousands of genomes. *G3 (Bethesda)* *1*, 457–470.

Knights, D., Silverberg, M.S., Weersma, R.K., Gevers, D., Dijkstra, G., Huang, H., Tyler, A.D., van Sommeren, S., Imhann, F., Stempak, J.M., et al. (2014). Complex host genetics influence the microbiome in inflammatory bowel disease. *Genome Med.* *6*, 107.

Kumar, V., Wijmenga, C., and Xavier, R.J. (2014a). Genetics of immune-mediated disorders: from genome-wide association to molecular mechanism. *Curr. Opin. Immunol.* *31*, 51–57.

Kumar, V., Cheng, S.-C., Johnson, M.D., Smeekens, S.P., Wojtowicz, A., Giannellos-Bourboulis, E., Karjalainen, J., Franke, L., Withoff, S., Plantinga, T.S., et al. (2014b). Immuno-chip SNP array identifies novel genetic variants conferring susceptibility to candidaemia. *Nat. Commun.* *5*, 4675.

Lee, M.N., Ye, C., Villani, A.-C., Raj, T., Li, W., Eisenhaure, T.M., Imboya, 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, Y., Oosting, M., Deelen, P., Ricaño-Ponce, I., Smeekens, S., Jaeger, M., Matzaraki, V., Swertz, M.A., Xavier, R.J., Franke, L., et al. (2016). Inter-individual variability and genetic influences on cytokine responses to bacteria and fungi. *Nat. Med.* *22*, 952–960.

Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S.A., Harney, E., Stewardson, K., Fernandes, D., Novak, M., et al.

- (2015). Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* 528, 499–503.
- Netea, M.G., Wijmenga, C., and O'Neill, L.A.J. (2012). Genetic variation in Toll-like receptors and disease susceptibility. *Nat. Immunol.* 13, 535–542.
- Pillai, S.M., and Meredith, D. (2011). SLC36A4 (hPAT4) is a high affinity amino acid transporter when expressed in *Xenopus laevis* oocytes. *J. Biol. Chem.* 286, 2455–2460.
- R Development Core Team (2015). R: A language and environment for statistical computing (R Foundation for Statistical Computing).
- Ramirez-Ortiz, Z.G., Prasad, A., Griffith, J.W., Pendergraft, W.F., 3rd, Cowley, G.S., Root, D.E., Tai, M., Luster, A.D., El Khoury, J., Hacohen, N., and Means, T.K. (2015). The receptor TREML4 amplifies TLR7-mediated signaling during antiviral responses and autoimmunity. *Nat. Immunol.* 16, 495–504.
- Ricaño-Ponce, I., Zernakova, D.V., Deelen, P., Luo, O., Li, X., Isaacs, A., Karjalainen, J., Di Tommaso, J., Borek, Z.A., Zorro, M.M., et al.; BIOS Consortium; Lifelines Cohort Study (2016). Refined mapping of autoimmune disease associated genetic variants with gene expression suggests an important role for non-coding RNAs. *J. Autoimmun.* 68, 62–74.
- Roederer, M., Quaye, L., Mangino, M., Beddall, M.H., Mahnke, Y., Chattopadhyay, P., Tosi, I., Napolitano, L., Terranova Barberio, M., Menni, C., et al. (2015). The genetic architecture of the human immune system: a bioresource for autoimmunity and disease pathogenesis. *Cell* 161, 387–403.
- Roest, H.-J., van Gelderen, B., Dinkla, A., Frangoulidis, D., van Zijderveld, F., Rebel, J., and van Keulen, L. (2012). Q fever in pregnant goats: pathogenesis and excretion of *Coxiella burnetii*. *PLoS ONE* 7, e48949.
- Schirmer, M., Smeekens, S.P., Vlamakis, H., Jaeger, M., Oosting, M., Franzosa, E.A., Jansen, T., Jacobs, L., Bonder, M.J., Kurilshikov, A., et al. (2016). Linking the human gut microbiome to inflammatory cytokine production capacity. *Cell* 167, this issue, 1125–1136.
- Schramek, S., and Galanos, C. (1981). Lipid A component of lipopolysaccharides from *Coxiella burnetii*. *Acta Virol.* 25, 230–234.
- Shah, T.S., Liu, J.Z., Floyd, J.A., Morris, J.A., Wirth, N., Barrett, J.C., and Anderson, C.A. (2012). optiCall: a robust genotype-calling algorithm for rare, low-frequency and common variants. *Bioinformatics* 28, 1598–1603.
- Smeekens, S.P., Ng, A., Kumar, V., Johnson, M.D., Plantinga, T.S., van Dieën, C., Arts, P., Verwiël, E.T.P., Gresnigt, M.S., Fransen, K., et al. (2013). Functional genomics identifies type I interferon pathway as central for host defense against *Candida albicans*. *Nat. Commun.* 4, 1342.
- Swertz, M.A., Dijkstra, M., Adamusiak, T., van der Velde, J.K., Kanterakis, A., Roos, E.T., Lops, J., Thorisson, G.A., Arends, D., Byelas, G., et al. (2010). The MOLGENIS toolkit: rapid prototyping of biosoftware at the push of a button. *BMC Bioinformatics* 11 (Suppl 12), S12.
- ter Horst, R., Jaeger, M., Smeekens, S.P., Oosting, M., Swertz, M.A., Li, Y., Kumar, V., Diavatopoulos, D.A., Jansen, A.F.M., Lemmers, H., et al. (2016). Host and Environmental Factors Influencing Individual Human Cytokine Responses. *Cell* 167, this issue, 1111–1124.
- Tigchelaar, E.F., Zernakova, A., Dekens, J.A.M., Hermes, G., Baranska, A., Mujagic, Z., Swertz, M.A., Muñoz, A.M., Deelen, P., Cénit, M.C., et al. (2015). Cohort profile: Lifelines DEEP, a prospective, general population cohort study in the northern Netherlands: study design and baseline characteristics. *BMJ Open* 5, e006772.
- Wang, W., Mize, G.J., Zhang, X., and Takayama, T.K. (2010). Kallikrein-related peptidase-4 initiates tumor-stroma interactions in prostate cancer through protease-activated receptor-1. *Int. J. Cancer* 126, 599–610.
- Ward, L.D., and Kellis, M. (2012). HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* 40, D930–D934.
- Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorf, L., and Parkinson, H. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 42, D1001–D1006.
- Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3, 160018.
- Yamada, A., Suzuki, D., Miyazono, A., Oshima, K., Kamiya, A., Zhao, B., Takami, M., Donnelly, R.P., Itabe, H., Yamamoto, M., et al. (2009). IFN-gamma down-regulates Secretoglobin 3A1 gene expression. *Biochem. Biophys. Res. Commun.* 379, 964–968.
- Yamamoto, H., and Kemper, C. (2014). Complement and IL-22: partnering up for border patrol. *Immunity* 41, 511–513.
- Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42, 565–569.
- Zaitlen, N., and Kraft, P. (2012). Heritability in the genome-wide association era. *Hum. Genet.* 131, 1655–1664.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
ProlastinC (AAT)	Grifols	N/A
BSA	SIGMA	A7030
Critical Commercial Assays		
Human IL-1beta ELISA Kit	RD Systems	DY201
Human IL-6 ELISA Kit	PeliKine Compact	M9316
Human TNF- α ELISA Kit	RD Systems	DY210
Human IL-22 ELISA Kit	RD Systems	DY782
Human IL-17 ELISA Kit	RD Systems	DY317
Human IFN γ ELISA Kit	PeliKine Compact	M9333
Human AAT ELISA Kit	RD Systems	DY1268
Human Resistin ELISA Kit	RD Systems	DY1359
Human Leptin ELISA Kit	RD Systems	DY398
Human Adiponection ELISA Kit	RD Systems	DY1065
Human IL-1Ra ELISA Kit	RD Systems	DRA00B
Human IL-18Bp ELISA Kit	RD Systems	DY119
IgM, IgG, IgA	Beckman Coulter	In house
IgG subclasses	Binding Site	BN II Combi Kit
25-hydroxy vitamin D3 measurement	LCMSMS	In house
Testosterone measurement	LCMSMS	In house
Progesterone measurement	LCMSMS	In house
Plasma IL-1beta	Protein Simple	Simple Plex cartridges
Plasma IL-6	Protein Simple	Simple Plex cartridges
Plasma IL-18	Protein Simple	Simple Plex cartridges
Plasma VEGF	Protein Simple	Simple Plex cartridges
Mouse IL-1beta	RD Systems	DY401
TruSeq RNA sample preparation kit v2	Illumina	RS-122-2001
Deposited Data		
ELISA cytokine measurements and other cohort data:	This paper/ BBMRI-NL data infrastructure	https://hfgp.bbMRI.nl/ ; http://bbMRI-eric.eu/bbMRI-eric-directory
RNA sequencing of 88 individuals	This paper/ BBMRI-NL data infrastructure	https://hfgp.bbMRI.nl/
Climatological data	Koninklijk Nederlands Meteorologisch Instituut	http://projects.knmi.nl/klimatologie/daggegevens/selectie.cgi
Human genome for RNaseq mapping: GRCh37.75	Ensembl	http://ftp.ensembl.org/pub/release-75/fasta/homo_sapiens/dna/
Experimental Models: Organisms/Strains		
Human PBMCs and Human Monocytes	Primary/Healthy volunteers	N/A
<i>Bacteroides fragilis</i>	BD Biosciences, Franklin Lakes	NCTC 10584
<i>E. coli</i>	N/A	ATCC 25922
<i>Staphylococcus aureus</i>	N/A	ATCC 29213
LPS	Sigma-Aldrich, St. Louis	<i>E. coli</i> serotype 055:B5
<i>Mycobacterium tuberculosis</i>	N/A	H37Rv
<i>C. burnetii</i>	Bundeswehr Institute for Microbiology, Munich	Nine Mile RSA493
<i>B. burgdorferi</i>	N/A	ATCC strain 35210

