Supplemental Information

Supplemental Results

Establishment and quality control of the multiplex gene expression pipeline for the measurement of localization-dependent gene expression fingerprints

We aimed for establishment and control of a robust data analysis pipeline to assess multiplex qPCR data for small bulk cell populations and for single cells (1). Prerequisite to qPCR analysis of bulk cells is accurate normalization using an internal control with constant expression values, independent of experimental conditions. Since no such universal control gene exists, identification and verification of several candidate genes is of utmost importance (2). Out of eleven potential reference genes, *B2M* and *PGK1* were selected as reference genes in HeLa cells. The choice was based on low variance of expression over all conditions (Fig. S3a, left), expression levels in-range with those of selected genes (Fig. S3a, right) and robust regression coefficients of CT values (Fig. S3b) both at the level of small bulk cell populations and of single cells.

For each experiment, we ran multiple gene expression arrays, each corresponding to infection with WT or one of five *Shigella* mutants. In order to assess their comparability, the cDNA from the same uninfected control samples was applied to all arrays within one experiment, and both their bulk cell CTs (CT_{bc}) and single cell CTs (CT_{sc}) were used to test the inter-chip variance (Fig. S3c, left and right, respectively). Only experiments that showed small differences (+/−1 CT) of ΔCT_{bc} and CT_{sc} from median values were used for further analysis.

Next, we aimed at controlling the experimental conditions and assuring reproducibility. Assessing the impact of CCF4 treatment on transcriptional signatures

by PCA revealed a diffuse, non-separable distribution of unloaded and CCF4-loaded cells, indicating no global impact of CCF4 on gene expression (Fig. S3d).

To stop infection and enzymatic activities of transcription, as well as to preserve the overall availability of mRNA, cells were shifted to, and further processed on, ice. Intact living single cells were directly sorted into lysis buffer. PCA showed that transcriptional signatures of *Shigella*-infected cells strongly separate from uninfected cells along the first PC axis, indicating significantly different gene expression signatures activated upon bacterial challenge (Fig. S3e). Furthermore, cells sorted immediately (0 h) or only after a delay on ice of 1 h or 2 h, yielded non-separable data points, indicating that this had no global impact on gene expression profiles (Fig. S3e). This suggests that both transcription and infection were stalled at 4°C.

In addition, the homogeneity of the results originating from multiple independent experiments demonstrates the sensitivity and reproducibility of our analytical approach (Fig. S3).

Quality assessment of single cell analysis

In order to rule out detection biases between lowly and highly expressed genes, we examined the concordance between the average CT_{bc} (obtained from 20 cells) and their corresponding cumulative CT_{sc} (cCT_{sc}), a computed CT value corresponding to the sum of expression from 20 single cells. Two out of three independent experiments showed that for low cCT_{sc}, deviations from CT_{bc} were in the order of magnitude of technical noise (+/−1 CT), whereas high cCT_{sc} of poorly expressed genes strongly diverged from CT_{bc}, most likely due to low signal-to-noise ratios in the single cell expression measurements (Fig. S7a). To distinguish between valid and potentially non-specific, low expression measurements, and thus to maximize global concordance

between single and bulk cell measurements, a cutoff (CT_{max}) was defined, above which all CTs were regarded as non-expressed (see Experimental Procedures). To this end, two global measures of concordance, the absolute mean difference between CT_{bc} and the cCT_{sc} (red line) and the average root mean square deviation (RMSD) (blue line) were evaluated for a range of CT cutoffs (Fig. S7b). This revealed a CT_{max} of 23 to yield globally optimal concordance (Fig. S7b). Applying this cutoff retained valid measurements for 90.5% of all gene-stage pairs (green line) and 82 out of 96 genes showed only moderate deviations between the cCT_{sc} and CT_{bc} (RMSD≤1.5). To further investigate the behavior of genes, for which the cCT_{sc} showed low concordance with CT_{bc}, two types of box plots were generated. The first displays the differences between the cCT_{sc} and CT_{bc} on a per-gene basis for all infection stages (Fig. S7c). The second plot shows cCT_{sc} values overlaid on their corresponding bulk cell measurements (Fig. S7d). Most of the poorly concordant genes were pro-inflammatory genes, such as *TNFα*, *HBD3*, *IL-8* and *CCL20*, whose deviance from the bulk cell data was often restricted to individual infection stages (e.g. *IL-8*, Fig. S7d). Taken together, applying a CT_{max} of 23 removed low, non-specific single cell expression measurements, while still allowing analysis of ~90% of the investigated genes.

Supplemental Materials and Methods

Cell culture, infection experiments and imaging of Caco-2 TC7 cells

Intestinal epithelial Caco-2 TC7 cells (kindly provided by P. Sansonsetti) were grown in DMEM supplemented with 10% decomplemented FCS, 1% HEPES, 1% non-essential amino acids and 1% penicillin/streptomycin (100 U/ml and 100 µg/ml, respectively) at 37°C and 10% CO_2. For experimentation with polarized epithelial cells, 10^6 cells were seeded onto 6-well polyester membrane transwell inserts (8 μm pore size, Corning) and grown upside down for 14 days with medium replacement every 2-3 days (Fig. S4c). To control polarization, the transepithelial resistance (TER) was surveyed and transcriptional profiles of differentiated (TER>700 Ohm/cm^2) and non-differentiated cells were compared (Fig. S4d). For the FRET assay, transwells were washed 3x with PBS and loaded with 0.25 µM CCF4/AM substrate (Invitrogen) on both sides of the transwell inserts. For basolateral infection, transwell plates containing bacteria at a MOI 75 in the insert were spun for 5 min at 200 x g and 37°C to facilitate transmembrane migration of bacteria and were further incubated for 25 min at 37°C to promote bacterial invasion. Subsequently, cells were extensively washed and further incubated for 2.5 h at 37°C.

For imaging of Caco-2 cells, both sides of the inserts were fixed in 4% PFA for 30 min, incubated in PBS/1% BSA for 30 min at RT and stained with anti-villin antibody (Abcam) for 1 h at RT, followed by 1 h incubation at RT with anti-mouse FITC-conjugated antibody (Jackson). Subsequently, membranes were mounted on cover slides using ProlongGold anti-fading kit (Invitrogen) containing DAPI to stain nuclei. Microscopic imaging was performed on a Spinning disc microscope (Perkin Elmer) using a 40x objective with excitation at 405 nm (for DAPI), 488 nm (for Villin) and 561 nm (for dsRed *Shigella*). Images of 23 confocal sections were acquired at 1 µm

intervals using a Hamamatsu C9100-50 camera and were further processed by maximum projection using the Volocity software (Perkin Elmer).

Imagestream

For Imagestream analysis, cells were trypsinized and fixed in 4% PFA for 30 min and then resuspended in PBS. High-throughput multispectral fluorometric technology was used to control and verify FACS-based analysis with microscopic images using the imaging flow cytometer (ImageStream ISX, Amnis Corporation, Seattle, WA). For each sample, 15,000 – 25,000 cells were imaged and data were analyzed using the manufacturer's software (IDEAS, Amnis Corporation). Briefly, fluorometric compensation was digitally calculated based on single-stain controls. Focused single cells were selected based on the digital plot of aspect ratio, area and gradient of the bright field images. Distinct populations of cells were identified based on the fluorescence signals and the spot count algorithm was used to identify attached or internalized bacteria.

Extended Information on single cell data processing and analysis

Estimation of a specific expression threshold. In order to assess concordance of single and bulk cell measurements, we aimed at comparing the average expression value of bulk cells (CT_{bc}) and the sum of expression values of single cells, the so-called cumulative single cell CT (cCT_{sc}). To compute cCT_{sc}, non-normalized, CT_{sc} were transformed into expression threshold (ET) values by subtracting them from a given maximal specific CT value (CT_{max}), added up to create an *in silico* pool and finally transformed back into CT space (3). To determine CT_{max}, a cutoff, above which all CT_{sc} measurements exceeding that threshold will be discarded, a range of CT

values between 20 and 40 was tested. To this end, two measures of concordance were applied: the mean difference between cCT_{sc} and CT_{bc} over all infection stages and the RMSD.

$$RMSD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(cCT_{SC}(i) - \text{mean}(CTs_{BC}(i))\right)^2}$$, n: number of measured conditions

To determine a CT_{max} yielding the best global concordance between single cell and bulk cell data, the mean difference and RMSD were averaged over all genes in the combined data set. Fig. S6b demonstrates that at a cutoff CT_{max} of 23±0.02, both measures attained optimal values, while retaining 90.5% of all measured gene-stage pairs (Fig. S7b). Thus, all subsequent single cell analysis was based on ET values computed using a cutoff CT_{max} of 23.

Differential expression analysis. In the likelihood ratio test (LRT), the mean expression (μ) and proportion (π) of single cells expressing a given gene are simultaneously compared and the goodness-of-fit of two alternative models to the measured data is determined (4). In the null model, parameters μ and π are jointly estimated. The alternative model estimates separate parameter values for the two experimental units to be compared (i.e. two stages of infection). Goodness-of-fit of the two models is individually evaluated by the likelihood function:

$$L(\theta \mid y,v) = \prod_{k} \pi_k^{\,n_k}(1-\pi_k)^{I-n_k} \prod_{i \in S_k} g(y_{ik} \mid \mu_k, \sigma^2)$$

y and v are the vectors of ET values for a given gene across the two groups, θ is the vector of parameters to be fitted to the data, S_k is the set of cells expressing the gene in group k (i.e. $S_k = \{i : v_{i,k}=1\}$), $n_k = \Sigma_i v_{i,k}$ is the number of cells expressing the gene in group k and g is the density function of the log-normal distribution with parameters μ_k and σ^2.

