Identification of Parameters of Host Cell Vulnerability during Salmonella Infection by Quantitative Image Analysis and Modeling

Jakub Voznica, Christophe Gardella, Ilia Belotserkovsky, Alexandre C Dufour, Jost Enninga, Virginie Stévenin

IAI Accepted Manuscript Posted Online 30 October 2017
Infect. Immun. doi:10.1128/IAI.00644-17

Identifying parameters of host cell vulnerability during *Salmonella* infection by quantitative image analysis and modeling

Jakub Voznica[1,2,3], Christophe Gardella[4,5], Ilia Belotserkovsky[6], Alexandre Dufour[2], Jost Enninga[1], Virginie Stévenin[1*]

[1]Institut Pasteur, Dynamics of Host-Pathogen Interactions Unit, 25 rue du Dr. Roux, 75724 Paris, France

[2]Institut Pasteur, BioImage Analysis Unit, 25 rue du Dr. Roux, 75724 Paris, France

[3]Ecole Normale Supérieure de Cachan, Université Paris-Saclay, 61 avenue du Président Wilson, 94235 Cachan, France

[4]Laboratoire de Physique Statistique, CNRS, UPMC and Ecole Normale Supérieure, 24, rue Lhomond, 75005 Paris, France

[5]Institut de la Vision, INSERM and UPMC, 17 rue Moreau, 75012 Paris, France

[6]Institut Pasteur, Molecular Microbial Pathogenesis Unit, 25 rue du Dr. Roux, 75724 Paris, France

*For correspondence:

Virginie Stévenin virginie.stevenin@ens-cachan.fr

Key words:

Salmonella enterica serovar Typhimurium, cooperative behavior, cell vulnerability, single-cell heterogeneity, mathematical modeling.

Running title:

Cell vulnerability during *Salmonella* infection

ABSTRACT

Salmonella targets and enters epithelial cells at permissive entry sites: some cells are more likely to be infected than others. However, the parameters that lead to host cell heterogeneity are not known. Here, we quantitatively characterized host cell "vulnerability" towards *Salmonella* infection based on imaged parameters. We performed successive infections of the same host cell population followed by automated high-throughput microscopy and observed that infected cells have higher probability of being re-infected. Establishing a predictive model we identified two combined origins of host cell vulnerability: the pathogen-induced cellular vulnerability emerging from *Salmonella* uptake and persisting at later stage of the infection, and the host cell-inherent vulnerability. We linked the host cell inherent vulnerability with its morphological attributes such as the local cell crowding, and with host cell cholesterol content. This showed that the probability of *Salmonella* infection success can be forecast from morphological or molecular host cell parameters.

INTRODUCTION

Salmonella enterica serovar Typhimurium (*S*. Typhimurium) is a Gram-negative bacterium that causes enteric diseases in many vertebrates after ingestion of contaminated food or water. Salmonellosis is one of the most common causes of food-borne diseases in humans and is considered to be major public health and global economic problem (1). After oral uptake, more than 99% of *S*. Typhimurium are killed in the stomach or in the gut (2). The surviving bacteria reach the distal ileum where they invade non-phagocytic intestinal epithelial cells (3). *In vitro* experiments have shown that *S*. Typhimurium invasion of host cells occurs after a phase of bacterial “Near Surface Swimming” (NSS) on the epithelial layer. The bacteria scan the surface and eventually stop and dock at a “selected” host cell (4), (5). Docking is irreversible (6) and followed by injection of *Salmonella* effectors into the host cell through a Type 3 Secretion System (T3SS), leading to the formation of ruffles that engulf the incoming bacterium (7), (8). Upon internalization *S*. Typhimurium either develops inside a *Salmonella*-Containing Vacuole (SCV) or it ruptures the SCV to escape into the cytoplasm where the pathogen replicates at a high rate, a phenomenon called hyperreplication (HR) (9, 10).

The mechanism by which *S*. Typhimurium targets specific host cellular sites for its entry remains debated. Santos and colleagues suggested that mitotic cells are selected due to increased cholesterol accumulation at the cell surface during metaphase (11). By contrast, Misselwitz and colleagues proposed that physical obstacles and forces that occur during the process of NSS lead to the targeting of topologically prominent sites, such as dividing cells or membrane ruffles (4). Finally, Lorkowski and colleagues have reported that the invasion of *S*. Typhimurium at the ruffle site is a highly cooperative effort (12), (6). Indeed, co-infection of WT and non-invasive *S*. Typhimurium mutants resulted in the entry of both strains inside the host cells: non-invasive *S*. Typhimurium mutants are trapped at ruffle sites and concomitantly internalized within the host cell, following the active invasion by WT *S*. Typhimurium. However, the cooperative effect between intracellular and entering bacteria remains poorly understood at later stage of the infection.

An increasing number of studies have highlighted the relevance of intrinsic cellular heterogeneity within eukaryotic monocultures. After seeding, cells display a dynamic range of variability in their morphology depending on their local microenvironment, including the local density, and the peripheral or central positioning within cellular islets (13). This heterogeneity results in differences of transcription (14), (15), lipid composition (15), (13) and sensitivity towards infections (13). Such cell-to-cell variations have been studied during viral infection

93 revealing that simian virus 40 and mouse hepatitis virus present a population-determined
94 pattern of infection associated with differential cell local crowding (13). In the context of
95 bacterial infection, cell targeting has been related to bacterial cooperation at the entry site
96 and evaluated at the whole population level using Colony Forming Unit (CFU) counting or
97 flow cytometry analysis (12), but so far not *in situ* at the single cell level.
98
99 Here we investigated the susceptibility of epithelial host cells within the same cell population
100 to become infected by *S*. Typhimurium. Our analysis revealed that some cells are more likely
101 to be infected by *Salmonella* than others. We termed them “vulnerable cells”. The cell
102 vulnerability was characterized in a quantitative manner by automated high-content imaging
103 through double sequential infections with a delay of 1 to 3 h between the bacterial
104 challenges. The number of intracellular bacteria per cell as well as the corresponding host
105 cell parameters were assessed, such as cell perimeter, local density, and number of infected
106 neighboring cells. Using a mathematical model, we showed that host cell vulnerability can be
107 induced by a first bacterial uptake but also emerged from its intrinsic morphological and
108 micro-environmental characteristics.

RESULTS

Sequential infections allow studies of *Salmonella* cooperation at the single cell level

We carried out a microscopy-based double infection assay to explore possible links between host cell vulnerability and successive bacterial infections of epithelial cells (**Fig.1**). HeLa cells grown in 96-well plates were subjected to a first infection with green *S*. Typhimurium expressing the fluorescent protein GFP (SL_{GFP}) for 30 min followed by elimination of the extracellular bacteria via gentamicin treatment and washing. The cells were then incubated for 1, 2 or 3 h before being subjected to a second wave of infection with red *S*. Typhimurium expressing the fluorescent protein dsRed (SL_{dsRed}). Extracellular bacteria were again eliminated in the same way, and the host cells were stained with CellMask and DAPI before automated image acquisition of the entire culture wells (**Fig.1A**). The obtained images were analyzed with CellProfiler, a widely used image analysis software (16), (17) (**Fig.1B**). The differently labeled bacteria and the stained host cells enabled us to distinguish and quantify distinct cellular populations: those cells infected during the 1st infection (I_1) or not (noI_1), those infected during the 2nd infection (I_2) or not (noI_2), as well as the associated subpopulations ($I_1\&I_2$, $noI_1\&noI_2$, $I_1\&noI_2$ and $noI_1\&I_2$) (**Fig.1C**). We based our analysis on comparing the probabilities of infection in these subpopulations.