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<i>C. albicans</i> blastoconidia	N/A	strain ATCC MYA-3573, UC 820
<i>Aspergillus fumigatus</i>	N/A	V05-27
<i>Cryptococcus gattii</i>	N/A	A1M-R265, AFLP type 6
Influenza virus	N/A	pH1N1 A/Netherlands/602/09
Pam3Cys	EMC microcollections (L-2000)	N/A
PolyI:C	InvivoGen	N/A
PHA	Sigma	N/A
Sequence-Based Reagents		
Life Technologies Globin Clear kit	Illumina	www.illumina.com
Paired End cluster kit	Illumina	www.illumina.com
HiSeq2500 SBS Sequencing reagents	Illumina	www.illumina.com
Infinium CoreExome-24 v1.1 kit	Illumina	www.illumina.com
Software and Algorithms		
R programming language	R Development Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.	https://www.R-project.org/
Custom scripts in the r programming language based on function like: <code>lm {stats}</code> , <code>cor {stats}</code> , <code>cor.test {stats}</code>	R Development Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.	https://www.R-project.org/
RNA sequencing mapping: STAR (version 2.3.0)	Dobin et al., 2013	https://github.com/alexdobin/STAR
RNA read counting: HTSeq (version 0.5.4p3)	Anders et al., 2015	http://www-huber.embl.de/users/anders/HTSeq/doc/overview.html

CONTACT FOR REAGENT AND RESOURCE SHARING

Request should be directed and will be fulfilled by lead author Y.L. (y.li01@umcg.nl).

EXPERIMENTAL MODEL AND SUBJECT DETAILS**Ethics statement**

The HFGP study was approved by the Ethical Committee of Radboud University Nijmegen, the Netherlands (no. 42561.091.12). Experiments were conducted according to the principles expressed in the Declaration of Helsinki. Samples of venous blood were drawn after informed consent was obtained.

Population cohorts

The study was performed in two independent cohorts of ~500 and ~200 healthy individuals of Western European ancestry from the Human Functional Genomics Project (500FG and 200FG cohorts, see www.humanfunctionalgenomics.org). The 500FG cohort comprises 534 adults from Nijmegen, the Netherlands (237 males and 296 females, age range 18–75 years). The 200FG cohort comprises individuals from the ‘Geldersch Landschap’, ‘Hoge Veluwe’, ‘Twickel’, and ‘Kroondomein het Loo’ in the Netherlands (77% males and 23% females, age range 23–73 years old).

Stimuli**Bacteria**

Bacteroides fragilis (NCTC 10584) grown anaerobically overnight at 37°C on blood agar plates (BD Biosciences, Franklin Lakes) was inoculated in 20 mL pre-warmed and pre-reduced Brain Heart Infusion broth (BD Diagnostics, Basel) and again grown anaerobically overnight at 37°C until reaching a stationary growth phase mimicking growth conditions in abscesses. Bacterial suspensions were washed three times in phosphate-buffered saline (PBS; B. Braun Medical B.V., Melsungen) and heat-killed at 95°C for 30 min. Before

heat-killing, aliquots of bacterial suspensions were taken to determine colony-forming unit (CFU) counts. Heat-killed bacteria were washed again and after adjusting the concentration in PBS to 1×10^8 CFU/mL, stored at -80°C . *B. fragilis* was used in the stimulation experiments as 1×10^6 /mL.

E. coli ATCC 25922 was grown overnight in culture medium, washed three times with PBS, and heat-killed for 60 min at 80°C . *Staphylococcus aureus* strain ATCC 29213 was grown overnight in culture medium, washed twice with cold PBS, and heat-killed for 30 min at 100°C ; both *E. coli* and *S. aureus* were used in a final concentration of 1×10^6 /mL. Success of heat-inactivation was confirmed by cultures.

LPS (*E. coli* serotype O55:B5), a TLR4 ligand, was purchased from Sigma-Aldrich (St. Louis) and an extra purification step was performed as described previously (Hirschfeld et al., 2000). Purified LPS was tested in TLR4^{-/-} mice for the presence of contaminants and did not show any TLR4-independent activity. CpG (ODN M362) was purchased from InvivoGen (San Diego) and used at a final concentration of 10 $\mu\text{g}/\text{mL}$.

Cultures of H37Rv *Mycobacterium tuberculosis* (MTB) were grown to mid-log phase in Middlebrook 7H9 liquid medium (Difco, Becton Dickinson, East-Rutherford) supplemented with oleic acid/albumin/dextrose/catalase (OADC) (BBL, Becton Dickinson), washed three times in sterile saline solution, heat-killed and then disrupted using a bead beater, after which the concentration was measured using a bicinchoninic acid (BCA) assay (Pierce, Thermo Scientific, Rockville).