The LRT statistic is defined by the ratio of the likelihoods of the null and alternative model. It asymptotically follows a χ^2 distribution with two degrees of freedom under the null hypothesis (5), which allows the computation of p-values for differential expression.

Gene-pair correlation analysis. We computed Spearman's rank correlation coefficients over the 20 measured ET values for all pairs of genes in a given experiment *e* and infection stage *s*:

$$c_{e,s}(g_i,g_j) = \rho(ETs_{e,s}(g_i), ETs_{e,s}(g_i))$$

To asses significance of these correlations, we compared them to a background set of N_B=10,000 correlations computed over randomly permuted orderings of the 20 ET values and defined an empirical p-value:

$$P_{e,s}(g_i,g_j) = \left(\sum_{b=1}^{N_B} I(c_{e,s}(g_i, perm_b(g_j)), c_{e,s}(g_i,g_j)) + 1 \right) / N_B + 1, \text{ with}$$

$$I(c_1,c_2) = \begin{cases} 1, \text{ if } \text{abs}(c_1) \geq \text{abs}(c_2) \\ 0, \text{ otherwise} \end{cases}$$

Gene-pair correlations and significances of the two retained experiments were aggregated into a combined correlation c_{comb} and p-value P_{comb} (Fig. S6e):

$$c_{comb\,s}(g_i,g_j) = \begin{cases} \sum_{e=1}^{2} c_{e,s}(g_i,g_j)/2 & \text{if } \text{sign}(c_{1,s}(g_i,g_j)) = \text{sign}(c_{2,s}(g_i,g_j)) \\ c_{\arg\min_e (P_{e,s}(g_i,g_j)),s}(g_i,g_j) & \text{otherwise} \end{cases}$$

$$P_{comb\,s}(g_i,g_j) = \begin{cases} \prod_{e=1}^{2} P_{e,s}(g_i,g_j) & \text{if } \text{sign}(c_{1,s}(g_i,g_j)) = \text{sign}(c_{2,s}(g_i,g_j)) \\ P_{\arg\min_e (P_{e,s}(g_i,g_j)),s}(g_i,g_j) & \text{otherwise} \end{cases}$$

Pathway correlation analysis. Based on these measures, we evaluated the degree of correlation within common response pathways, i.e. pro-inflammation, apoptosis and stress (see list at the end of this section). The degree of correlation c_{sum} between genes within a pathway *pw* at a given infection stage *s* was defined as follows:

$$c_{sum_s}(pw) = \sum_{i,j \in pw,\, i \neq j} -\log(P_{comb_s}(g_i, g_j))$$

A score for changes in the c_{sum} values between different infection stages was defined as follows:

$$\Delta c_{sum_{pw}}(s_1, s_2) = c_{sum_{s_1}}(pw) - c_{sum_{s_2}}(pw)$$

To assess significance of these changes, we compared them against an empirical background distribution based on randomly drawn pseudo-pathways and defined a statistic *t* as detailed in the following pseudo-code:

```
NPseudoPathways = 200
NRuns = 50
empPs = ()

# compare deltaCsum(pw,s1,s2) to values obtained on random pseudo - pathways
for ( n in NRuns ) {
        pps = drawPseudoPathways( NPseudoPathways )
        deltaCsBg = ()
        for ( pp in pps ){
                append( deltaCsBg , deltaCSum(pp,s1,s2) )
        }
        append( empPs , (sum(deltaCsBg >= deltaCsPw) +1) / (NPseudoPathways +1) )
}
mu = mean(empPs)
sigma = sd(empPs)

# define summary statistic t
if ( mu + sigma < 0.05 )
        t = 3
elif ( mu < 0.05 )
        t = 2
elif ( mu - sigma < 0.05 )
        t = 1
else
        t = 0
```

Pseudo-pathways were drawn with gene expression levels and size matching the respective pathway of interest. *t* scores for significant decrease of correlation were computed analogously and assigned values from -3 to 0.

List of gene sets used for pathway correlations.

Pro-inflammation: *IL-8, TNFα, IFNα14, IFNβ, IL-6, HBD3, CCL2, CCL20, CXCL-1, CXCL-2, IL-18, CXCL10, NFKBIA, NFKBIE, NFKB1, NFKB2, c-Jun, COX2*

Apoptosis: *IFNβ, TNFAIP3, Bnip3, RELA, CLARP2, Birc2, Birc3, TNFAIP8, BECN1, Gadd45α, XRCC5, NFκB1, JUNB, CHOP, CYR61*

Stress response genes: *CHOP, ATF3, Xbp1, CDKN1A, IL-8, IL-6, IFNβ, NFKBIE, COX-2, CCNE1, CCNA2, CCNB1*

Supplemental Figures

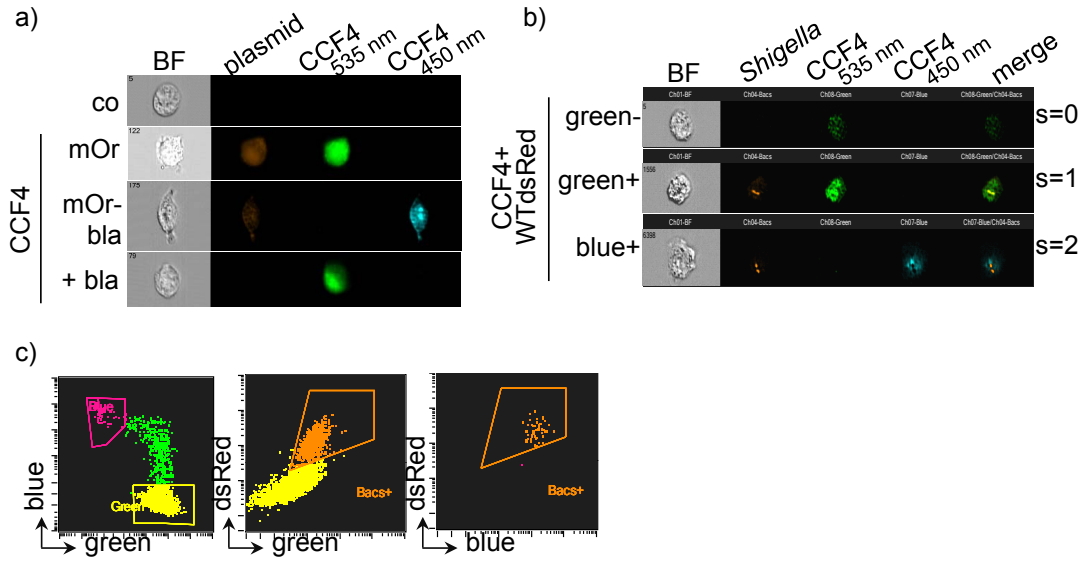
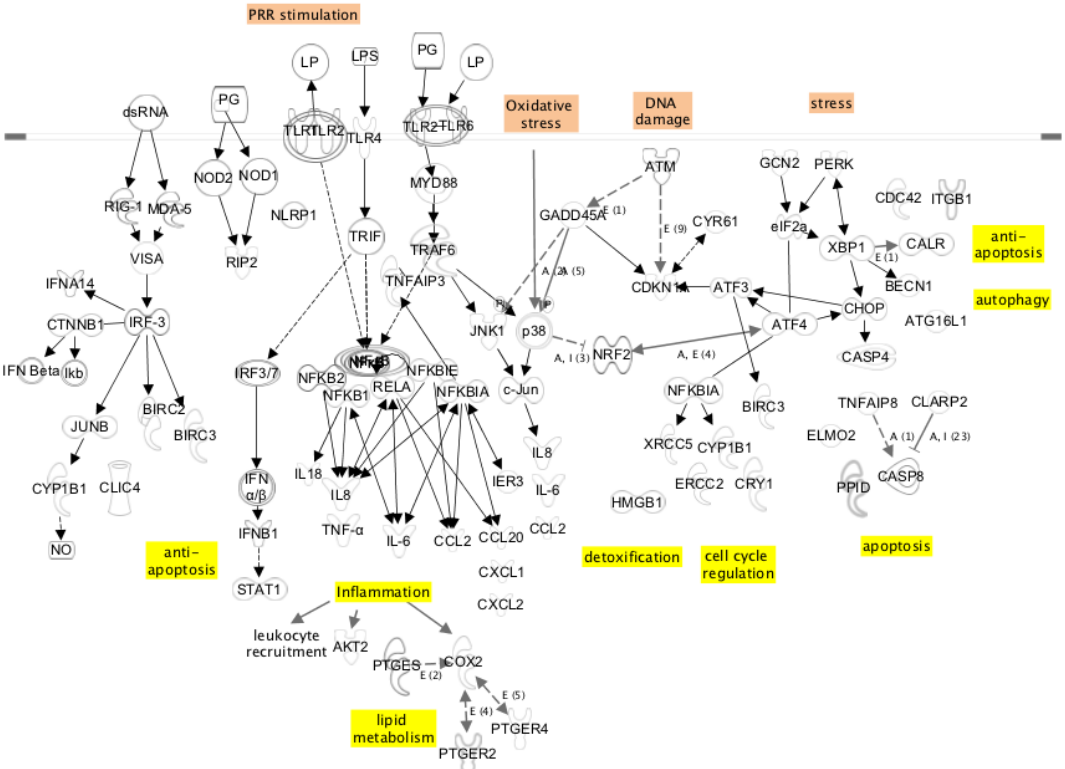


Figure S1 (relates to Fig. 1a) Image stream analysis of different stages of bacterial infection

The compatibility of the CCF4/β-lactamase assay for discrimination and accurate purification of bacterial infection stages via a FACS-based methodology was tested using the Imagestream (ISX). 10,000 HeLa cells were analyzed in the indicated bright field (BF) and fluorescence channels. a) HeLa cells were left untreated, transfected with an empty (mOr) or a β-lactamase- and mOrange-containing (mOr-bla) plasmid to obtain cytosolic expression of β-lactamase or treated with soluble β-lactamase (+bla) to exclude de-FRET upon extracellular presence of β-lactamase. CCF4 was loaded where indicated. b) To discriminate stages of infected cells, CCF4-loaded cells were infected with dsRed-fluorescent WT *Shigella* at a MOI 25 for 1 h. The number of intracellular bacteria (s) was determined using the Amnis software's spot count tool. c) Cells were appropriately gated and analyzed using the ISX dot plot tool.