Salmonella cooperates for entry at ruffles

In order to test the reliability of our method, we analyzed first if we could detect the ruffle-dependent cooperation between individual salmonellae during host cell entry, previously observed in infected HeLa and MDCK cells (4), (12). To do this we determined first the time window during which ruffle-associated cooperation could potentially occur by performing time-lapse microscopy of *Salmonella* infection of HeLa cells transiently expressing GFP-tagged actin (**Fig. 1D**). Time series of 90 min at 3 min intervals provided image sequences with forming and disappearing ruffles. In most of the cases, we observed the uptake of one to two bacteria per ruffle, and we saw ruffle disappearance in less than 15 min (**Movie.S1**). We noticed that the more bacteria were engulfed by the ruffles, the longer we could detect the presence of these ruffles. Therefore, newly arriving bacteria prompted additional growth of the ruffles (**Movie.S2**). We quantified the ruffle lifetime by measuring the delay of their disappearance after the entry of the last bacterium. The few cases of very high infection (>5 bacteria/ruffle) that could not be properly analyzed were excluded. Quantification revealed an average ruffle lifetime of 13 min and that 90% of the ruffles completely disappeared after 24 min (**Fig.1D**). Labeling Caco-2 cells with the membrane dye FM 4-64, we observed that the ruffle lifetime for infected Caco-2 cells were similar to infected HeLa cells.

We then challenged HeLa and Caco-2 cells with SL_{GFP} and SL_{dsRed} at the same time and compared the probability for SL_{dsRed} to infect the same cell containing simultaneously SL_{GFP} with those that did not contain SL_{GFP} (**Fig.1E**;**F;G**)**;** *see Materials and Methods for details*. The probability of SL_{dsRed} infection was significantly higher in a cell infected by SL_{GFP} than in a cell not infected by SL_{GFP}, both for HeLa (**Fig.1E**) and Caco-2 (**Fig.1F**) cells. The repartition of the different populations of infected cells shows a much larger overlap between the cells co-infected with SL_{GFP} and SL_{dsRed} than one would anticipate theoretically for two independent infections (**Fig.1G**). Thus, the efficiency of *Salmonella* invasion of an individual epithelial cell depends on the concomitant invasion of the same cell by other salmonellae. Interestingly, increasing the multiplicity of infection (MOI) in HeLa cells (**Fig.1E**) resulted in a significant increase of the SL_{dsRed} infection in cells infected by SL_{GFP}, but not in cells not infected by SL_{GFP}. This result confirmed that the direct effect of an MOI increase is a higher number of bacteria that infect certain cells rather than an increase in the overall number of cells that become infected. In addition to the previously reported *Salmonella* cooperative entry in HeLa and MDCK cells (4), (12), we showed here that this cooperation also takes place in Caco-2 cells, suggesting that this phenomenon is universal during *Salmonella* entry in epithelial cells. Taken together, these results validated that our system was operational.

The probability of being re-infected by *Salmonella* is higher for already-infected cells, even after the disappearance of the entry ruffles

To study long-term and ruffle-unrelated cooperative events of *Salmonella* co-infections, we set up the sequential infections with a delay of 1 h between the two infection waves, killing extracellular bacteria in between through gentamicin treatment. Scanning our time-lapse movies, we were ensured that this time lag led to the complete disappearance of any remaining entry ruffles from the first infection. In addition, we extended the delay between the two sequential infections to 2 h and 3 h (*see* **Fig.1A)**. We compared the different populations of cells infected during the 2nd infection (population I_2), depending on whether they were already infected during the first wave of infection (population $I_2 \mid I_1$) or not (population $I_2 \mid noI_1$) for HeLa (**Fig.2A**) and Caco-2 (**Fig.2B**) cells. For both tested cell types, it was significantly more probable for a cell infected the 1st time to be re-infected the 2nd time compared to a cell not previously infected. We propose that such cells are somehow more vulnerable for future infection.

During all sequential infection experiments we also controlled the overall infection efficiencies of SL_{GFP} and SL_{dsRed} at all measured time points (1st: SL_{GFP} - 2nd: SL_{dsRed} or in the reverse order) (**Fig.S1**)**.** In all cases, the percentage of cells infected by each fluorescent *Salmonella* was similar to cells subjected to single (control) or sequential infections, underlining that

sequential infections did not change the overall infection efficiency for the differently colored salmonellae. Nevertheless, we noticed a decrease in the amount of infected cells between the early infection and later time points. This effect is most likely due to the technically unavoidable gentamicin treatment between infections. Besides, SL_{GFP} showed a higher infectivity than SL_{dsRed} for each condition explained by general deleterious effects of the heterologously overexpressed fluorescent proteins on *Salmonella* infectivity, and by the partial loss of dsRed expression observed by us and others. Taking into account these issues, we took advantage of the observed consistency of the differences of infection efficiency between the initial and the successive infections, and between SL_{GFP} and SL_{dsRed}. This consistency allows comparative analyses of the ratio of the different infection probabilities, and it provided us with an analytical tool for precise quantification independently of the variances of the differently colored bacteria and technical hurdles of sequential infection.

We defined a "vulnerability score" as the conditional probability for a cell to be infected during the 2nd infection after it had already been infected during the 1st one ($I_2 \mid I_1$), divided by the conditional probability for a cell to be infected during the 2nd infection when it had not been previously infected ($I_2 \mid noI_1$) (described in details in *Materials and Methods*). We also analyzed the changes of the vulnerability score in time comparing cells subjected to sequential infections with 1, 2 and 3 h delays (**Fig.2B** and **Fig.S2** for detailed representation of the conditional probability for each replicate). Surprisingly, the vulnerability score appeared un-altered. We obtained similar results when reversing the order of the tested pathogens, infecting first with SL_{dsRed} and then with SL_{GFP} (**Fig.S3**). It was not possible to shorten the delay between infections to less than 1 h due to the ruffle influence, and we could not extend it beyond 3 h due to potential release of hyper-replicative (HR) bacteria from the first infection into the extracellular medium that could then re-infect new cells during the 2nd wave of infection. Altogether, these results showed that, after ruffle disappearance, the infected cells remain more vulnerable to a new infection than the non-infected ones, and this vulnerability is stable in time.