C. burnetii Nine Mile RSA493 (NM) phase I (a gift from the Bundeswehr Institute for Microbiology, Munich, Germany) was cultured on buffalo green monkey cells, and the numbers of *Coxiella* DNA copies were determined using TaqMan real-time polymerase chain reaction as described (Roest et al., 2012). Lipopolysaccharide (LPS) phase determination was performed by sodium dodecyl sulfate polyacrylamide gel electrophoresis and silver staining, using purified phase I (RSA493) and phase II (RSA439) *C. burnetii* NM LPS (kindly provided by R. Toman) as controls (Hitchcock and Brown, 1983; Schramek and Galanos, 1981). *C. burnetii* was inactivated by heating for 30 min at 99°C .

B. burgdorferi, ATCC strain 35210, was cultured at 33°C in Barbour-Stoenner-Kelley (BSK)-H medium (Sigma-Aldrich) supplemented with 6% rabbit serum. Spirochetes were grown to late-logarithmic phase and examined for motility by dark-field microscopy. Organisms were counted using a Petroff-Hauser counting chamber. Bacteria were harvested by centrifugation of the culture at 3000 g for 30 min, washed twice with sterile PBS (pH 7.4), and diluted in the specified medium to the required concentrations of 1×10^6 spirochetes per mL. The serum-resistant *Salmonella typhimurium* strain (phage type 510) was grown at 37°C for 12 hr in nutrient broth (BHI Oxoid, Nepean). Bacteria were collected and centrifuged and then washed three times in PBS. Aliquots were stored at -20°C throughout the study.

Fungi

Heat-killed *C. albicans* blastoconidia (strain ATCC MYA-3573, UC 820) in a concentration of 10^6 CFU/mL were used throughout this study. A clinical isolate of *Aspergillus fumigatus* V05-27 was used for stimulations. Isolates were grown on YAG agar plates for 3 days at 37°C . Fungal spores in the presence of sterile 0.1% Tween 20 in PBS were harvested by gentle shaking, washed twice with PBS, filtered through a 40-mm pore size cell strainer (Falcon, Vienna) to separate conidia from contaminating mycelium. They were then counted by a hemacytometer, suspended at a concentration of 10^8 spores/mL and heat-killed. A final concentration of 1×10^7 /mL was used in the experiments. A clinical isolate of *Cryptococcus gattii* (A1M-R265, APLP type 6) was freshly grown on Sabouraud dextrose agar plates. Afterward a suspension in PBS was heat-killed at 56°C for 24 hr and then quantified at a wavelength of 530 nm on a spectrophotometer. Killing efficiency as well as bacterial and fungal contamination was checked using Sabouraud dextrose and blood agar plates, respectively. Aliquots were stored at -20°C throughout the study.

Influenza virus culture and inactivation. Influenza virus strain pH1N1 A/Netherlands/602/09 (kindly provided by Prof. Ron Fouchier, Erasmus MC, Rotterdam) was grown in the allantoic fluid of embryonated chicken eggs. Viral titers were determined by three independent plaque assays performed on Madin-Darby canine kidney (MDCK) cells. To inactivate the pH1N1 strain, β -propiolactone (BPL) (Acros Organics, Morris Plains) in citrate buffer (125 mM sodium citrate, 150 mM sodium chloride [pH 8.2]) was added to the pH1N1 virus to a final concentration of 0.1% and incubated for 24 hr at 4°C under continuous slow shaking. Inactivated virus was subsequently snap-frozen and stored at -80°C . Virus inactivation was confirmed by three passages in MDCKs where no virus could be detected by plaque assay following the third passage.

TLR ligands and non-microbial stimuli

Pam3Cys, a TLR1/2 ligand, was purchased at EMC microcollections (L-2000) and used in a final concentration of 10 $\mu\text{g}/\text{mL}$.

PolyI:C, a TLR3 ligand, was purchased from InvivoGen and used at a final concentration of 100 $\mu\text{g}/\text{mL}$. PHA was purchased from Sigma and used at a final concentration of 10 $\mu\text{g}/\text{mL}$. Palmitic acid was purchased from Sigma-Aldrich. Human albumin (Albuman) was purchased from Sanquin (Amsterdam). Stock palmitic acid was dissolved in 100% ethanol. Palmitic acid (C16.0) and human albumin were conjugated by warming to 37°C in a water bath before adding together in a 1:5 ratio. The mixture was sonicated for 20–25 min and kept at 37°C until use. The vehicle control for 50 μM C16.0 consisted of 0.025% albumin and 0.025% ethanol. MSU crystals were formed by dissolving 1.0 g of uric acid and 0.48 g sodium hydroxide in 400 mL of sterile water. The pH was adjusted to 7.2 and the solution was sterilized by heating it for 6 hr at 120°C . No LPS contamination was detected by Limulus amoebocyte lysate assay.

METHOD DETAILS

PBMC collection and stimulation experiments

After obtaining informed consent, venous blood was drawn from the cubital vein of volunteers into 10 mL EDTA Monoject tubes (Medtronic, Dublin). The PBMC fraction was obtained by density centrifugation of EDTA blood diluted 1:1 in pyrogen-free saline over Ficoll-Paque (Pharmacia Biotech, Uppsala). Cells were washed twice in saline and suspended in medium (RPMI 1640) supplemented with gentamicin (10 mg/mL), L-glutamine (10 mM) and pyruvate (10 mM). Addition of antibiotics such as gentamycin is a standard method used to avoid contamination of cultures, and it does not influence the ability to induce cytokine production by PBMCs or macrophages (data not shown). The cells were counted in a Coulter counter (Beckman Coulter, Pasadena) and the number was adjusted to 5×10^6 cells/mL. A total of 5×10^5 PBMCs in a total volume of 200 μ L per well were incubated at 37°C in round-bottom 96-well plates with the different stimuli, as indicated above. After 24 h (for early cytokines IL-1 β , TNF- α , IL-6, IL-8, and IL-10), or 7 days of incubation (for IFN- γ and IL-17), supernatants were collected and stored at -20°C until assayed. The stimulation periods were chosen based on extensive studies that showed that 24 hr stimulation was best suited for assessing monocyte-derived cytokines. When cells were cultured for 7 days, this was done in the presence of 10% human pooled serum.

For validation experiments, PBMCs were incubated with recombinant IL-1Ra (R&D Systems, Minneapolis) 10 μ g/ml, 30 min before stimulation with *C. neoformans*. The cytokine production capacity in the absence or presence of IL-1Ra was measured in the supernatants as described above.

Macrophage differentiation and stimulation

We cultured $> 5 \times 10^5$ monocytes in flat-bottom plates with 10% human serum at 37°C and 5%CO₂ for 6 days. After differentiation, the medium was removed and the differentiated macrophages were stimulated for 24 hr. Supernatants were collected and stored in -20°C until used for ELISA.

Whole-blood stimulation experiments

100 μ L of heparin blood was added to a 48-well plate and then stimulated with 400 μ L stimulus (final volume 500 μ L) for 48 hr at 37°C and 5%CO₂. Supernatants were collected and stored in -20°C until used for ELISA.