202

203 **Figure S2 (relates to Fig. 1) Ingenuity pathway network of selected genes**

204 Overview of the molecular signaling pathways linked with the selected genes used for
205 the transcriptional analysis of HeLa cells. The overview was created using the
206 “pathway tool” of the IPA program. Continuous and dashed arrows indicate direct or
207 indirect interactions, respectively. Only interactions leading to transcriptional
208 activation or expression are shown.

209

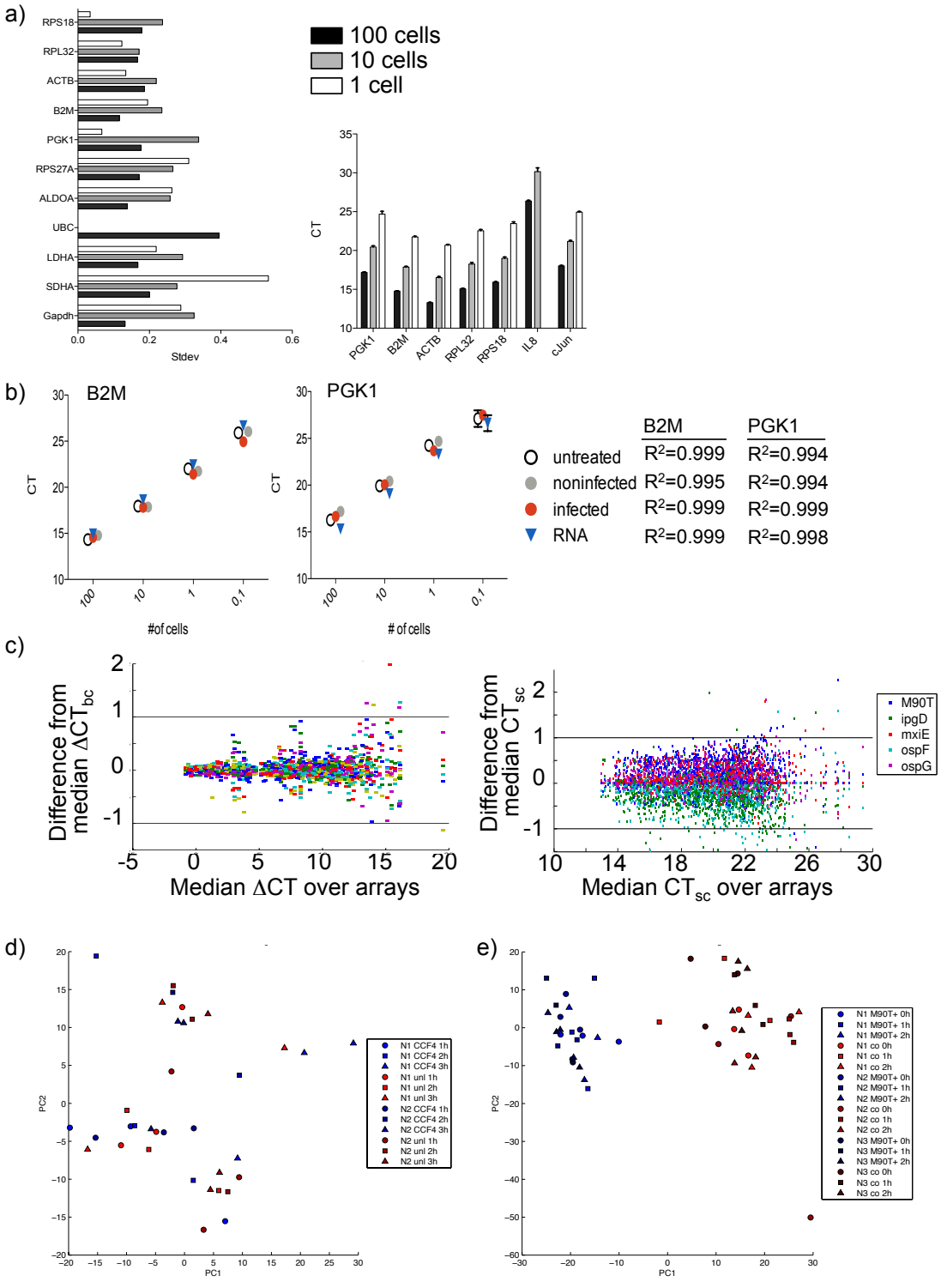


Figure S3 (relates to Fig. 1) Establishment of a robust protocol for multiplex transcriptional analysis of bulk and single cells

a,b) Choice of two (*B2M* and *PGK1*) out of eleven reference genes (see also supplemental results text). 150 ng RNA of either sorted untreated, uninfected or WT

Shigella-infected cells or of commercial HeLa RNA were reverse-transcribed and pre-amplified with 18 cycles. cDNA was titrated to obtain concentrations equivalent to 100, 10, 1 and 0.1 cell(s) per volume of cDNA (2.9 μl) and applied to the dynamic array (Biomark). Samples were run in triplicates and obtained CT values were quality-controlled as described in the experimental procedures. a) The variance of eleven potential reference genes was assessed by plotting the standard deviations (Stdev) of mean CT values (left). The five genes that showed the least variance were further tested for suitable expression levels in range with weakly (*IL-8*) and highly (*c-Jun*) expressed genes from our panel by plotting the CT + Stdev values of uninfected samples from indicated titrations (right). b) To investigate correlation coefficients (R^2), the CT values for *B2M* and *PGK1* from indicated samples were plotted against the equivalent number of cells.

c) Since transcriptional analysis was performed on multiple arrays within one experiment, control samples from the same experiment were repeatedly re-run on all arrays in order to assess their comparability across arrays. Therefore, the difference of normalized ΔCT_{bc} values (left) or of non-normalized CT_{sc} values (right) obtained from all control samples was plotted against the corresponding median expression over the arrays. While the plot shows only expression of control samples, arrays are labeled according to the *Shigella* strain used for infection of their other samples, i.e. as M90T WT, *ΔmxiE, ΔipgD, ΔospF, ΔospG*.

d) CCF4 loading has no global impact on gene expression profiles. PCA of unloaded (unl, red symbols) and CCF4-loaded (CCF4, blue symbols) cells incubated for further 1 h (circles), 2 h (squares) and 3 h (triangles) at 37°C after removal of CCF4 from the cells and washing. Samples of two independent experiments (N1, N2) were run in triplicates.

e) Temperature shift to and further incubation of cells at 4°C stalled transcription. Samples were sorted immediately (+0 h circles), or kept on ice for further 1 h (squares) or 2 h (triangles) before sort. PCA shows transcriptional information of untreated (co, red symbols) and WT *Shigella*-infected (M90T, blue symbols) cells obtained from three independent experiments (N1, N2, N3) run in triplicates.