Cell vulnerability to secondary infection can be predicted from the number of intracellular bacteria

So far, we only considered the character "infected" or "non-infected" for each cell after SL_{GFP} and SL_{dsRed} infections that provides global trends on their interaction. To further exploit our data we quantified the number of bacteria per host cell and related the obtained numbers with the previously extracted vulnerability scores. The distribution of intracellular bacteria

inside infected cells at 2.5 h post-infection (pi) showed that most of the cells contained a few bacteria, and the proportion of cells decreased drastically when the number of intracellular bacteria increases. Overall, we were able to distinguish three groups of infected cells: the ones containing one to two intracellular bacteria (35% of the global population), the ones containing three to eight intracellular bacteria (39% of the global population) and the ones containing more than nine intracellular bacteria (26% of the global population), corresponding respectively to low, medium and high infections (**Fig.3A)**.

We compared the vulnerability score of these three infection groups during sequential double infections (**Fig.3B**). This analysis revealed that the more bacteria had entered in a given host cell during the first infection, the more it was likely that this cell became re-infected. Such tendencies still emerged when the bacteria were not grouped, but analyzed individually, underlining the robustness of this result (**Fig.S4**).

Then, we investigated how the level of bacterial uptake during the second infection depends on the number of intracellular bacteria of the first infection. For this we quantified the probability for a cell to be highly infected during the second infection as a function of the efficiency of the first uptake (**Fig.3C**). We found that the more intracellular bacteria had been internalized during the first infection, the more likely the host cells were to engulf a high amount of new bacteria during the second infection. Therefore, we propose that cell vulnerability is maintained from the first to the second infection.

Cell vulnerability as intrinsic or induced property

The results from the sequential infections (**Fig.2** and **Fig.3**) provided quantitative scores of cell vulnerability towards *Salmonella* infection. We secondly investigated the origin of the observed cell vulnerability. Two possibilities can be anticipated: (i) the cellular vulnerability would be an intrinsic host cell attribute (hypothesis 1: “intrinsic vulnerability”) or (ii) it would be induced by bacterial uptake (hypothesis 2: “induced vulnerability”) (**Fig.4A**). In theory, these hypotheses can be distinguished by the observable difference in the probability of the 2nd wave of infection occurring in previously non-infected cells $P(I_2 \mid noI_1)$ as depicted in the two schemes of **Fig.4B** and described as follows: In the case of vulnerability as intrinsic attribute, the probability of infection $P(I_2 \mid noI_1)$ would be lower than $P(I_{2Ctr})$ as the pool of vulnerable cells would have already been partially consumed during the 1st sequential infection, whereas it would remain conserved in the control (**Fig.4B-left**). In the case of induced vulnerability, the probability of infection $P(I_2 \mid noI_1)$ would be similar to $P(I_{2Ctr})$, as the cells would be considered with equivalent vulnerabilities before their first infection (**Fig.4B-right**). The experimental data obtained did not show a significant difference between $P(I_2 \mid$

nol$_1$) and P(I$_{2Ctr}$) (t-test p-value >0,05) (**Fig.4C**), suggesting that vulnerability may be induced by bacterial uptake (**Fig.4B,** *hypothesis 2*). Taking into account the small percentage of cells belonging to the studied subpopulations we caution that the absence of a statistically significant difference between these populations did not allow to exclude the first hypothesis of host cell inherent vulnerability.

Single cell vulnerability to *Salmonella* infection is a combination of intrinsic and induced vulnerability

Considering that the subpopulation comparison could not exclude an involvement of inherent vulnerability, we developed a mathematical model to evaluate the relative contribution of induced and inherent vulnerability to the overall cell vulnerability towards *Salmonella* infection. To investigate the contribution of cell parameters at a single-cell level, we measured different intrinsic variables that could influence the cellular vulnerability, namely the cell morphology (cell perimeters, cell circularity), the local environment (local cell density, number of infected and non-infected neighboring cells), and the above-analyzed features of the *Salmonella* infection (delay between infections, load of intracellular bacteria per cell from I$_1$) (**Fig.5A**). We extracted all these elements using Icy, an image analysis software (18) being recently used for *Salmonella* infection studies *in situ* (19) (*see **Fig.S5A** for illustration of Icy cell segmentation*).

First, we analyzed the distribution of distinct cellular parameters in either infected or in non-infected HeLa (upper panels) and Caco-2 (lower panels) cell populations (**Fig.5A**). Caco-2 cells were cultured at high confluence so that the cells formed a continuous polarized monolayer (*see Materials and Methods*). For both cell types, the infected cells displayed distinct cellular features in comparison to the non-infected cells, such as a higher local crowding reflected by a higher number of neighboring cells in direct contact. Comparing the relative correlations of the cellular parameters, we highlight the presence of strong links between many of them (**Fig.S5B-C, S6 and S7**). In particular cell morphology is highly dependent on the local micro-environment, such as the local cell density that negatively correlates with the cell perimeter in HeLa and Caco-2 cells. Interestingly, cells that were infected during the second bacterial challenge are more likely to be nearby cells that were infected during the first bacterial challenge (“infected neighbor cells”) than by non-infected neighbor cells. Thus *Salmonella* infection of one cell increases the probability of its neighboring cells to be subsequently infected.

To quantify the direct involvement of each studied parameter on the overall cell vulnerability we developed a statistic modeling approach adapted to our high-throughput microscopy dataset on sequential *Salmonella* infection. This model is based on a logistic regression that is able to predict the infection efficiency at a single cell level from cellular parameters. We measured the contribution of each parameter to the prediction by estimating how well the model predicts compared to a model that would ignore one parameter; as described in *Materials and Methods* (**Fig.5B**). Taken separately, the load of intracellular bacteria resulting from I_1 directly improved the prediction of cell vulnerability towards subsequent infection (**Fig.5B**). Thus, host cell vulnerability is induced by bacterial uptake, which is in line with our experimental data. In addition, the host cell parameters linked to cell morphology and local environment also significantly improved the model prediction of infection for HeLa and for Caco-2 cells (*see* **Table.S1** and **Table.S2** *for model details and the value of the coefficients*). Together, our modeling approach revealed that single host cell vulnerability to *Salmonella* infection is a combination of intrinsic and bacterial-induced vulnerability.

We quantified their relative involvement by calculating the model-based fold change in the probability of infection of a cell not infected and having a low score of inherent vulnerability with a cell infected and/or having a high score of inherent vulnerability (**Fig.5C**). This showed that induced and intrinsic vulnerability have both a strong impact on the overall cell vulnerability. Interestingly, the induced vulnerability is more prevalent for *Salmonella* infection of HeLa cells (2.2 fold-increase) than infection of Caco-2 cells (1.3 fold-increase), whereas the inherent vulnerability plays a more prominent role for Caco-2 cell infections (2,6 fold-increase) than for HeLa cells (1.6 fold-increase). From these findings we conclude that the analyzed host cell parameters are differentially involved in relation to cell vulnerability towards *Salmonella* infection depending on the cell type. In particular, the local cell density increases the cell vulnerability for HeLa cells but reduces it for Caco-2 cells (**Fig.5D**). This could be explained by the polarization of the Caco-2 at high confluence and highlights the specificity of each predicted model for a given cell-type.