Cytokine measurements

Concentrations of human IL-1 β , IL-6, IL-10, TNF- α , IL-17, or IFN- γ were determined using specific commercial ELISA kits (PeliKine Compact, Amsterdam, or R&D Systems), in accordance with the manufacturer's instructions. Detection limits were 31 pg/mL (IL-6 (macrophages), 39 pg/mL (IL-1 β and IFN γ), 156 pg/mL (IL-6 (whole blood), IL-22) or 78 pg/mL (TNF- α and IL-17).

Cytokine clustering and variance analysis

Raw cytokine levels were first log-transformed, and cytokine measurements showing little/no variation across individuals were filtered out for the follow-up analysis. We excluded 17 stimulation-cytokine measurements that did not pass our quality control. Unsupervised hierarchical clustering was performed using Spearman's correlation as the measure of similarity. We used Levene's test to check the equality of variance of cytokine levels before and after stimulation.

Genotyping, quality control and imputation

DNA samples of 500 individuals were genotyped using the commercially available SNP chip, Illumina HumanOmniExpressExome-8 v1.0. The genotype calling was performed using Optical 0.7.0 (Shah et al., 2012) using default settings. Samples with a call rate ≤ 0.99 were excluded from the dataset, as were variants with a Hardy-Weinberg equilibrium (HWE) ≤ 0.0001 , call rate ≤ 0.99 and minor allele frequency (MAF) ≤ 0.001 . We identified 17 ethnic outliers by merging multi-dimensional scaling plots of samples with 1000 Genomes data and these were excluded from further analysis (Figure S7). This resulted in a dataset of 483 samples containing genotype information on 518,980 variants for further imputation. The strands and variant-identifiers were aligned to the reference Genome of the Netherlands (GoNL, Genome of the Netherlands Consortium, 2014) dataset using Genotype Harmonizer (Deelen et al., 2014a). The data were phased using SHAPEIT2 v2 (Delaneau et al., 2013) using GoNL as a reference panel. Finally, these data were imputed using IMPUTE2 (Howie et al., 2011) with GoNL as the reference panel (Deelen et al., 2014b). We selected SNPs that showed an INFO score ≥ 0.8 upon imputation for further cytokine QTL mapping.

Cytokine QTL mapping

Both genotype and cytokine data could be generated for a total of 442 individuals. We obtained cell count data measured by FACS for total lymphocytes, T cells, B cells, monocytes and NK-cells from 487 individuals from the 500FG cohort. In total, there were 409 samples with genotype, cytokine, and cell-count data. We excluded 17 samples due to genetic differences. We coded gender information either 0 for females or 1 for males. The actual age, gender, and cell-count information were included as covariables in the linear model to correct the cytokine distributions for QTL mapping. Raw cytokine levels were first log-transformed then mapped to genotype data using a linear regression model with age and gender as covariates. Since the stimulation-cytokine combinations cannot be regarded as completely independent (we observed strong correlations in the cytokine clustering analysis we performed), the total number of

independent tests among all phenotypes will be much less than the number of phenotypes measured, and therefore correcting for the number of SNPs tested for each trait should be sufficient. We considered a p value of $< 5 \times 10^{-8}$ to be the threshold for significant cytokine QTLs.

Cytokine heritability estimation

Using the GCTA tool (Yang et al., 2010), we fitted a linear mixed model to each of the 112 cytokine phenotypes and used restricted maximum likelihood to estimate the variance explained by the approximately 8 million SNPs. Following our method for cQTL mapping, we also included age, gender, and cell counts as covariates for estimating the proportions of phenotypic variance explained by the SNPs. It should be noted that the 95% confidence intervals for the heritability estimate are wide due to the sample size ($N < 1000$) in this study (Zaitlen and Kraft, 2012).

Expression QTL analysis to prioritize causal genes

RNA sequencing of 629 peripheral blood samples from the Lifelines-Deep cohort (Tigchelaar et al., 2015) were investigated to map *cis*-eQTLs. The eQTL mapping strategy and data have been described in detail (Ricaño-Ponce et al., 2016). Briefly, *cis*-eQTL analysis was performed on transcript-SNP combinations for which the distance from the center of the transcript to the genomic location of the SNP was ≤ 500 kb. Associations were tested by non-parametric Spearman's rank correlation test and a p value < 0.05 was considered significant. We also employed HaploReg database v4.1 (Ward and Kellis, 2012) (www.broadinstitute.org/mammals/haploreg/haploreg.php) to extract publicly available eQTL results from blood tissue for cytokine QTL SNPs.

Genotype-dependent gene expression analysis at rs6834581

The PBMCs from 70 individuals in the Lifelines-Deep cohort (Tigchelaar et al., 2015) were stimulated with or without *Candida albicans* as previously described (Smeekens et al., 2013; Kumar et al., 2014b). The RNA sequencing analysis of this dataset has been described by previously (Li et al., 2016). There were 7 CC, 25 CT and 38 TT individuals carrying these genotypes at rs6834581. We combined CC and CT in one group and compared the median expression-fold changes of genome-wide transcripts between two genotype groups (CC+CT versus TT). We focused on the absolute fold changes between these two groups and show the expression levels of the top 30 genes in Figure S4.

Candidate genes located within a *cis*-window of approximately 500 kb of all significant cytokine QTL loci were further tested to see if they responded to any of the pathogens using RNA-seq data from PBMCs of eight individuals, which were stimulated by *Pseudomonas aeruginosa*, *Streptococcus pneumoniae*, *Mycobacterium tuberculosis*, *Candida albicans*, *Aspergillus fumigatus*, and IL-1 α (Li et al., 2016).

Intersection of ENCODE enhancers and regions under positive selection

To perform enhancer enrichment analysis on cytokine QTL SNPs, we extracted all proxies for the 17 cQTLs ($R^2 \geq 0.8$; CEU population as a reference) to have 186 variants. We then intersected these variants with enhancer data of 127 different cell lines available in HaploReg tool v4.1 (www.broadinstitute.org/mammals/haploreg/haploreg.php). HaploReg calculates a background frequency of enhancer overlap in each cell type using two background sets of SNPs. It compares the data with all independent GWAS loci associated in the European population, and with a second set of background SNPs consisting of all 1000 Genomes variants with a frequency $> 5\%$. Initial enrichment of enhancers is calculated using the Binomial test and we applied the Bonferroni correction for multiple testing to define significance levels (0.05/127 cell lines).

Extraction of infectious disease associated SNPs

SNPs associated with a number of infectious diseases that showed a p value $< 9.99 \times 10^{-6}$ were extracted using the GWAS catalog (<http://www.genome.gov/gwasstudies>). As of December 2014, there were two studies on leprosy, two studies on malaria, four studies on tuberculosis, four studies on chronic hepatitis C infection, one study on HPV seropositivity, one study on Dengue shock syndrome, and one study on meningococcal susceptibility. From a systematic search of the literature, we extracted SNPs associated with susceptibility to other infectious diseases but not reported in the GWAS catalog. We found three studies on invasive aspergillosis and two studies on pneumococcal disease (see the list published in Li et al. (2016)).

GWAS SNP extraction and enrichment analysis

GWAS SNPs from the GWAS catalog and their proxies ($r^2 \geq 0.8$ from a 500 kb window) were first extracted, which provided a list of SNPs associated to 122 different human traits and diseases. We selected diseases/traits for which at least 10 independent SNPs were reported to be associated. We then binned these GWAS SNPs into eight categories based on their association to closely related human phenotypes (cancer, immune-mediated diseases, infectious disease, heart-related traits, blood-related traits, metabolic traits, height, and Type 2 diabetes-related traits). Duplicated SNPs were removed from further analysis. We then intersected the SNPs of each category with cQTLs that showed $p < 0.05$ in our study. The Fisher exact test was applied to test the over-representation of cQTL SNPs in infectious disease SNPs using the height-associated SNPs as reference.