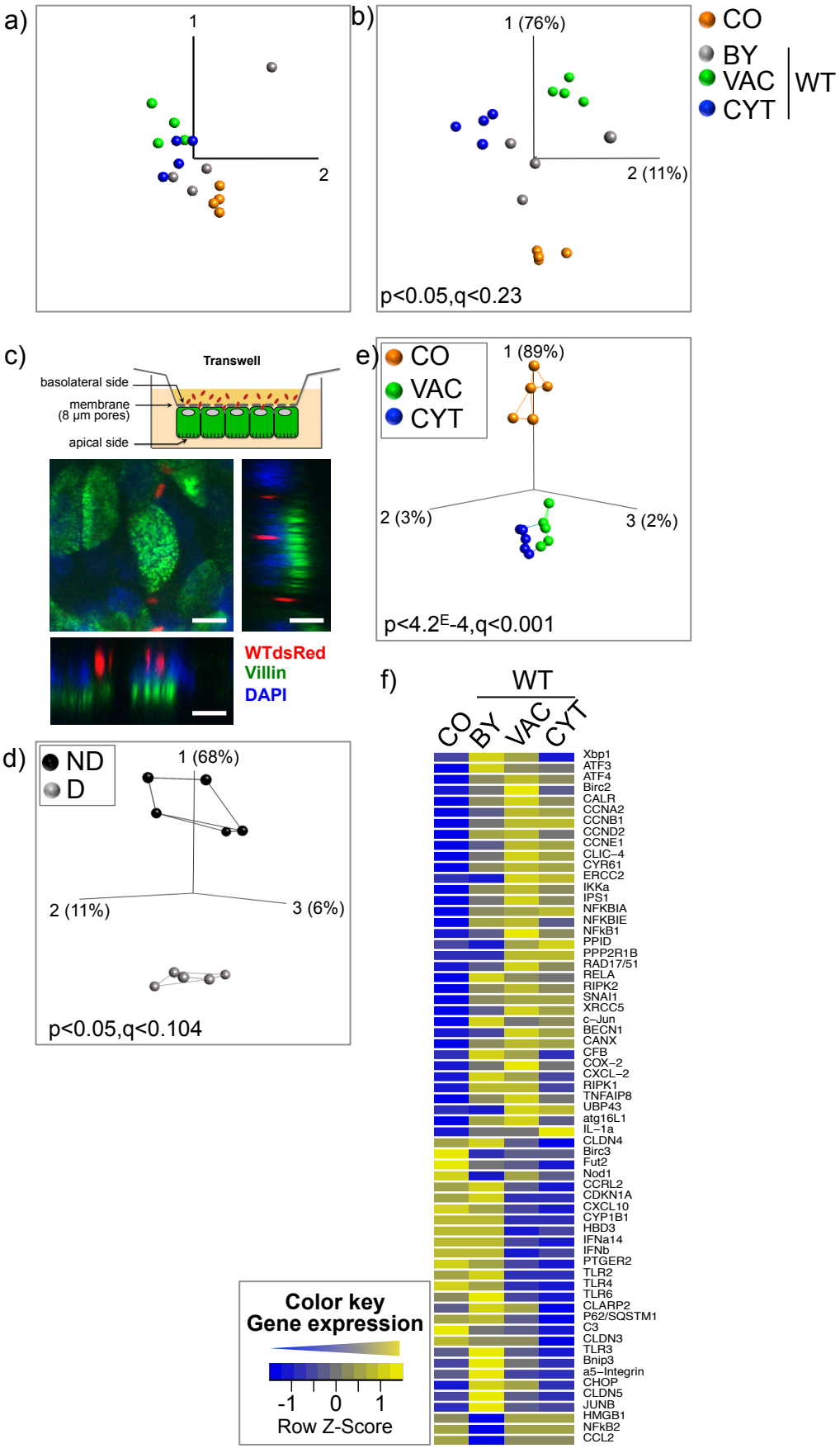

a)
1
2
b)
1 (76%)
2 (11%)
p<0.05,q<0.23
CO
BY
VAC
CYT
WT
c)
Transwell
basolateral side
membrane
(8 µm pores)
apical side
WTdsRed
Villin
DAPI
d)
ND
D
1 (68%)
2 (11%)
3 (6%)
p<0.05,q<0.104
e)
CO
VAC
CYT
1 (89%)
2 (3%)
3 (2%)
p<4.2E-4,q<0.001
f)
WT
CO
BY
VAC
CYT
Xbp1
ATF3
ATF4
Birc2
CALR
CCNA2
CCNB1
CCND2
CCNE1
CLIC–4
CYR61
ERCC2
IKKa
IPS1
NFKBIA
NFKBIE
NFkB1
PPID
PPP2R1B
RAD17/51
RELA
RIPK2
SNAI1
XRCC5
c–Jun
BECN1
CANX
CFB
COX–2
CXCL–2
RIPK1
TNFAIP8
UBP43
atg16L1
IL–1a
CLDN4
Birc3
Fut2
Nod1
CCRL2
CDKN1A
CXCL10
CYP1B1
HBD3
IFNa14
IFNb
PTGER2
TLR2
TLR4
TLR6
CLARP2
P62/SQSTM1
C3
CLDN3
TLR3
Bnip3
a5–Integrin
CHOP
CLDN5
JUNB
HMGB1
NFkB2
CCL2
Color key
Gene expression
-1
0
1
Row Z-Score

Figure S4 (relates to Fig. 1) Localization-dependent transcriptional signatures of WT *Shigella*-infected HeLa and polarized Caco-2 cells

a and b) PCA plots (corresponding to S2 and 3 Movies, respectively) of transcriptional signatures of bulk cells (dots) at indicated infection stages (BY, VAC, CYT) of WT *Shigella* infection and of a non-infected control (CO). No (a) or indicated cutoffs for p- and q-value (b) were applied.

c) Upper panel: Infection model of basolateral bacterial infection of Caco-2 TC7 cells with dsRed-expressing WT *Shigella* using Transwell plates (see details in Suppl. Exp. Procedures). Lower panel: Maximum projection of confocal images of polarized Caco-2 cells infected with dsRed-fluorescent WT *Shigella* at 3 h p.i. showing villi (green) at the apical side and nuclei (blue). Orthogonal (top left), xz (bottom left) and yz (top right) view, scale bar 10 μm.

d,e) PCA plots of bulk cell transcriptional signatures (dots) of non-differentiated (ND) and differentiated (D) Caco-2 cells (d) or of differentiated Caco-2 cells at indicated infection stages (VAC, CYT) of *Shigella* WT-infected and of non-infected control (CO) cells (e). Connections between the two nearest neighbors (lines) and percentages of the captured variability by each PC are shown. Representative data from one out of three independent experiments are shown.

f) Heat map shows median ΔCT values of bulk cell samples of Caco-2 cells transformed into a Row Z-score of indicated genes at indicated conditions of *Shigella* WT infection (BY, VAC, CYT) or uninfected control (CO). All genes with at least one significant gene expression change at one of the conditions are shown. The color key indicates the row Z-scores ranging from +2 (high expression, yellow) to −2 (low expression, blue). Data from two independent experiments at a p-value cutoff of 0.05 are shown.

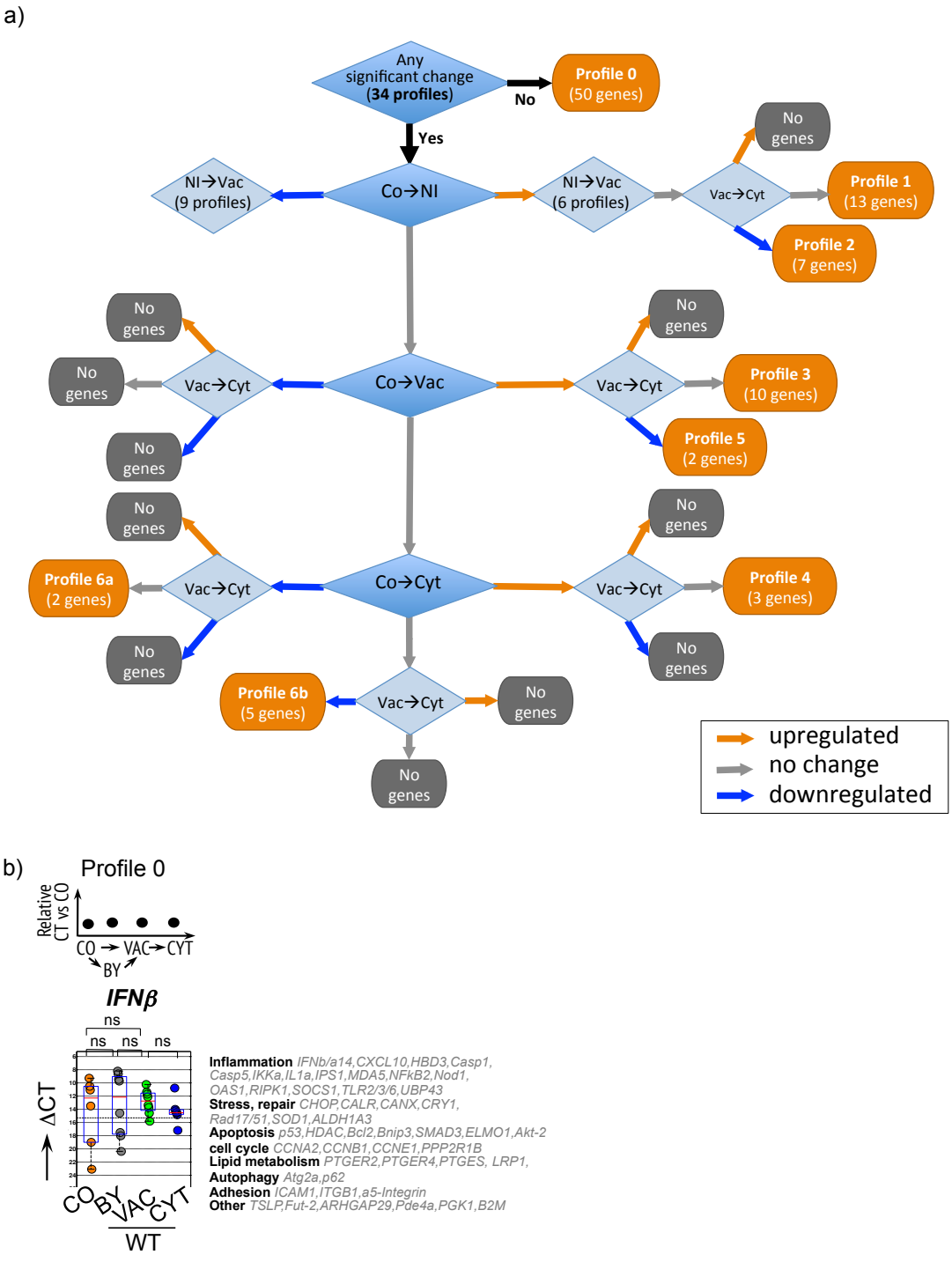


Figure S5 (relates to Fig. 3) Computational procedure for gene expression profiling

a) The flow chart represents the computational analytical procedure used to classify the transcriptional signatures into localization-dependent gene expression profiles (see Fig. 3 upper panels). It uses a decision procedure based on the measured gene

expression changes at the indicated transitions from one stage of bacterial infection to another, i.e. CO→BY→VAC→CYT (represented by blue rectangles). This yields a theoretical maximum of 34 possible profiles (represented by rounded squares or only summarized by numbers in brackets). Decisions were made based on whether a significant increase (orange arrows), decrease (blue arrows) or no significant change (grey arrows) of gene expression was detected. Orange symbols represent profiles to which experimentally identified signatures could be assigned. Profiles 6a and b were pooled, as the outcome (decrease only upon cytosolic localization) is identical. Grey symbols represent profiles to which no signatures could be assigned.