We also investigated whether the first infection affects the inherent host cell parameters, we compared the correlation between parameters that were identified as being either involved or not involved in the inherent vulnerability of the cell (**Fig.S8**). As their correlations were similar in infected and non-infected cells we concluded that *Salmonella* infection did not impact the implication of the studied inherent cell parameters.

Reliability of the model-based prediction of infection

To investigate the spatial distribution of the cell vulnerability among the cell population, we generated “vulnerability maps” from the original images of the cell population after labeling

each cell nucleus with a color corresponding to its probability of infection (**Fig.6A**). Notably, we could confirm that on average the infected cells were properly assigned with a higher prediction score to be infected than the non-infected ones (*see **Fig.S9** for quantification*). Based on our vulnerability maps, the predicted infected cells showed a very good overlap or were in close vicinity with the experimentally infected cells (**Fig.6A**). This illustrates the reliability of our approach in a qualitative way, and it also underlines the impact of local micro-environment on cell vulnerability. We went on to quantify the veracity of the HeLa and Caco-2 adapted models when confronted with 100 experimentally measured infected and 100 experimentally measured non-infected cells. For both cell-types, models allowed a good prediction in the majority of the cases, 62% for HeLa and 66% for Caco-2, respectively (**Fig.6B**). Taken together, these results attest that the probability of *Salmonella* infection success can be forecast at the near single-cell level based on host cell parameters.

Involvement of cellular cholesterol level as an inherent vulnerability factor

To investigate the molecular players that are linked to the inherent cell vulnerability towards *Salmonella* infection, we analyzed the plasma membrane composition as a main feature known to be relevant to *Salmonella* infection. We focused on cholesterol as the cells at low crowding present a higher amount of free cholesterol than the cells at high crowding (15). We monitored the relationship between global cellular cholesterol levels and host cell targeting performing *Salmonella* infection of HeLa cells for 30 min, followed by cholesterol labeling via filipin staining. Although filipin is the most commonly used tool to assess cholesterol content, it also displays very fast photobleaching properties (20). Thus, the automatic acquisition of an entire 96-wells plate would introduce a strong bias due to the loss of filipin signal during the acquisition. To circumvent this technical issue, we carried out flow cytometry acquisition and analysis (**Fig.7**). For each experiment, we binned the total cell population into five subpopulations corresponding to the increasing cellular levels of cholesterol that we classified as 1 to 5, with each subpopulation containing 20% of the total cells (*see **Fig.S10** for FACS gating details*). Comparing the number of infected cells in these different subpopulations with different amounts of cholesterol in HeLa (**Fig.7A**) and Caco-2 (**Fig.7B**) cells, we revealed that the probability of infection correlates in both cases with the cholesterol levels. Increasing the cholesterol level corresponds to a decrease of the probability of *Salmonella* infection in HeLa cells (**Fig.7A**), however it also corresponds to an increase of the infection in Caco-2 cells (**Fig.7B**). Thus, similarly to the cell density, the cholesterol level is a host cell parameter allowing to estimate the cell vulnerability towards *Salmonella* infection in a cell-type dependent manner.

DISCUSSION

Cellular heterogeneity describes cases in which genetically identical cells present different behaviors and morphologies. This biological phenomenon is commonly present in an epithelial layer of an individual as well as within a monolayer of cultured cells. Despite the realization of the importance of cellular heterogeneity, its study has only become feasible during recent years, mainly thanks to the implementation of novel technologies such as imaging and computer-assisted analyses. In the context of pathogen infection, this heterogeneity produces cells unequally vulnerable or resistant, which impacts on the overall infection.

We investigated the cell vulnerability of epithelial cells for *S*. Typhimurium infection. According to our results, infected cells display a strikingly higher probability of being re-infected with *Salmonella*, even after the disappearance of membrane ruffles. We obtained similar results in two relevant epithelial cell lines, HeLa and Caco-2, suggesting that this represents a conserved propensity towards *Salmonella* infection. The measured cellular vulnerability remained unaltered for all measured time-points ranging from a delay of 1 h to 3 h between the infections. Attributing a "vulnerability score" to the challenged cells, we showed a higher vulnerability score in cells that had been previously infected, and we found that this score increased with the amount of intracellular bacteria contained by a given cell. This result raises the issue of the bacterial impact on the cell vulnerability. Therefore, we aimed at distinguishing inherent cell vulnerability from the one induced by bacterial uptake (**Fig.4A,** *hypothesis 1 and 2 respectively*) exploiting the imaging data obtained via a high-content analytical pipeline. This allowed visualization of the infection *in situ* and provided a large number of associated cellular parameters. We quantified the implication of specific parameters associated with individual cells on the cell vulnerability towards *Salmonella* infection. It appeared clearly that the efficiency of early bacterial uptake during the first infection directly determines cell vulnerability. Thus *Salmonella* induces an increase in the cell vulnerability towards subsequent infections.

While long-term cooperation among bacteria has been extensively studied for the communities of bacteria living in a common extracellular environment (21), little is known about the cooperation between intracellular and extracellular bacteria leading to increased bacterial uptake. Nevertheless, this phenomenon has been investigated more extensively for many viruses, including bacteriophages (22), influenza virus (23), poxviruses (24, 25), flaviviruses (26, 27) alphaviruses (28), and alphaherpesviruses (29). Generally, those works have demonstrated that the first virus to infect a cell has the capacity to prevent co-infection of other viruses belonging either to the same strain, or to more distantly related or unrelated

strains. It is termed “superinfection exclusion” and may protect limited cellular resources and promote the replication and dissemination of the originally infecting virus. By analogy, the increased probability of cellular re-infection by *Salmonella* can be phrased as a “superinfection promotion”. It remains to be clarified if such process is relevant for all intracellular bacteria. For instance and in contrast to *Salmonella* infection, Jorgensen *et al* reported that the *Chlamydia* effector protein CPAF secreted from bacteria within mature inclusions prevents those that are still extracellular to invade (30). Thus, CPAF could be a factor mediating *Chlamydia* resistance towards superinfection.

Our approach also allowed the relative quantification of the impact of different host cell parameters on the inherent vulnerability of host cell to *Salmonella* infection. In particular, morphological attributes and local cell crowding are highly linked with this vulnerability. Cell crowding as a major determinant for the probability to become infected has been proposed by Snijder and colleagues in the context of viral infection. They showed that during infections by the simian virus SV40 or the mouse hepatitis virus (MHV), the targeted cells have different localization within cell islets (13). SV40 and MHV infect preferentially either peripheral or central cells, a phenomenon that is linked to the differential expression levels of focal adhesion kinase and the presence of sphingolipid GM1 at the plasma membrane of the challenged host cells. Thus, similarly to several viral infections, the probability of infection of a single cell by *Salmonella* is influenced by its local environment.