Enrichment for positive selection

The cytokine QTLs at different thresholds (1×10^{-7} , 1×10^{-6} , 1×10^{-5} , 1×10^{-4}) were intersected with loci under positive selection (Mathieson et al., 2015). The distribution of positive selection p values on cQTL SNPs were compared with a randomly selected set of non-significant cQTL SNPs ($p > 0.01$) by using Kolmogorov–Smirnov test.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were performed in R. Unsupervised hierarchical clustering was performed using Spearman's correlation as the measure of similarity. The Levene's test to check the equality of variance of cytokine levels before and after stimulation. R-package Matrix-eQTL, was used for cytokine QTL mapping, where linear model was applied with age, gender, and cell-count information included as covariables. The Fisher exact test was applied to test the over-representation of cQTL SNPs in infectious disease SNPs using the height-associated SNPs as reference.

DATA AND SOFTWARE AVAILABILITY

Online database

All data used in this project have been meticulously cataloged and archived in the BBMRI-NL data infrastructure (<https://hfgp.bbmri.nl/>) using the MOLGENIS open source platform for scientific data (Swertz et al., 2010). This allows flexible data querying and download, including sufficiently rich metadata and interfaces for machine processing (R statistics, REST API) and using FAIR principles to optimize Findability, Accessibility, Interoperability and Reusability (Wilkinson et al., 2016).

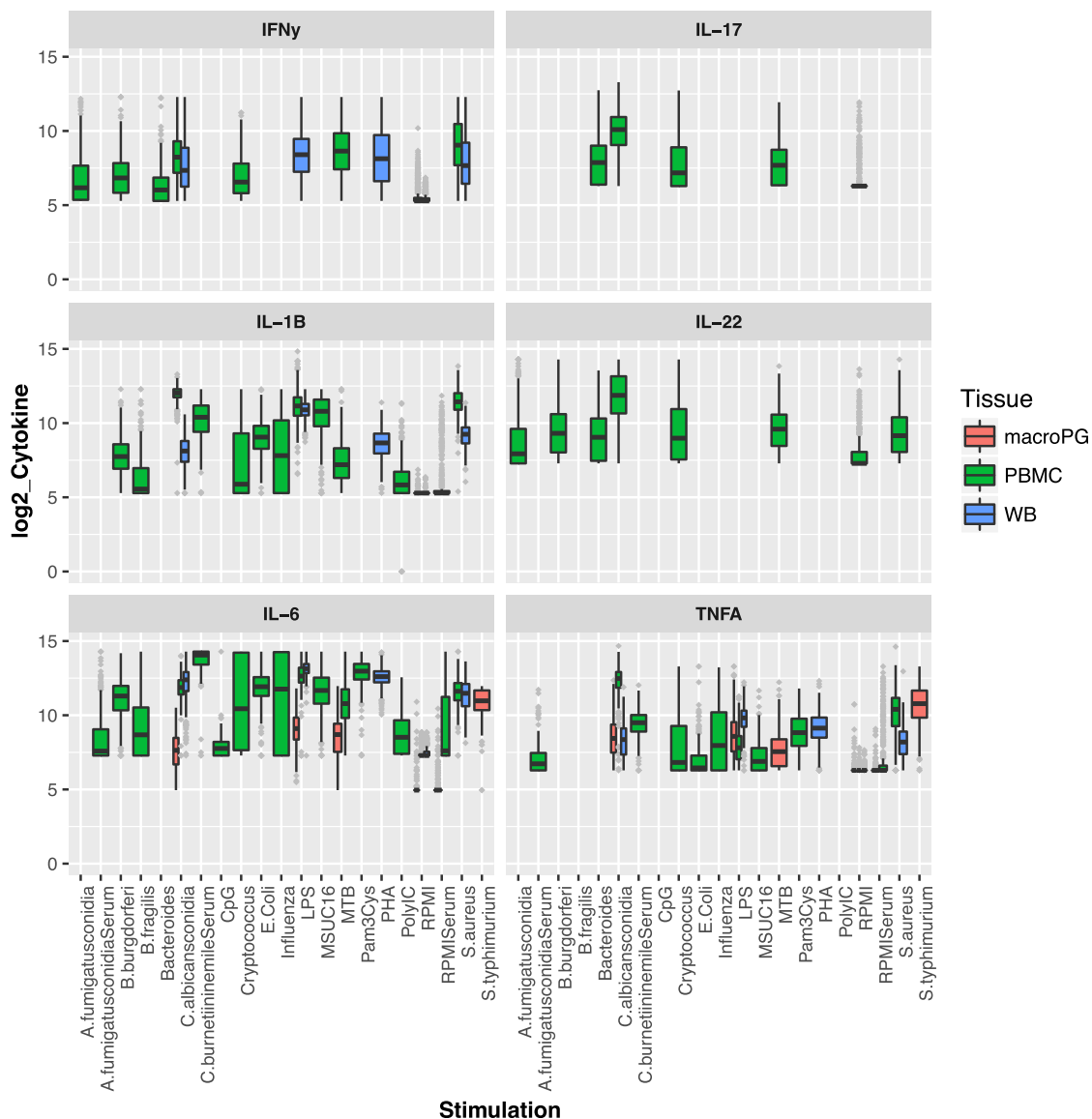


Figure S1. Increased Levels of Cytokines upon Stimulation, Related to Figure 1

The box plots of 6 different cytokines in 500 individuals upon stimulation. The y-axis depicts the log₂ transformed cytokine levels. The x-axis shows different stimulations used to induce cytokine production in different tissues. The color legend indicates the different tissue systems used for stimulation.