b) Summary of profile 0 to which genes were assigned that showed no significant change at the transition of the indicated stages upon WT *Shigella* infection. The box plot shows ΔCT values of quadruplicates of bulk cell samples from two independent experiments for *IFNβ*. Median expression value (red) as well as the 25% and 75% quartiles (boxes) are shown. Significances: Mann-Whitney U test with *$p<0.05$, **$p<0.005$, ***$p<0.001$, otherwise not significant.

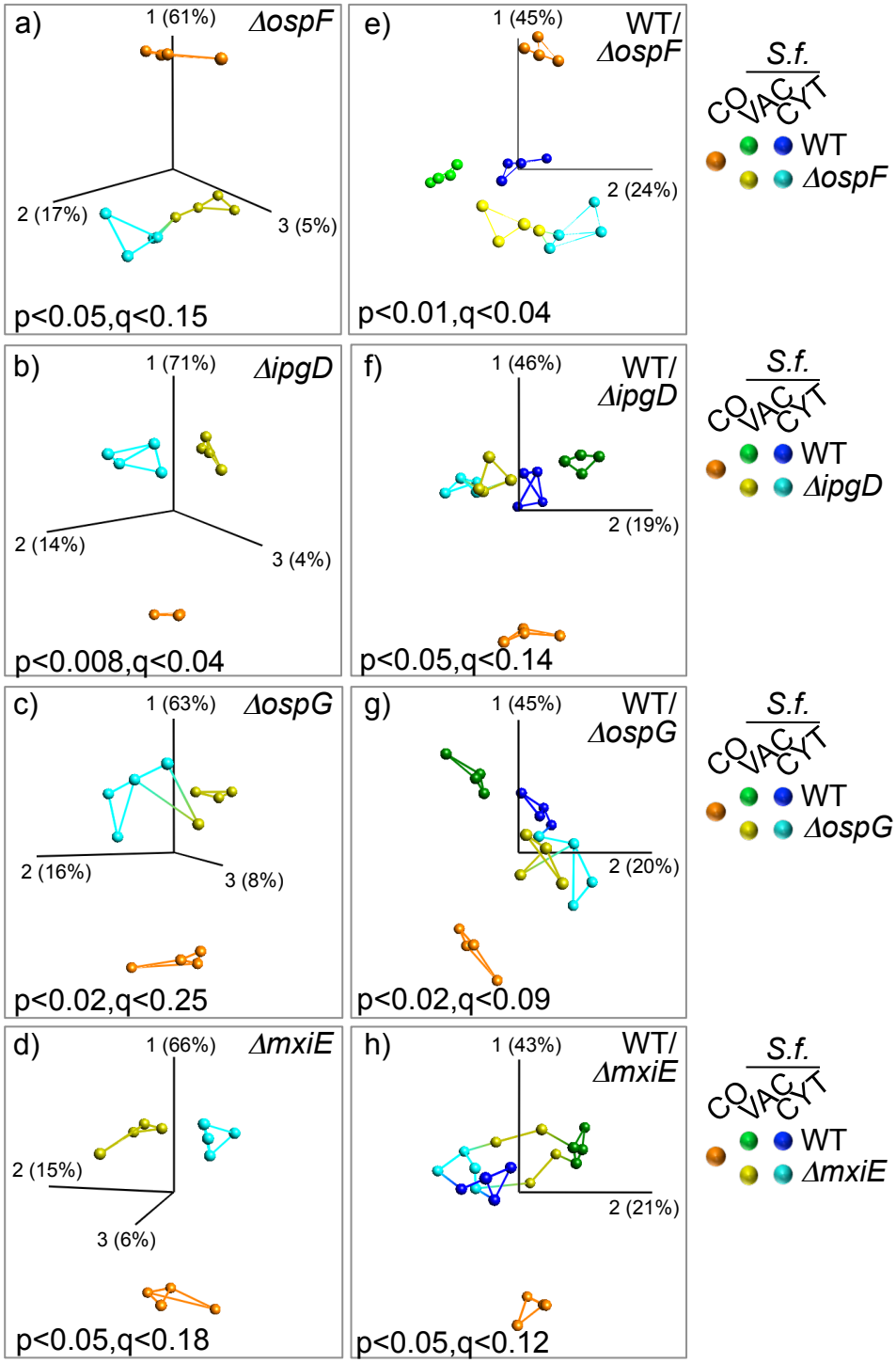


Figure S6 (relates to Fig. 4) Transcriptional signatures of WT- and mutant *Shigella*-infected cells

a-d) PCA plots of bulk cell transcriptional signatures (dots) of non-infected control (CO) cells and of cells at indicated infection stages (VAC, CYT) with a) *ΔospF*, b)

299 *ΔipgD*, c) *ΔospG*, d) *ΔmxiE Shigella*. Cutoffs for p- and q-values and nearest neighbor
300 connections (lines) are shown.
301 e-h) PCA plots show bulk cell transcriptional signatures (dots) of cells infected with
302 WT or with e) *ΔospF*, f) *ΔipgD*, g) *ΔospG*, h) *ΔmxiE Shigella* at the vacuolar or the
303 cytosolic stage (VAC, CYT) and of non-infected control (CO) cells. The two nearest
304 neighbor connections (lines) at indicated cutoffs for p- and q-value are shown. All
305 PCA plots show data from one out of three independent experiments and correspond
306 to Movies S5-8, respectively.
307

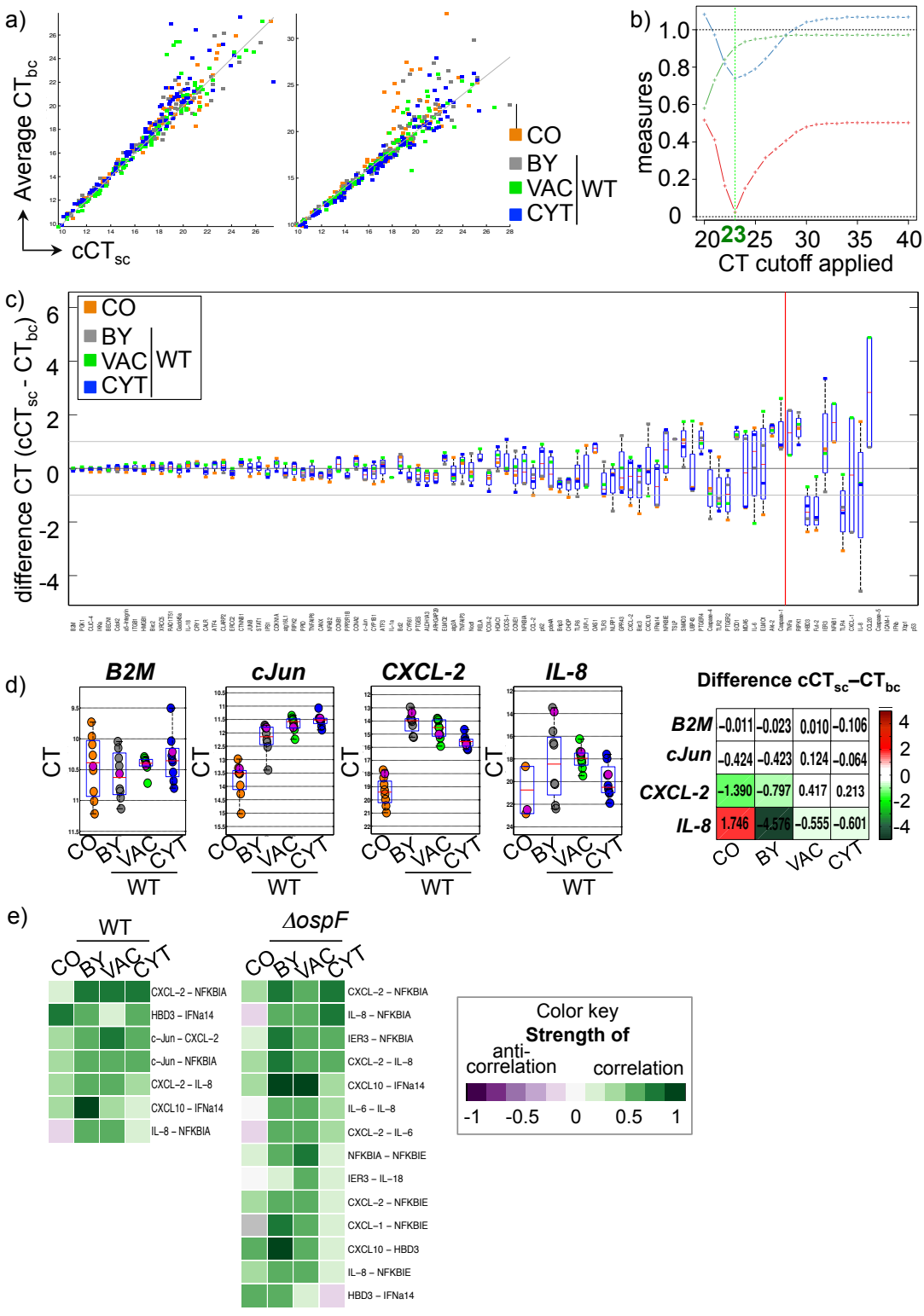


Figure S7 (relates to Fig. 5) Establishment of a robust protocol for single cell transcriptional analysis

a) Scatter plots depict concordance of single and bulk cell measurements of two independent experiments before a CT cutoff has been applied. Cumulative CT_{sc}