Our analytical tools will be useful for further studies on *Salmonella*, and for other researchers working on different intracellular bacterial pathogens, such as *Chlamydia*, *Listeria* or *Shigella* (*see Materials and Methods*). We revealed that some cells are indeed intrinsically more vulnerable to *Salmonella* and will be targeted by the bacteria first. Most of the tested parameters appeared to be relevant for model-based infection prediction but are differentially involved in the cell vulnerability depending on the cell-type studied. Developing an adapted model based on host cell parameters we could forecast the probability of *Salmonella* infection success at the near single-cell level. Interestingly, the number of infected neighboring cells is highly increased in the population of infected cells. Cases of bacterial uptake impacting on the cells neighboring the infection (called bystander cells) have been previously reported for *Shigella* that induces an IL-8 immune response after NFκB activation detectable from 2 h pi in 70% of the bystander cells (31). However, it is not known whether the neighboring cells are also more susceptible to *Shigella* entry.

Because of our lack of knowledge of host factors that are involved in the early attachment, such as potential entry receptors, it remains difficult to identify the molecular mechanisms

that establish the differential vulnerability during *Salmonella* infection. Although receptors for direct recognition of *Salmonella* have been proposed, such as the cystic fibrosis transmembrane conductance regulator (CFTR) (32) and the epithelium growth factor receptor (EGFR) (33), many cell types infected by *Salmonella* do not express them (34). Therefore, it has been proposed that recognition mechanisms likely involve more ubiquitous factors (35). To explore the molecular cues involved in the inherent heterogeneity of host cell vulnerability, we decided to investigate the membrane lipid composition, in particular cellular cholesterol. We found that the cholesterol amount at single cell level in HeLa and Caco-2 cells correlates with the vulnerability of these cells to *Salmonella* infection. In HeLa cells *Salmonella* preferentially targets cells with low amounts of cholesterol. However, in Caco-2 cells, *Salmonella* preferentially targets cells with high amounts of cholesterol. Interestingly, these results on an implicated host molecule are in agreement with the morphological feature of local density. Frechin and colleagues reported that cells at high density contain lower amounts of cholesterol (15). Besides, at high density HeLa and Caco-2 cells display an increase or a decrease in inherent cell vulnerability, respectively. This is in line with the correlation that we reported between the cholesterol level and host cell vulnerability. The molecular role of cholesterol during *Salmonella* infection is still under debate. Several studies have demonstrated that the *Salmonella* SipB effector and translocon component requires cholesterol for proper functioning (35, 36). In this context, it should be noted that the translocons operate in small cholesterol-rich microdomains at the plasma membrane and cannot be linked readily to the overall cholesterol levels. Furthermore, those studies were based on sterol sequestering agents and biosynthesis inhibitors. Contrastingly, Gilk and colleagues have shown that cholesterol is not essential for *Salmonella* invasion and intracellular replication inside host cells using an original mouse model (37). In our study we highlighted that non-treated HeLa cells with a low amount of global cellular cholesterol are preferentially targeted by *Salmonella,* which does not exclude a potential involvement of cholesterol at the subcellular level. Santos and colleagues have also reported that the preferential invasion of hTERT-RPE1 and HeLa mitotic cells by *Salmonella* was SipB and cholesterol dependent (11). However, the low amount of mitotic cells in the whole population (< 4%) may have a limited impact on the overall inherent vulnerability of the host cell population. Thus our observation that the most vulnerable HeLa cells display a low cholesterol level is not in contradiction with previous publication on cholesterol involvement during *Salmonella* infection process.

In conclusion, our study represents a first step in understanding *Salmonella* cell targeting and provides a path for the identification of cellular and bacterial factors involved in host cell vulnerability. Such factors could be targeted to render a cell more resistant to pathogen

infections, allowing potential new therapeutic strategies. Together, our study delineates in a quantitative manner the importance of vulnerable cell recognition and bacterial cooperation for cell targeting by *S*. Typhimurium.

MATERIALS AND METHODS

Bacterial Strains

The following *S.* Typhimurium were used: SL1344 (wild type), SL1344 pM965 (*Salmonella*-GFP) described by Stecher *et al* (38), and SL1344 pGG2 (*Salmonella*-dsRed) obtained after transformation of SL1344 with the pGG2 plasmid described by Lelouard *et al* (39). Bacteria were grown in Lysogeny Broth (LB) medium supplemented with 0.3 M NaCl and ampicillin at 50 µg/ml at 37ºC in an orbital shaker.

Cell Culture

All cell culture reagents were purchased from Invitrogen unless otherwise stated. Human epithelial HeLa cells (clone CCL-2 from ATCC) were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS), at 37 ºC, 5% CO_2. HeLa cells were plated at a concentration of $1.5x10^4$ cells/well in glass-bottom 96-wells plates 24 h before infection, so that they displayed about 80% of confluence on the infection day. Intestinal epithelial Caco-2 TC7 cells (kindly provided by P. Sansonsetti) were grown in DMEM supplemented with 10% FBS at 37°C, 10% CO_2. Caco-2 cells were plated at a concentration of $3.5x10^4$ cells/well in glass-bottom 96-wells plates 48 h before infection, so that they displayed a polarized (but not differentiated) continuous monolayer on the infection day. All infection assays were performed in EM buffer (120 mM NaCl, 7 mM KCl, 1.8 mM $CaCl_2$, 0.8 mM $MgCl_2$, 5 mM glucose, 25 mM HEPES, pH 7.4). HeLa cells were transfected with pEGFP-actin plasmid DNA (40) from a maxiprep, using the X-tremeGENE 9 DNA transfection reagent (Roche) for 48 h.

Double Infection Assays

For invasion experiments, overnight bacterial cultures were sub-cultured 1/20 and grown until late exponential/early stationary phase. Before infection, bacteria were gently washed and resuspended in EM buffer. Bacteria were added to the cells at an MOI of 30 corresponding to CFU, and incubated for 30 min at 37 ºC, 5% or 10% CO_2 for HeLa or Caco-2 cells, respectively. Non-internalized bacteria were eliminated by washing 3 times with warm EM buffer and incubated for 1, 2 or 3 h at 37 ºC, 5% or 10% CO_2 for HeLa or Caco-2 cells, respectively. Adding EM buffer containing 100 µg/ml gentamicin for 1 h killed extracellular bacteria. The concentration of gentamicin was then decreased to 10 µg/ml and 10% FBS was added to the medium. At the desired time points, the cells were washed again in EM buffer to eliminate the remaining gentamicin and re-infected with a fresh batch of sub-cultured bacteria following the same protocol. After killing the extracellular bacteria again by