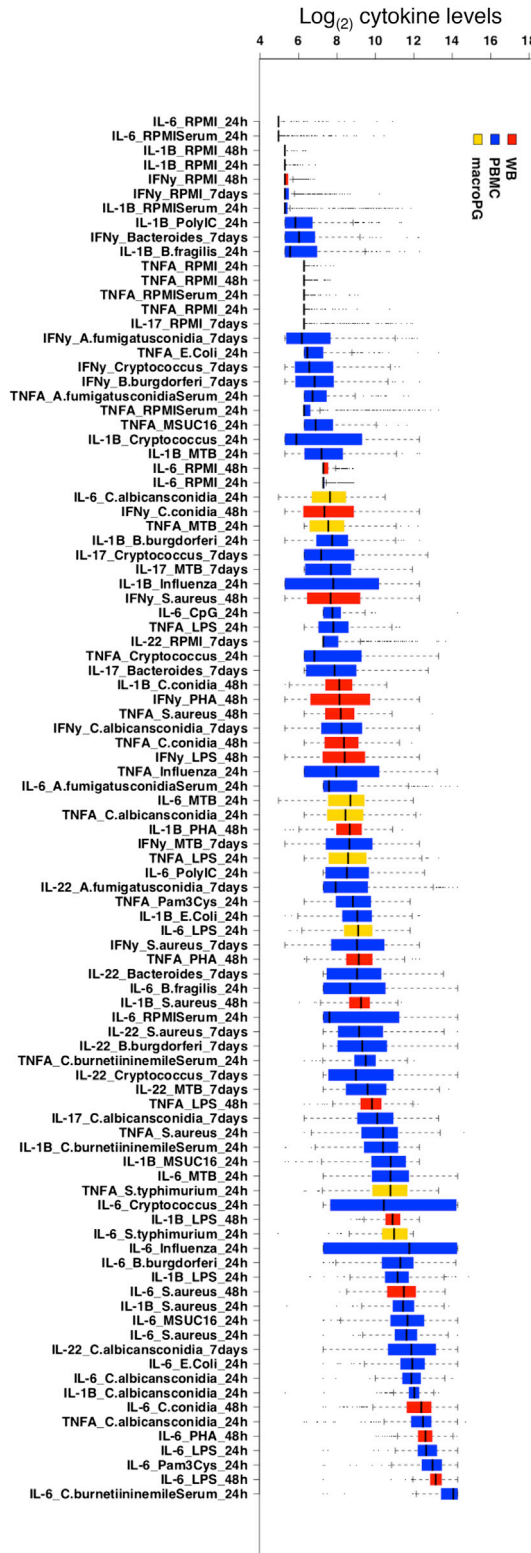
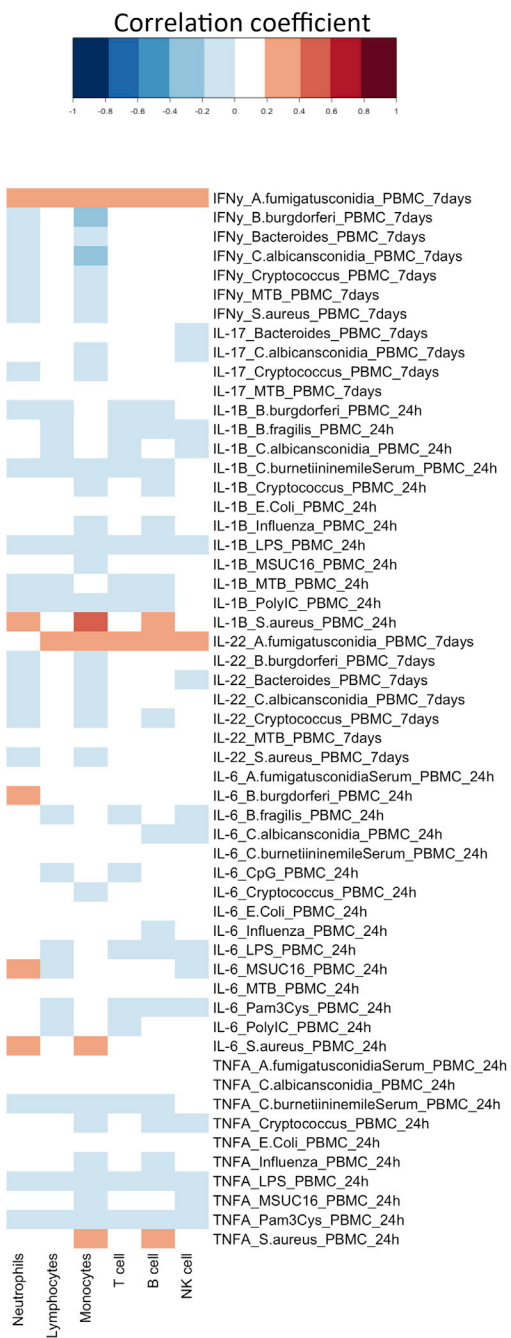


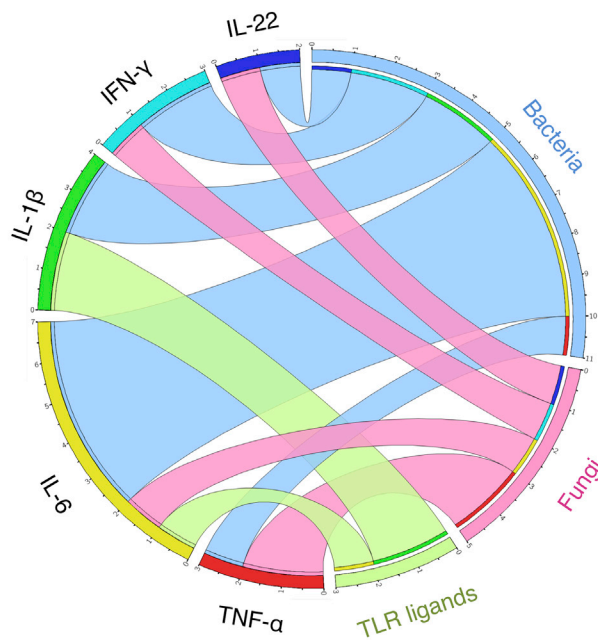
Figure S2. Increased Inter-individual Variation upon Stimulation, Related to Figure 1

Box plots of cytokine levels (x axis) induced upon stimulation (y axis) in 500 individuals sorted based on the median values. The color legend shows the different tissues used for stimulation.

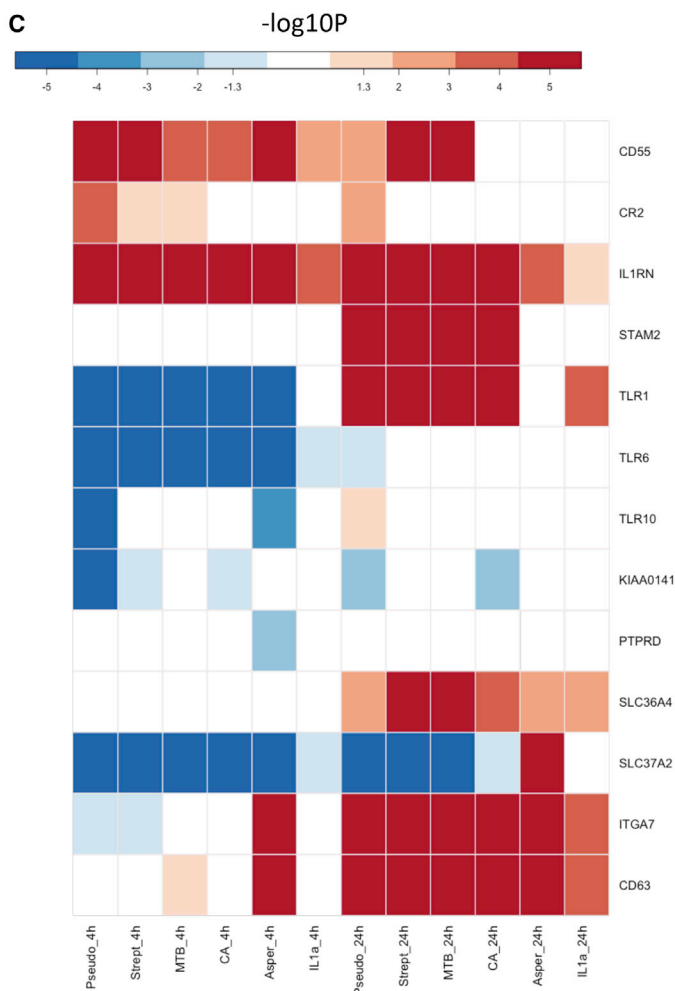
A



B



C



(legend on next page)

Figure S3. Genome-wide Significant Cytokine QTLs, Related to Figure 4

(A) Correlation between PBMC-derived cytokines and cell counts.