(cCTsc) and the average CT_{bc} for indicated stages of bacterial infection or uninfected control cells are shown. Note the high concordance of data with CTs below 23.

b) Definition of the CT cutoff at 23. For a range of CT_{max} cutoffs the absolute mean difference between CT_{bc} and cCT_{sc} (red line), the average RMSD (blue line) and the ratio of retained valid measurements after applying indicated cutoffs (green line) were plotted, which revealed that a CT_{max} of 23 leads to a globally optimal concordance.

c) Box plots show the differences between cCT_{sc} and CT_{bc} on a per-gene basis for all indicated infection stages. The average difference of CTs between single cell and bulk cell measurements is plotted for the 96 measured genes. Grey lines indicate the deviation of +/−1 CT from absolute concordance (black line). The red line indicates for which genes the RMSD between cCT_{sc} and CT_{bc} are above 1.5.

d) Box plots display CT_{bc} values. cCT_{sc} (pink dots) for indicated genes are overlaid on their corresponding bulk cell measurements. The table shows the difference between cCT_{sc} and CT_{bc}. Coloring indicates whether cCT_{sc} over- (green) or underestimates the expression measured in bulk cells (red).

e) Heat maps show Spearman correlation coefficients (positive correlation: green, anti-correlation: purple) between indicated pairs of inflammatory genes. Correlations within uninfected control cells (CO) or within cells at indicated stages of WT- (left) and *ΔospF*-infected cells (right) are shown. All gene pairs with at least one significant correlation ($p_{comb} < 0.01$) within the tested conditions are shown.

Table S1 (relates to Fig. 1) List of gene expression assays used for HeLa and Caco-2 cells with annotations manually assigned based on information retrieved from Pubmed, IPA and the literature

	accession number	Gene name	Annotation 1	Annotation 2	Annotation 3	Annotation 4
1	NM_000291	***PGK1***	reference gene			
2	NM_000994	***B2M***	reference gene			
3	NM_000034	*SDHA*	reference gene			
4	NM_001101	*GAPDH*	reference gene			
5	NM_002046	*LDHA*	reference gene			
6	NM_005566	*ALDOA*	reference gene			
7	NM_002954	*RPS27A*	reference gene			
8	NM_022551	*ACTB*	reference gene			
9	NM_004048	*UBC*	reference gene			
10	NM_004168	*RPL32*	reference gene			
11	NM_021009	*RPS18*	reference gene			
12	NM_003998	***NFkB1***	apoptosis	inflammation	PRR signaling	Transcription factor
13	NM_002229	***JUNB***	apoptosis	cell cycle	Transcription factor	
14	NM_002228	***c-Jun***	inflammation	apoptosis	stress response	Transcription factor
15	NM_002502	***NFkB2***	inflammation	Transcription factor		
16	NM_020529	***NFKBIA***	inflammation	PRR signaling	apoptosis	Transcription factor
17	NM_004556	***NFKBIE***	inflammation	Transcription factor		
18	NM_004083	***CHOP***	ER-stress	UPR	apoptosis	Transcription factor
19	NM_001674	***ATF3***	ER-stress	AA starvation	UPR	Transcription factor
20	NM_182810	***ATF4***	ER-stress	UPR	Transcription factor	
21	NM_005080	***Xbp1***	ER-stress	UPR	Transcription factor	
22	NM_003804	***RIPK1***	PRR signaling	apoptosis	stress response	
23	NM_020746	***IPS1***	PRR signaling	IFN signaling		
24	NM_003821	***RIPK2***	PRR signaling	inflammation		
25	NM_001278	***IKKa***	PRR signaling			
26	NM_001033053	***NLRP1***	PRR signaling	apoptosis		
27	NM_006092	***Nod1***	PRR	PRR cytosol	apoptosis	inflammation
28	NM_022168	***MDA5***	PRR	PRR cytosol	DNA repair	apoptosis
29	NM_003264	***TLR2***	PRR	PRR membrane		
30	NM_138554	***TLR4***	PRR	PRR membrane		
31	NM_003265	***TLR3***	PRR	PRR membrane		
32	NM_006068	***TLR6***	PRR	PRR membrane		
33	NM_000584	***IL-8***	pro-inflammation	cyto-/chemokines		
34	NM_000594	***TNFa***	pro-inflammation	cyto-/chemokines	apoptosis	

Table S1 (continued)

35	NM_002172	***IFNa14***	pro-inflammation	cyto-/chemokines		
36	NM_002176	***IFNb***	pro-inflammation	cyto-/chemokines	apoptosis	
37	NM_000600	***IL-6***	pro-inflammation	cyto-/chemokines		
38	NM_018661	***HBD3***	pro-inflammation	cyto-/chemokines	AMP	
39	NM_002982	***CCL2***	pro-inflammation	cyto-/chemokines		
40	NM_001130046	***CCL20***	pro-inflammation	cyto-/chemokines		
41	NM_001511	***CXCL-1***	pro-inflammation	cyto-/chemokines		
42	NM_002089	***CXCL-2***	pro-inflammation	cyto-/chemokines		
43	NM_000575	***IL-1a***	pro-inflammation	cyto-/chemokines	cell cycle	apoptosis
44	NM_001562	***IL-18***	pro-inflammation	cyto-/chemokines		
45	NM_001565	***CXCL10***	pro-inflammation	cyto-/chemokines		
46	NM_000963	***COX-2***	inflammation	lipid metabolism	apoptosis	
47	NM_138937	***RegIIIbeta***	inflammation	adhesion	acute phase response	
48	NM_002128	***HMGB1***	inflammation	apoptosis	DNA repair	
49	NM_0033292	***Caspase-1***	inflammation	pyroptosis	inflammasome act.	
50	NM_001225	***Caspase-4***	inflammation	pyroptosis	inflammasome act.	ER-stress
51	NM_004347	***Caspase-5***	inflammation	pyroptosis	inflammasome act.	
52	NM_006290	***TNFAIP3***	anti-inflammation	apoptosis	UPR	PRR signaling
53	NM_139266	***STAT1***	IFN signaling	apoptosis		
54	NM_003745	***SOCS-1***	IFN signaling	metabolism		
55	NM_017414	***UBP43***	IFN signaling			
56	NM_002534	***OAS1***	IFN signaling			
57	NM_005038	***PPID***	heast shock resp.	protein folding		
58	NM_004343	***CALR***	cytoskeleton	ER-stress	cell cycle	
59	NM_001746	***CANX***	cytoskeleton	ER-stress	protein transport	
60	NM_000400	***ERCC2***	DNA repair	cell cycle	apoptosis	
61	NM_002873	***RAD17/51***	DNA repair	cell cycle		
62	NM_004075	***CRY1***	DNA repair			
63	NM_000693	***ALDH1A3***	detoxification	apoptosis		
64	NM_000499	***CYP1A1***	detoxification	cell cycle	S phase reg	
65	NM_000104	***CYP1B1***	detoxification			
66	NM_000454	***SOD1***	detoxification			
67	NM_013943	***CLIC-4***	cell cycle	cytoskeleton	IFN signaling	
68	NM_001904	***CTNNB1***	cytoskeleton	cell cycle		
69	NM_001791	***Cdc42***	cytoskeleton			
70	NM_078467	***CDKN1A***	cell cycle	cell cycle progression	apoptosis	
71	NM_181699	***PPP2R1B***	cell cycle	cell cycle progression		
72	NM_001238	***CCNE1***	cell cycle	G1/S transition, S	DNA repair	
73	NM_031966	***CCNB1***	cell cycle	G2/M transition, M		
74	NM_001237	***CCNA2***	cell cycle	G2/M transition, S/G2		
75	NM_017974	***Atg16L1***	autophagy	starvation	stress response	
76	NM_015104	***Atg2A***	autophagy			

Table S1 (continued)