518 a 1 h of incubation with EM buffer containing 100 µg/ml gentamicin, the cells were fixed with
519 4% paraformaldehyde at room temperature for immunofluorescence analysis.
520
521 **Microscopy**
522 All image acquisitions were performed on a Nikon inverted widefield microscope using a
523 20x/0.5NA air objective, an automatic programmable XY-stage and the Nikon perfect focus
524 system. For sequential infections of HeLa and Caco-2 cells, 161 fields were imaged per well
525 and four channels per field were captured using a CoolSnap2 camera (Roeper Scientific).
526 Nuclei and cells were stained using DAPI (excitation/emission wavelengths: 350/470 nm)
527 and the cell bodies with CellMask DeepRed Plasma Membrane Stain
528 (ThermoFisherScientific, excitation/emission wavelengths: 640/670 nm) respectively. Caco-2
529 cells were stained with the FM® 4-64 membrane dye (Invitrogen) before time lapse imaging
530 (excitation/emission wavelengths: 558/734 nm). Quantification of the ruffle timing was
531 performed on the same microscope, using a 20x/0.5NA air objective and time intervals of 3
532 min for 90 min. Time lapse imaging of ruffles was performed on a DeltaVision widefield
533 microscope using a 60x/1.42 NA oil objective and z-stacks with a spacing of 500 nm. The
534 images were subsequently de-convolved using DeltaVision Elite integrated software.
535
536 **Cholesterol measurements**
537 HeLa and Caco-2 cells were challenged with SL_{GFP} for 30 min before trypsinization, fixation
538 with 4% paraformaldehyde at room temperature and incubation with 16ug/mL filipin complex
539 from *Streptomyces filipinensis* (Sigma-Aldrich). This treatment was directly followed by FACS
540 measurement on *BD FACS CANTO* cytometer using the excitation/emission wavelengths of
541 405/450 nm and 488/530 nm for filipin and GFP fluorescence detection respectively. Infected
542 and non-infected cells were distinguished using the green fluorescence emitted by SL_{GFP}
543 (*see* ***Fig.S10*** *for gating details*). Data were processed using FlowJo software.
544
545 **Image Analysis**
546 All images were analyzed with two open source software: CellProfiler (http://cellprofiler.org/)
547 and Icy (http://icy.bioimageanalysis.org/). CellProfiler was used to detect each single cell
548 and the number of its intracellular salmonellae expressing either GFP or dsRed. The
549 following modules were used during the analysis: *IdentifyPrimaryObjects* recognized nuclei
550 and bacteria; *IdentifySecondaryObjects* identified cells (here the secondary objects) by
551 extending the nuclear area previously recognized; *RelateObjects* assigned bacteria within
552 individual cells. Icy was used for accurate detection of cell borders and the cellular
553 microenvironment analysis. We used a graphical environment called *Protocols* for the
554 development of an analytical pipeline including the following plugins: *HK-Means* that identify

nuclei by pre-filtering the signal to identify objects within a size range; *Spot Detector* that identify bacteria; *Active Contours* that identify the edges of the plasma membrane by propagating the Region of Interest (ROI) detected for the nuclei; and *Javascript* that parent the ROI of cells with bacteria, to measure local cell density and to distinguish which neighboring cells are infected by which bacteria.

Probability

$P(I_2|I_1)$ means "Probability of the 2nd sequential infection, knowing that the cell has been infected by the 1st one" and is calculated as follows:

$$\boldsymbol{P(I_2|I_1) = P(I_1\&I_2)/P(I_1)}$$

Where *$P(I_1)$ = [Number of cells in I_1 / Total number of cells],* and *$P(I_1\&I_2)$ = [Number of cells in $I_1\&I_2$ / Total number of cells].*

$P(I_2|noI_1)$ means "Probability of the 2nd sequential infection, knowing that the cell has not been infected by the 1st one" and is calculated as follows:

$$\boldsymbol{P(I_2|noI_1) = P(I_2\&noI_1)/P(noI_1)}$$

Where *$P(noI_1)$ = [Number of cells in noI_1 / Total number of cells],* and *$P(I_2\&noI_1)$ = [Number of cells in $noI_1\&I_2$ / Total number of cells].*

Model

We modeled the influence of multiple parameters on the probability of a second infection. A Boolean variable Y represents the second infection: It is equal to 1 for infected cells and 0 otherwise. Its probability is predicted by the following seven parameters: *Load of infection* (LOI) represents the number of infecting bacterium during the first infection, separated in 4 groups corresponding to no (0 bacteria), low (1 or 2), medium (3 to 8) or high (9+) infection. *Delay* is a categorical variable corresponding to the delay between the 1st and the 2nd infections (1, 2 or 3 h). *Infected neighbor cells* (X1) refers to the number of cells in contact that had been infected during the first infection. *Non-Infected neighbor cells* (X2) refers to the number of cells in contact which had not been infected during the first infection. *Local Cell Density* (X3) is the number of cells present in a vicinity of 100 µm. The distance is calculated between the center of the nuclei. *Cell perimeter* (X4) is the length of the perimeter of the cell (in µm) obtained after segmentation. *Circularity* (X5) refers to the cell circularity defined as: "4π*area/perimeter^2". This parameter is higher for circular cells, and lower for cells that are elongated or have complex shape, but does not depend *a priori* on the cell size. In practice we used to its square root. The probability of Y during the second infection is modeled as:

$$\boldsymbol{P(Y = 1 \mid X1,\ldots,Xn) = 1 / [1 + exp(-(aLOI + aDelay + a1\,X1 + \ldots + a5\,X5))]}$$

where aLOI (resp. aDelay) has a different value for each of the LOI categories (resp. Delay categories), and $a_1,...,a_5$ are constants. All parameters were learned by maximizing the likelihood of the model, e.g. the probability of the observed data as measured by the model. We used 115 000 and 327 000 cells to train and test the model for HeLa and Caco-2 cells respectively. We divided the cell population into two random sets; the training set ($9/10^{th}$ of the cells per replicate) and testing set ($1/10^{th}$ of the cells) and computed the likelihood of infection observed in the testing set. The higher the likelihood, the better the parameters of the model predicted infection. We repeated this procedure 100 times. To measure the improvement of infection prediction by taking into account each parameter, the likelihood of the complete model was compared (on a log scale) with the likelihood of seven models ignoring each time one parameter. This difference of log-likelihood is reported in **Fig.5B**.

Quantification of the impact of a parameter towards cell vulnerability was obtained by applying our statistical model to the 1^{st} and the 3^{rd} quantile values of a given parameter, while other parameters were kept equal at their median values. We obtained the probabilities of the second infection for these two sets and reported their ratio. In **Fig.5D**, the arrows "↗" and "↘" correspond to a ratio above and under 1 respectively. The parameters-values corresponding to a low inherent vulnerability of HeLa and Caco-2 cells were the following: local cell density (1^{st} quantile and 3^{rd} quantile respectively), cell perimeter (1^{st} quantile), infected neighboring cells (median), non-infected neighboring cells (median), circularity (median and 3^{rd} quantile respectively). The parameters-values corresponding to a high inherent vulnerability of HeLa and Caco-2 cells were the following: local cell density (3^{rd} quantile and 1^{st} quantile respectively), cell perimeter (3^{rd} quantile), infected neighboring cells (median), non-infected neighboring cells (median), circularity (median and 1^{st} quantile respectively).

Models reliability was evaluated using 100 infected and 100 non-infected cells and quantifying the amount of "good predictions" among those cells. We repeated this procedure 100 times and showed the average. As a comparison, a random model would provide approximately 50% of "good predictions".