(B) Summary of all genome-wide significant cytokine QTLs.

(C) Prioritized causal genes by differential expression analysis for genome-wide significant cytokine QTLs.

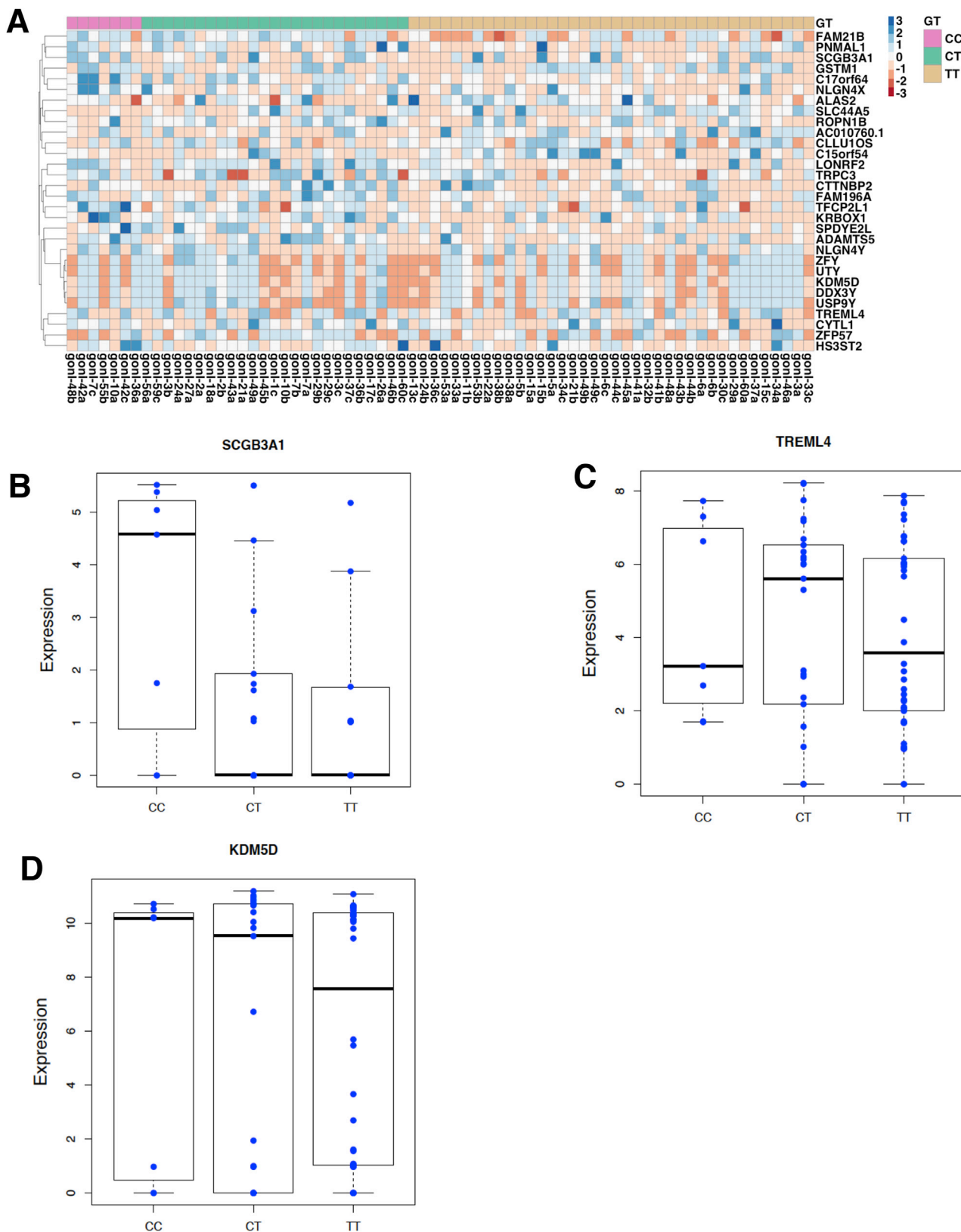


Figure S4. *TLR1-6-10* Locus Genotype Stratified Gene Regulation upon *Candida* Stimulation, Related to Figure 5

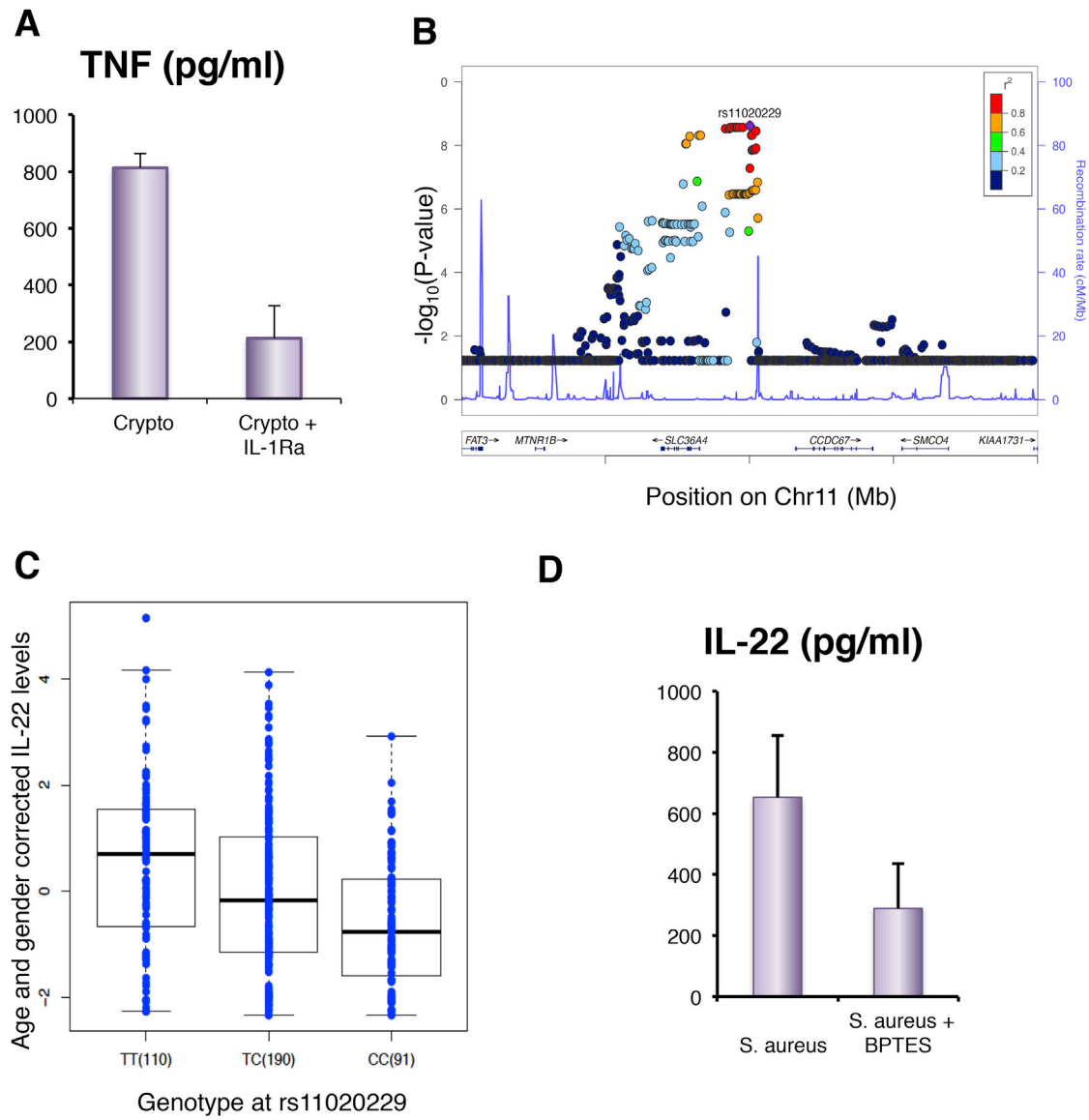


Figure S5. Cytokine QTLs for TNF and IL22 Production, Related to Figure 5

(A) Validation of *Cryptococcus* induced cytokine TNF production.

