- STP1 = phenol-preferring phenol sulfotransferase 1
- P-selectin glycoprotein ligand
- gamma-interferon inducible gene IP-30
- HMG-2 = high mobility group 2 protein
- Net = ELK3
- R-PTP-alpha = receptor protein tyrosine phosphatase alpha
- CD49C = Integrin alpha 3
- Arachidonate 5-lipoxygenase = 5-lipoxygenase = 5-LO
- Mx8
- c-myc
- SIP-110 = signaling inositol polyphosphate 5 phosphatase
- WIP/HS PRPL-2 = WASP interacting protein
- CD49C = Integrin alpha 3
- PKC beta = Protein kinase C, beta
- Tis11d = ERF-2 = growth factor early response gene
- ELF-1 = ets family transcription factor
- SH3P18 = SH3 domain-containing protein
- PCM-1 = autoantigen pericentriol material 1
- pLK = homologue of Drosophila polo serine/threonine kinase
- CD105 = endoglin
- Cytotoxic ligand TRAIL receptor
- CD103 alpha = Integrin alpha-E
- ATM
- GRK5 = G protein-coupled receptor kinase 5
- Glutathione peroxidase 1
- IL-16
- CD62L = L-selectin = LAM-1
- c-myb
- CD31 = PECAM-1
- CD33
- SLAP = src-like adapter protein
- CD11A = Integrin, alpha L = LFA-1 alpha chain
- Transforming growth factor, beta receptor
- PARP
- MDR1 = Multidrug resistance-associated protein 1
- calcium/calmodulin-dependent protein kinase II
- CD64
- TTG-2
- C-C chemokine receptor 1 = CC CK1
- HNPP = nuclear phosphoprotein
- MAPKAP kinase (3pK)
- cytohesin-1
- CD115 = CSF-1 receptor
- thymosin beta-4
- Myeloid cell nuclear differentiation antigen
- Peroxisome assembly factor-2 (PAF-2)
- Peroxisome biogenesis disorder protein 1 (PEX1)
Using the bacterial diversity data set, 1,208 unique genes were selected based on having a Locus Link assignment (http://www.ncbi.nlm.nih.gov/LocusLink). \( p \) values for the frequency of 184 GO annotation terms (those that had at least 5 occurrences in the data set) were calculated based on the hypergeometric distribution. Those annotation terms found to have \( p \) values less than 1% within the common induction (red) and common repression (green) clusters are displayed with their GO identification numbers and example genes from the 1,208-gene data set. The GO annotations for genes in the entire bacterial diversity data set were extracted in batch from SOURCE (http://genome-www.stanford.edu/source), using the Locus Link identification number for each gene. The resulting 1,208 unique annotated genes were used for the analysis, among which 184 annotation terms occurred for at least 5 genes. We compared the frequency of these 184 annotations in the data set as a whole to that in the common induction and common repression clusters, using the hypergeometric distribution to calculate \( p \) values (Tavazoie, S., Hughes, J. D., Campbell, M. J., Cho, R. J. & Church, G. M. (1999) \textit{Nat. Genet.} 22, 281-285) (see equation below). Let \( N = 1208 \) denote the total number of genes under consideration and \( A \), the number of these genes with a particular annotation. The chance of observing at least \( x \) genes with that annotation in a random subset of \( n \) genes is given by

\[
p(x; N, A, n) = 1 - \sum_{i=0}^{\min(A, n)} \binom{A}{i} \left( \frac{N - A}{n-i} \right) \binom{N}{n-i}, \text{ where } \binom{N}{n} = \frac{N!}{n!}\left( N - n \right)!
\]
**Common induction**

- Detection of bacteria/bacterial products
- Systemic effects
  - Fever (IL-1, IL-6)
  - Mobilization of metabolites (TNFα)
  - Shock (TNFα)
  - Production of acute-phase proteins (IL-6)
- Cell-to-cell signaling
  - Activation (IL-1, IL-6, TNFα)
  - Stimulation of differentiation (GM-CSF, LIF)
  - Chemoattraction (MIP-1, MIP-2, PARC)
- Intracellular signaling
  - +/- regulation of NF-kB signaling pathway (NFκB1, Rel, IκB)
- Metabolic processes
  - (Adenosine deaminase, Thioredoxin)
  - +/- regulation of apoptosis (cIAP, LTA)