Statistical analysis

The statistical analysis was performed using R and GraphPad Prism. T-tests were used to evaluate the significance of the results, referred like *, **, *** for p-values <0.05, <0.01, and <0.001, respectively.

Supplemental information

629 The pipeline used on CellProfiler and on ICY, as well as the R code used to generate the
630 model can be provided by the authors.

631 **ACKNOWLEDGMENTS**

632

633 We thank Jennifer Fredlund and Andrew Rutenberg for their help during the initial phase of
634 the project, Adrien Sauvaget, Claude Loverdo, Kristine Schauer and Uriel Hazan for
635 productive discussions, Mariana Ferrari for her help with the FACS experiments and all the
636 members of the DIHP unit and BioImage Analysis Group for helpful interactions. VS was
637 supported by a Ph.D. fellowship from the University Paris Diderot attributed by the ENS
638 Cachan, Université Paris-Saclay. JE is member of the LabEx consortia IBEID and
639 MilieuInterieur. JE also acknowledges support of from the ANR (grant StopBugEntry and
640 AutoHostPath) and the ERC (CoG EndoSubvert).

641 **REFERENCE**
642

643 1. Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, Jones TF, Fazil A,
644 Hoekstra RM. 2010. The Global Burden of Nontyphoidal Salmonella Gastroenteritis.
645 Clinical Infectious Diseases 50:882-889.
646 2. Carter PB, Collins FM. 1974. The route of enteric infection in normal mice. The
647 Journal of experimental medicine 139:1189-203.
648 3. Watson KG, Holden DW. 2010. Dynamics of growth and dissemination of Salmonella
649 in vivo. Cellular Microbiology 12:1389-1397.
650 4. Misselwitz B, Barrett N, Kreibich S, Vonaesch P, Andritschke D, Rout S, Weidner K,
651 Sormaz M, Songhet P, Horvath P, Chabria M, Vogel V, Spori DM, Jenny P, Hardt
652 WD. 2012. Near surface swimming of salmonella Typhimurium explains target-site
653 selection and cooperative invasion. PLoS Pathogens 8:9.
654 5. Vonaesch P, Cardini S, Sellin ME, Goud B, Hardt WD, Schauer K. 2013. Quantitative
655 insights into actin rearrangements and bacterial target site selection from
656 SalmonellaTyphimurium infection of micropatterned cells. Cellular Microbiology
657 15:1851-1865.
658 6. Misselwitz B, Kreibich SK, Rout S, Stecher B, Periaswamy B, Hardt WD. 2011.
659 Salmonella enterica serovar typhimurium binds to hela cells via fim-mediated
660 reversible adhesion and irreversible type three secretion system 1-mediated docking.
661 Infection and Immunity 79:330-341.
662 7. Haraga A, Ohlson MB, Miller SI. 2008. Salmonellae interplay with host cells. Nature
663 Reviews Microbiology 6:53-66.
664 8. LaRock DL, Chaudhary A, Miller SI. 2015. Salmonellae interactions with host
665 processes. Nature Reviews Microbiology 13:191-205.
666 9. Knodler LA. 2015. Salmonella enterica: Living a double life in epithelial cells. Current
667 Opinion in Microbiology 23:23-31.
668 10. Santos JC, Enninga J. 2016. At the crossroads: Communication of bacteria-
669 containing vacuoles with host organelles. Cellular Microbiology 18:330-339.
670 11. Santos AJM, Meinecke M, Fessler MB, Holden DW, Boucrot E. 2013. Preferential
671 invasion of mitotic cells by Salmonella reveals that cell surface cholesterol is maximal
672 during metaphase. Journal of Cell Science 126:2990-2996.
673 12. Lorkowski M, Felipe-López A, Danzer CA, Hansmeier N, Hensel M. 2014. Salmonella
674 enterica invasion of polarized epithelial cells is a highly cooperative effort. Infection
675 and Immunity 82:2657-2667.

13. Snijder B, Sacher R, Rämö P, Damm E-M, Liberali P, Pelkmans L. 2009. Population context determines cell-to-cell variability in endocytosis and virus infection. Nature 461:520-523.
14. Liberali P, Snijder B, Pelkmans L. 2014. Single-cell and multivariate approaches in genetic perturbation screens. Nature Reviews Genetics 16:18-32.
15. Frechin M, Stoeger T, Daetwyler S, Gehin C, Battich N, Damm E-M, Stergiou L, Riezman H, Pelkmans L. 2015. Cell-intrinsic adaptation of lipid composition to local crowding drives social behaviour. Nature 523:88-91.
16. Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin Da, Chang JH, Lindquist Ra, Moffat J, Golland P, Sabatini DM. 2006. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome biology 7:R100.
17. Kamentsky L, Jones TR, Fraser A, Bray MA, Logan DJ, Madden KL, Ljosa V, Rueden C, Eliceiri KW, Carpenter AE. 2011. Improved structure, function and compatibility for cellprofiler: Modular high-throughput image analysis software. Bioinformatics 27:1179-1180.
18. de Chaumont F, Dallongeville S, Chenouard N, Hervé N, Pop S, Provoost T, Meas-Yedid V, Pankajakshan P, Lecomte T, Le Montagner Y, Lagache T, Dufour A, Olivo-Marin J-C. 2012. Icy: an open bioimage informatics platform for extended reproducible research. Nature Methods 9:690-696.
19. McQuate SE, Young AM, Silva-Herzog E, Bunker E, Hernandez M, de Chaumont F, Liu X, Detweiler CS, Palmer AE. 2016. Long-term live-cell imaging reveals new roles for Salmonella effector proteins SseG and SteA. Cellular Microbiology doi:10.1111/cmi.12641.
20. Boutte Y, Men S, Grebe M. 2011. Fluorescent in situ visualization of sterols in Arabidopsis roots. Nat Protoc 6:446-56.
21. Davey ME, O'toole GA. 2000. Microbial biofilms: from ecology to molecular genetics. Microbiology and molecular biology reviews : MMBR 64:847-67.
22. Cumby N, Davidson AR, Maxwell KL. 2012. The moron comes of age. Bacteriophage 2:225-228.
23. Huang IC, Li W, Sui J, Marasco W, Choe H, Farzan M. 2008. Influenza A virus neuraminidase limits viral superinfection. J Virol 82:4834-4843.
24. Doceul V, Hollinshead M, van der Linden L, Smith GL. 2010. Repulsion of superinfecting virions: a mechanism for rapid virus spread. Science (New York, NY) 327:873-876.

25. Laliberte JP, Moss B. 2014. A novel mode of poxvirus superinfection exclusion that prevents fusion of the lipid bilayers of viral and cellular membranes. Journal of virology 88:9751-68.

26. Zou G, Zhang B, Lim P-Y, Yuan Z, Bernard Ka, Shi P-Y. 2009. Exclusion of West Nile virus superinfection through RNA replication. Journal of virology 83:11765-11776.