(B) Regional plot of the association of *SLC36A4* locus with *S. aureus* induced IL-22 levels in PBMCs.

(C) Boxplot of *S. aureus* induced IL-22 levels in PBMCs at *SLC36A4* locus.

(D) Validation of *S. aureus* induced cytokine IL22 production

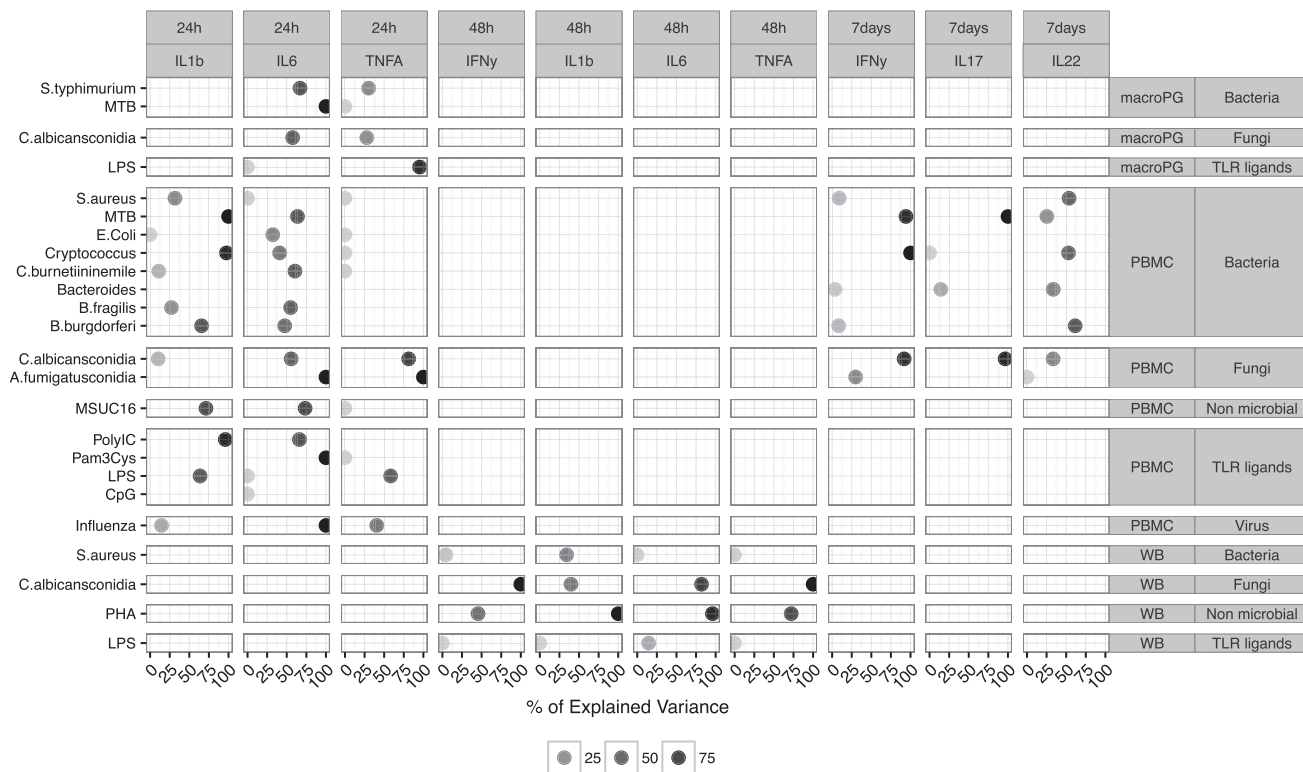


Figure S6. Proportion of Explained Variance of Cytokine Levels by Genetics, Across All the Measurements, Related to Figure 3

A summary of all the estimates of cytokine variance explained by genome-wide SNP data after age, gender, and cell-count correction is shown. The estimates <25% are shown in gray, and the estimates >50% are shown in black.

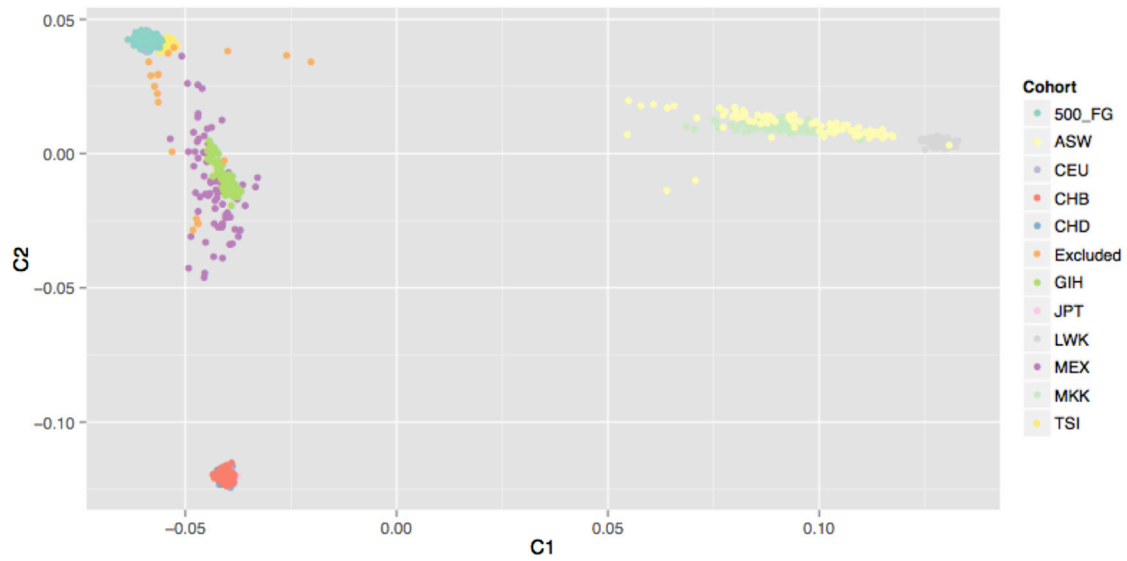


Figure S7. Multidimensional Scale Analysis of Genotype Data from 500FG Cohort, Related to Figures 3 and S6 and STAR Methods

Genome-wide SNP data was used to perform multidimensional scaling analysis across different populations, including 500FG cohort (cohorts are indicated in different colors). The x axis and y axis indicate the first two principal components differentiating different population cohorts. We analyzed 500FG cohort to map cytokine production QTLs in this study.