77	NM_004964	***HDAC1***	apoptosis	anti-inflammation	cell cycle
78	NM_001924	***Gadd45a***	apoptosis	cell cycle	G2/M transition
79	NM_005902	***SMAD3***	apoptosis	cell cycle	
80	NM_000546	***p53***	apoptosis	cell cycle	DNA repair
81	NM_133171	***ELMO2***	apoptosis	cytoskeleton	endocytosis
82	NM_014800	***ELMO1***	apoptosis	cytoskeleton	endocytosis
83	NM_021141	***XRCC5***	DNA repair	apoptosis	
84	NM_003897	***IER3***	apoptosis	immed. early response	
85	NM_001626	***Akt-2***	apoptosis	lipid metabolism	
86	NM_004052	***Bnip3***	apoptosis		
87	NM_021975	***RELA***	apoptosis		
88	NM_005306	***GPR43***	inflammation	lipid metabolism	
89	NM_014350	***TNFAIP8***	apoptosis		
90	NM_003766	***BECN1***	anti-apoptosis	autophagy	stress response
91	NM_000633	***Bcl2***	anti-apoptosis	cell cycle	G1/S transition
92	NM_003879	***CLARP2***	anti-apoptosis		
93	NM_001166	***Birc2***	anti-apoptosis		
94	NM_001165	***Birc3***	anti-apoptosis		
95	NM_001554	***CYR61***	apoptosis	adhesion	
96	NM_033668	***ITGB1***	adhesion	cell cycle	G1/S transition, S
97	NM_002205	***a5-Integrin***	adhesion	cell surface signaling	
98	NM_000201	***ICAM-1***	adhesion		
99	NM_000956	***PTGER2***	lipid metabolism		
100	NM_000958	***PTGER4***	lipid metabolism		
101	NM_004878	***PTGES***	lipid metabolism		
102	NM_002332	***LRP-1***	lipid metabolism		
103	NM_000511	***Fut-2***	metabolism	anti-apoptosis	
104	NM_006202	***Pde4A***	membrane ruffle		
105	NM_004815	***ARHGAP29***	GTPase-act. protein		
106	NM_138551	***TSLP***	commensal interaction		
Additional gene expression assays used for analysis of Caco-2 cells					
107	NM_000291	***PGK1***	reference gene		
108	NM_005566	***ALDOA***	reference gene		
109	NM_000619	***IFNg***	pro-inflammation	cyto-/chemokines	
110	NM_000882	***IL12A***	pro-inflammation	cyto-/chemokines	
111	NM_000585	***IL-15 v1/2***	pro-inflammation	cyto-/chemokines	
112	NM_172138	***IL28A***	pro-inflammation	cyto-/chemokines	
113	NM_172140	***IL29***	pro-inflammation	cyto-/chemokines	
114	NM_080389	***DEFB4A***	pro-inflammation	cyto-/chemokines	
115	NM_002982	***CCL2***	pro-inflammation	cyto-/chemokines	
116	NM_003965	***CCRL2***	pro-inflammation		
117	NM_000064	***C3***	pro-inflammation	complement	

Table S1 (continued)

118	NM_001710	***CFB***	pro-inflammation	complement
119	NM_001879	***MASP1***	pro-inflammation	
120	NM_000625	***NOS2***	antimicrobial activities	
121	NM_007052	***Nox1***	pro-inflammation	
122	NM_016931	***Nox4***	pro-inflammation	
123	NM_001041	***SI***	differentiation marker	
124	NM_005985	***SNAI1***	cell-cell junction regulation	
125	NM_020384	***CLDN2***	tight junction	
126	NM_001306	***CLDN3***	tight junction	
127	NM_001305	***CLDN4***	tight junction	
128	NM_003277	***CLDN5***	tight junction	
129	NM_001307	***CLDN7***	tight junction	
130	NM_006984	***CLDN10***	tight junction	

335
336

Table S2 (relates to Fig. 1c) Raw data: ΔCT values calculated from median average CT values and standard deviations from the 42 gene expression signatures which were identified to be significantly differentially regulated upon at least one of the indicated stages of WT *Shigella* infection of HeLa cells

	dCT values (median)				standard deviation			
	CO	BY	VAC	CYT	CO	BY	VAC	CYT
NFKBIA	6.91	3.84	3.66	4.22	0.39	0.69	0.21	0.19
CXCL-2	8.63	2.97	3.60	4.53	1.29	1.04	0.79	0.58
IL-8	12.75	7.25	7.02	9.45	2.95	3.54	1.05	1.66
NFKBIE	8.71	7.80	7.84	8.36	0.69	0.30	0.91	0.62
NFkB1	10.30	9.25	9.20	9.50	1.06	0.46	0.10	0.53
TLR4	9.91	7.49	8.26	10.14	4.13	3.92	1.13	3.40
IL-6	12.02	10.03	10.49	9.93	0.89	1.41	0.81	1.08
JUNB	2.93	1.19	1.84	2.29	0.64	0.65	0.43	0.34
ERCC2	6.12	5.76	5.90	5.86	0.42	0.37	0.44	0.56
IER3	9.97	8.04	8.72	9.42	1.64	1.19	0.37	0.68
ATF4	1.81	1.35	1.57	1.72	0.26	0.33	0.32	0.24
CDKN1A	4.81	4.11	4.47	4.75	0.60	0.48	0.41	0.53
COX-2	7.83	6.04	7.20	7.82	0.66	0.73	0.70	0.49
GPR43	8.37	7.98	12.14	11.85	2.07	3.62	3.53	3.36
IL-18	2.05	1.80	1.61	1.56	0.23	0.40	0.22	0.32
ATF3	3.28	2.44	2.12	2.19	0.58	0.57	0.17	0.27
Birc2	5.66	5.41	5.26	5.19	0.16	0.54	0.30	0.29
Birc3	8.45	6.44	5.70	6.04	1.21	1.21	0.67	0.34
CYR61	2.77	1.64	0.77	0.74	0.89	0.80	0.80	0.41
TNFAIP3	5.86	2.16	1.55	2.05	0.46	1.00	0.47	0.32
CCL-2	6.63	5.63	5.33	5.76	0.78	0.99	0.75	0.93
HDAC1	7.27	7.04	6.89	7.00	0.54	0.31	0.43	0.33
PPID	3.65	3.44	3.23	3.39	0.25	0.25	0.26	0.21
TNFAIP8	2.57	1.83	1.56	1.91	0.29	0.45	0.37	0.30
CLIC-4	0.77	0.59	0.46	0.57	0.17	0.39	0.18	0.10
HMGB1	0.67	0.61	0.39	0.69	0.23	0.14	0.18	0.20
CYP1B1	3.40	3.37	3.20	3.93	0.57	0.22	0.32	0.43
Xbp1	14.06	14.03	13.68	15.30	1.04	0.86	0.81	1.29
STAT1	5.71	5.66	5.49	5.89	0.28	0.36	0.23	0.21
c-Jun	3.04	0.92	0.78	0.58	0.55	0.84	0.22	0.30
RELA	8.44	7.90	7.97	7.79	0.37	0.81	0.36	0.56
RIPK2	5.80	5.25	5.20	5.09	0.22	0.51	0.27	0.25
BECN1	3.64	3.48	3.36	3.25	0.35	0.31	0.25	0.15
Cdc42	1.00	0.87	0.82	0.70	0.24	0.29	0.17	0.10
Gadd45a	6.13	4.83	4.15	3.52	0.46	0.78	1.00	0.93
NLRP1	8.88	9.59	9.64	9.89	0.82	1.05	1.03	1.43
CTNNB1	3.62	3.74	3.71	3.93	0.25	0.29	0.20	0.13
Caspase-4	8.17	8.22	8.41	7.82	0.58	1.15	0.80	0.44
XRCC5	0.10	0.00	-0.28	-0.18	0.25	0.43	0.28	0.13
ELMO2	7.26	7.33	6.63	6.65	0.18	0.64	0.31	0.27
CLARP2	6.13	6.12	5.67	5.93	0.28	0.68	0.50	0.31

Table S3 (relates to Fig. S4f) Raw data: ΔCT values calculated from median average CT values and standard deviations from gene expression signatures which were identified to be significantly differentially regulated upon at least one of the indicated stages of WT *Shigella*-activated gene expression signatures of Caco-2 cells