**Common repression**

- Chemotaxis to infection site
  - MCP-1, n-Formyl peptide receptor, CCR1, CD31
- Recognition & destruction of microbe
  - CD64, WASP, CYBB, NCF-1
- Presentation of microbial peptides
  - Cathepsin B, MHC Class II, IP30, CD13
chitinase precursor
APR=immediate-early-response gene
RGS1=regulator of G-protein signaling
APR=immediate-early-response gene
Unknown
Similar to CD95=Fas
IL-15 receptor alpha chain
Thioredoxin
Hs.172051 ESTs
Hs.105499 ESTs
Hs.125887 ESTs
Similar to polyprotein-encoding gene 1
IRF-7=interferon regulatory factor-7
Hs.104444 ESTs
Hs.230206 ESTs
PML
Staf50
HEM45=interferon-inducible PML nuclear bodies-associated protein
Glucocorticoid receptor
Hs.123664 ESTs
IFN-γ=interferon-gamma-inducible myeloid differentiation transcriptional activator
MxB=interferon-induced cellular resistance mediator protein
IFP35=interferon-induced leucine zipper protein
Unknown
Unknown sc_id = 4951
2'-5' oligoadenylate synthetase E
Hs.56009 ESTs
Hs.82554 ESTs
HEM45=gpISG20=interferon-inducible PML nuclear bodies-associated protein
HEM45=gpISG20=interferon-inducible PML nuclear bodies-associated protein
TRAIL=Apo-2 ligand
Similar to 5'-3' exonuclease=cytoplasmic exoribonuclease mXRN1p
Homo sapiens mRNA for deoxyribonuclease III dm3 gene
IP-10
Similar to myosin-binding protein C slow-type muscle
MIF=macrophage migration inhibitory factor
AIM2=interferon-inducible protein
Supplemental Tables & Figures:
(Enhanced versions also available at http://relman.stanford.edu/hostresponse)

**Supplemental Figure i:**
A subset of transcripts are induced exclusively by ionomycin/PMA treatment.

Data are the same as in Fig. 1a. Transcripts with multiple copies indicate multiple measurements.

**Supplemental Figure ii:**
A subset of genes in the common repression response are similarly repressed during *B. pertussis* infection of U937 cells.

Macrophages differentiated from the U937 cell line were infected with *B. pertussis* virulent (Bp338), avirulent (Bp537), and isogenic mutant strains (BpA2-6, BpTox6) as in Fig. 4c. Data were selected based on the same criteria as the bacterial diversity and dose response data sets, except genes were filtered based on expression of at least 1.5-fold below the mean for at least four of the sampled time points. Genes that overlapped in expression with the common repression cluster from the bacterial diversity data set are highlighted in red.

**Supplemental Table iii:**
Enrichment of GO annotation terms in the bacterial diversity data set.

Using the bacterial diversity data set, 1208 unique genes were selected based on having a LocusLink assignment ((http://www.ncbi.nlm.nih.gov/LocusLink). *p*-values for the frequency of 184 GO annotation terms (those that had at least 5 occurrences in the data set) were calculated based on the hypergeometric distribution. Those annotation terms found to have *p*-values less than 1% within the Common Induction (red) and Common Repression (green) clusters are displayed with their GO identification numbers and example genes from the 1208-gene data set.

The Gene Ontology annotations for genes in the entire bacterial diversity data set were extracted in batch from SOURCE (http://genome-www.stanford.edu/source), using the Locus Link identification number for each gene. The resulting 1208 unique annotated genes were used for the analysis, among which 184 annotation terms occurred for at least 5 genes. We compared the frequency of these 184 annotations in the data set as a whole to that in the common induction and common repression clusters, using the hypergeometric distribution to calculate *p*-values (Tavazoie, 1999 Nature Genetics) (see equation below). Let \(N\)=1208 denote the total number of genes under consideration and \(A\) the number of these genes with a particular annotation. The chance of observing at least \(x\) genes with that annotation in a random subset of \(n\) genes is given by

\[
p(x; N, A, n) = 1 - \sum_{i=0}^{x-1} \binom{A}{i} \binom{N-A}{n-i}, \text{ where } \binom{N}{n} = \frac{N!}{n!(N-n)!}.
\]
**Supplemental Figure iv:**
Physiological programs represented in the common response.

Physiological programs represented in the common induction (a) and common repression (b) response clusters. Representative examples in the “common induction” or “common repression” clusters associated with specific features of the physiological program of innate immunity are indicated.

**Supplemental Figure v:**
Modeling of the dose response reveals discrimination.

The response difference profiles (as in the fourth panel of Fig. 3) for 8 different genes are displayed in 2-dimensional format. The correlation of the fitted response to the observed response is displayed below each of the plots. The x-axis represents time, the y-axis, exposure dose, and the z-axis, the expression level (log2) of the zero-transformed data. The surface represents the least squares fit of the expression data to the model. The subtraction of the two expression surfaces as defined by $\Delta(d,t)$.

**Supplemental Figure vi:**
A subset of genes varies in expression among the donors.

As described in the text (Donor expression differences), the expression profiles of a cluster of genes identified in the bacterial diversity dataset (Fig.1a, gray bar), was assessed in the response of 3 separate donors, including a replicate experiment to verify the differential response of the first donor to *E. coli.*