	dCT (median)				standard deviation			
	CO	BY	VAC	CYT	CO	BY	VAC	CYT
Xbp1	8.51	7.96	8.09	9.37	1.01	1.24	1.06	0.70
ATF3	1.09	0.05	0.72	0.45	1.19	1.19	0.88	1.17
ATF4	0.16	-0.18	-0.51	-0.84	0.54	0.55	0.49	0.47
Birc2	1.05	0.75	0.91	0.47	0.57	0.80	0.78	0.69
CALR	-0.54	0.00	-0.96	-1.20	0.50	1.45	1.02	0.86
CCNA2	2.91	1.69	0.82	0.48	1.69	1.92	1.14	1.02
CCNB1	1.69	0.92	-0.12	-1.39	1.68	4.54	1.86	1.62
CCND2	7.18	7.31	6.81	6.61	1.54	2.25	1.03	1.15
CCNE1	8.89	8.96	7.48	7.33	1.84	2.48	1.54	0.93
CLIC-4	-0.04	-0.13	-0.50	-1.21	0.76	1.12	0.78	0.70
CYR61	2.61	3.19	0.45	-0.04	1.04	1.14	0.81	0.67
ERCC2	5.14	5.26	4.20	4.23	1.65	0.88	0.46	0.71
IKKa	4.33	4.62	3.41	3.48	0.71	0.59	0.52	0.69
IPS1	2.70	3.18	2.00	1.97	0.85	1.00	0.56	0.38
NFKBIA	2.77	2.85	2.07	1.54	0.82	0.63	0.54	0.60
NFKBIE	4.01	3.35	2.94	3.25	0.93	0.68	0.68	0.52
NFkB1	7.40	8.30	6.57	6.74	0.98	1.52	0.87	0.68
PPID	2.66	4.68	3.01	3.21	0.98	2.61	1.62	1.65
PPP2R1B	3.19	7.21	3.94	3.41	0.59	2.50	1.86	1.94
RAD17/51	0.96	1.45	1.00	0.70	0.59	0.92	0.60	0.56
RELA	6.96	6.30	6.00	5.92	1.71	0.88	0.78	0.67
RIPK2	2.82	2.55	2.14	1.64	0.74	0.52	0.64	0.49
SNAI1	7.22	5.82	5.73	5.32	1.23	2.24	0.83	0.67
XRCC5	-0.08	-0.41	-1.38	-1.80	0.68	0.82	0.56	0.43
c-Jun	-2.29	-3.30	-2.91	-3.38	0.81	0.67	0.67	0.39
BECN1	1.17	1.68	1.56	0.96	1.03	1.58	1.13	0.77
CANX	-1.35	-0.64	-1.79	-1.88	1.09	1.19	1.00	1.18
CFB	4.42	4.77	4.23	4.53	0.84	1.11	0.99	0.77
COX-2	5.90	7.50	5.53	5.02	2.12	2.00	1.28	0.74
CXCL-2	1.64	1.27	0.83	1.87	1.76	1.05	1.18	0.88
RIPK1	4.11	3.75	3.80	3.60	0.79	0.56	0.53	0.50
TNFAIP8	0.47	0.90	0.34	0.22	0.53	1.08	0.56	0.46
UBP43	3.68	5.13	3.78	3.74	0.88	3.64	1.87	1.35
atg16L1	3.17	3.87	3.28	2.94	0.63	1.59	0.85	0.79
IL-1a	8.09	7.24	7.71	7.75	1.94	2.84	1.25	0.81
CLDN4	2.66	4.45	4.66	4.80	0.88	2.69	1.77	1.82
Birc3	3.17	5.51	5.42	5.04	1.38	5.16	3.74	3.31
Fut2	3.78	5.17	5.74	6.63	2.14	5.29	5.30	4.13
Nod1	8.16	14.63	8.62	9.23	1.06	4.38	4.20	2.78
CCRL2	4.82	6.29	7.23	7.61	2.07	2.04	2.65	2.51
CDKN1A	4.16	3.35	5.05	5.19	1.24	5.34	3.94	3.62
CXCL10	8.38	10.05	12.79	12.98	3.38	6.14	4.22	4.54

Table S3 (continued)

CYP1B1	4.81	8.32	9.03	7.21	1.74	4.52	4.10	2.96
HBD3	4.65	6.20	9.86	7.67	1.52	3.85	3.43	2.50
IFNa14	5.26	6.62	9.09	11.06	1.23	4.14	3.87	3.31
IFNb	9.70	10.63	18.76	13.97	2.98	3.38	5.01	4.48
PTGER2	6.41	13.12	13.52	14.21	1.69	2.77	1.10	1.18
TLR2	3.63	4.01	6.90	5.41	1.19	2.53	1.74	1.95
TLR4	3.89	6.30	9.31	7.13	1.17	2.97	2.96	2.85
TLR6	5.73	6.58	8.48	8.70	2.13	1.68	2.46	2.71
CLARP2	1.51	1.43	2.02	2.34	1.18	0.62	0.87	0.36
p62	7.74	7.51	7.95	8.76	1.04	0.60	0.25	0.42
C3	9.91	14.15	11.65	13.07	1.74	3.57	3.52	3.33
CLDN3	5.80	6.74	7.03	7.14	2.17	2.27	1.66	1.13
TLR3	6.37	6.07	7.57	9.52	2.57	1.35	1.61	3.22
Bnip3	1.25	2.40	2.87	2.47	0.78	0.98	0.65	0.67
a5-Integrin	2.97	2.93	4.10	3.58	1.02	0.74	0.59	0.35
CHOP	-0.03	-0.40	0.44	0.53	1.42	0.86	1.56	0.69
CLDN5	8.29	8.27	9.17	9.34	2.16	1.82	1.05	3.30
JUNB	1.89	1.58	2.08	2.07	0.75	0.81	0.55	0.63
HMGB1	-0.04	2.31	0.97	0.91	2.49	3.44	1.87	1.66
NFkB2	2.41	3.48	2.16	2.42	0.61	3.63	2.52	1.78
CCL2	2.96	8.27	3.60	3.39	2.41	6.94	4.06	2.75

Table S4 (relates to Fig. 2a-d) Raw data: ΔCT values calculated from median average CT values and standard deviations of activated (p≤0.05, q≤0.25) gene expression signatures in bystander HeLa cells upon challenge with WT or different mutant strains of *Shigella*.

	dCT (median)						standard deviation					
	CO	BY WT	BY ospF	BY ospG	BY mxiE	BY ipgD	CO	BY WT	BY ospF	BY ospG	BY mxiE	BY ipgD
IL-6	11.94	10.03	9.69	11.27	9.98	10.58	0.97	1.41	2.18	1.82	1.09	2.55
IL-8	12.60	7.25	5.47	5.42	7.08	6.93	2.62	3.54	2.42	2.40	2.35	3.41
NFKBIE	8.79	7.80	7.91	8.03	7.78	8.29	0.62	0.30	0.69	0.44	0.39	0.69
CXCL-2	8.70	2.97	2.85	4.01	3.80	3.84	1.18	1.04	1.62	1.69	1.05	1.88
COX-2	7.76	6.04	5.97	6.80	5.92	6.38	0.61	0.73	1.05	0.96	0.93	1.07
NFKBIA	6.92	3.84	3.98	4.25	3.97	3.81	0.38	0.69	1.08	1.38	0.68	0.82
TNFAIP3	6.15	2.16	2.06	2.59	2.00	2.31	0.46	1.00	1.34	2.07	0.60	1.30
RIPK2	5.74	5.25	5.41	5.39	5.29	5.44	0.20	0.51	0.33	0.39	0.28	0.20
c-Jun	3.05	0.92	0.80	1.06	1.40	1.02	0.51	0.84	0.74	0.55	0.78	0.73
IER3	9.86	8.04	8.58	8.27	7.81	8.13	1.39	1.19	1.51	1.59	1.35	1.40
Birc3	5.71	5.41	5.30	5.45	5.37	5.38	0.23	0.54	0.39	0.34	0.34	0.48
Gadd45a	6.11	4.83	4.98	5.02	4.98	4.58	0.44	0.78	1.25	0.77	0.93	1.04
JUNB	2.90	1.19	1.19	1.65	1.56	1.57	0.67	0.65	0.94	1.07	0.85	0.87
TNFAIP8	2.57	1.83	1.95	2.13	1.81	1.94	0.30	0.45	0.41	0.30	0.20	0.40
CYR61	2.50	1.64	0.38	2.17	1.23	0.52	0.74	0.80	0.95	1.49	1.54	1.18
ERCC2	6.14	5.76	6.59	6.23	6.14	5.93	0.39	0.37	0.47	0.50	0.35	0.45
CDKN1A	4.96	4.11	4.52	4.23	4.56	4.55	0.56	0.48	0.75	0.95	0.60	0.51
ATF3	3.09	2.44	2.07	2.21	2.43	2.24	0.46	0.57	0.31	0.43	0.18	0.51
ATF4	1.88	1.35	1.53	1.63	1.30	1.50	0.28	0.33	0.28	0.35	0.41	0.36
CLIC-4	0.82	0.59	0.44	0.51	0.60	0.65	0.14	0.39	0.22	0.27	0.18	0.18

Movie S1 (relates to Fig. 1b) shows WT-induced transcriptional signatures at p≤0.005

Movie S2 (relates to Fig. S4a) shows WT-induced transcriptional signatures, no statistical cutoffs applied

Movie S3 (relates to Fig. S4b) shows WT-induced transcriptional signatures at p≤0.05

Movie S4 (relates to Fig. 2e, g and h) shows occurrence (top) or absence (middle and bottom) of p65 nuclear translocation after WT *Shigella* entry (top), upon cytochalasin D treatment prior and during WT *Shigella* challenge (middle) or upon *ΔmxiD Shigella* challenge (bottom).

Movie S5 (relates to Fig. S6e) shows WT- and *ΔospF* induced intracellular transcriptional signatures

Movie S6 (relates to Fig. S6f) shows WT- and *ΔipgD* induced intracellular transcriptional signatures

Movie S7 (**relates to Fig. S6g)** shows WT- and *ΔospG* induced intracellular transcriptional signatures

Movie S8 (relates to Fig. S6h) shows WT- and *ΔmxiE* induced intracellular transcriptional signatures

Supplemental References

1. **Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT.** 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments., vol 55, p 611-622.
2. **Andersen CL, Jensen JL, Ørntoft TF.** 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer research **64:**5245-5250.
3. **Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S.** 2006. Stochastic mRNA Synthesis in Mammalian Cells. PLoS Biology **4:**e309.
4. **McDavid A, Finak G, Chattopadyay PK, Dominguez M, Lamoreaux L, Ma SS, Roederer M, Gottardo R.** 2013. Data exploration, quality control and testing in single-cell qPCR-based gene expression experiments. Bioinformatics **29:**461-467.
5. **Wilks S.** 1938. The large-sample distribution of the likelihood ratio for testing composite hypotheses. The Annals of Mathematical Statistics.