27. Schaller T, Appel N, Koutsoudakis G, Kallis S, Lohmann V, Pietschmann T, Bartenschlager R. 2007. Analysis of hepatitis C virus superinfection exclusion by using novel fluorochrome gene-tagged viral genomes. Journal of virology 81:4591-603.

28. Karpf aR, Lenches E, Strauss EG, Strauss JH, Brown DT. 1997. Superinfection exclusion of alphaviruses in three mosquito cell lines persistently infected with Sindbis virus. Journal of virology 71:7119-7123.

29. Criddle A, Thornburg T, Kochetkova I, DePartee M, Taylor MP. 2016. gD-Independent Superinfection Exclusion of Alphaherpesviruses. Journal of Virology 90:4049-4058.

30. Jorgensen I, Bednar MM, Amin V, Davis BK, Ting JPY, McCafferty DG, Valdivia RH. 2011. The chlamydia protease CPAF regulates host and bacterial proteins to maintain pathogen vacuole integrity and promote virulence. Cell Host and Microbe 10:21-32.

31. Kasper CA, Sorg I, Schmutz C, Tschon T, Wischnewski H, Kim ML, Arrieumerlou C. 2010. Cell-cell propagation of NF-??B transcription factor and MAP kinase activation amplifies innate immunity against bacterial infection. Immunity 33:804-816.

32. Pier GB, Grout M, Zaidi T, Meluleni G, Mueschenborn SS, Banting G, Ratcliff R, Evans MJ, Colledge WH. 1998. Salmonella typhi uses CFTR to enter intestinal epithelial cells. Nature 393:79-82.

33. Pace J, Hayman MJ, Galán JE. 1993. Signal transduction and invasion of epithelial cells by S. typhimurium. Cell 72:505-514.

34. Jones BD, Paterson HF, Hall a, Falkow S. 1993. Salmonella typhimurium induces membrane ruffling by a growth factor-receptor-independent mechanism. Proceedings of the National Academy of Sciences of the United States of America 90:10390-10394.

35. Garner MJ, Hayward RD, Koronakis V. 2002. The Salmonella pathogenicity island 1 secretion system directs cellular cholesterol redistribution during mammalian cell entry and intracellular trafficking. Cellular Microbiology 4:153-165.

36. Hayward RD, Cain RJ, McGhie EJ, Phillips N, Garner MJ, Koronakis V. 2005. Cholesterol binding by the bacterial type III translocon is essential for virulence effector delivery into mammalian cells. Molecular Microbiology 56:590-603.

748 37. Gilk SD, Cockrell DC, Luterbach C, Hansen B, Knodler LA, Ibarra JA, Steele-
749 Mortimer O, Heinzen RA. 2013. Bacterial Colonization of Host Cells in the Absence of
750 Cholesterol. PLoS Pathogens 9.
751 38. Stecher B, Hapfelmeier S, Müller C, Kremer M, Stallmach T, Hardt W-d, Mu C. 2004.
752 Flagella and Chemotaxis Are Required for Efficient Induction of Salmonella enterica
753 Serovar Typhimurium Colitis in Streptomycin-Pretreated Mice Flagella and
754 Chemotaxis Are Required for Efficient Induction of Salmonella enterica Serovar
755 Typhimurium Colitis. Infect Immun 72:4138-4150.
756 39. Lelouard H, Henri S, De Bovis Ba, Mugnier Bnd, Chollat-Namy A, Malissen B,
757 M??resse Sp, Gorvel JP. 2010. Pathogenic Bacteria and Dead Cells Are Internalized
758 by a Unique Subset of Peyer's Patch Dendritic Cells That Express Lysozyme.
759 Gastroenterology 138:173-184.e3.
760 40. Ehsani S, Santos JC, Rodrigues CD, Henriques R, Audry L, Zimmer C, Sansonetti P,
761 Van Nhieu GT, Enninga J. 2012. Hierarchies of host factor dynamics at the entry site
762 of Shigella flexneri during host cell invasion. Infection and Immunity 80:2548-2557.
763

764 **FIGURE CAPTIONS**

765

766 **Fig.1. Double infections allow studies of *Salmonella* cooperation at the single cell**
767 **level. A. B. C.** Overview of the experimental workflow used in this study. **A.** Sequential
768 infection protocol: HeLa cells grown in 96-wells plates since 24 h were subjected for 30 min
769 to a first infection by SL_{GFP}. This was followed by elimination of extracellular bacteria by
770 gentamicin and incubation of the cells for 1, 2 or 3 h. The cells were subsequently challenged
771 by a second infection with SL_{dsRed} for 30 min. After removal of the extracellular bacteria, the
772 samples were fixed. Nuclei were stained with DAPI and cell membranes were stained with
773 CellMask before microscopic acquisition of the entire wells. **B.** Representative image of
774 SL_{GFP} and SL_{dsRed} internalized in HeLa cells. Host cell nuclei are visible through DAPI (in
775 blue), and cell membranes through CellMask (in grey). Scale bar correspond to 5µm. **C.**
776 Scheme of our statistical analysis of different subpopulations. The following cellular
777 populations can be distinguished: those cells infected during the 1st infection (I_1) or not (noI_1),
778 those infected during the 2nd infection (I_2) or not (noI_2), along with the related subpopulations
779 ($I_1\&I_2$, $noI_1\&noI_2$)**.** This scheme maps the case of two independent infections. **D.** Time
780 distribution of the ruffle disappearance during *Salmonella* infection followed in actin-GFP
781 transfected cells by time-lapse microscopy. **E. F.** Comparison of the conditional probability of
782 infection for two different populations during synchronous infection of SL_{GFP} and SL_{dsRed}. **E.**
783 Results obtained in HeLa cells. **F.** Results obtained in Caco-2 cells. The MOIs were chosen
784 to obtain in average 30% of the cells infected and calculated after CFU counting (n >=3). P-
785 values were obtained after t-test. **G.** Comparison of an independent model (left) with the
786 obtained data (right). The percentages are averaged from 6 independent experiments,
787 represented in **E** with an MOI of 30.

788

789 **Fig.2. The probability of being re-infected by *Salmonella* remains higher for already**
790 **infected cells after entry ruffle disappearance. A-B.** Conditional probability of infection for
791 two different populations during sequential infection with a delay of 2 h for HeLa cells (**A**) and
792 Caco-2 cells (**B**). Results were obtained from 3 independent experiments and P-values were
793 obtained after paired t-test. **C.** The vulnerability score was plotted for infection with a 1, 2 or 3
794 h delay before the second infection in HeLa cells. The red line corresponds to $P(I_2 \mid I_1)=P(I_2 \mid$
795 $noI_1)=1$ indicating the independence of the infections I_2 and I_1. Values above the red line
796 correspond to $P(I_2 \mid I_1) > P(I_2 \mid noI_1)$ indicating a cooperation between infections. Values below
797 the red line correspond to $P(I_2 \mid I_1) < P(I_2 \mid noI_1)$ indicating a competition between infections.

Results were obtained from 3 independent experiments per time-point, and P-values were obtained after unpaired t-test.

Fig.3. Cell vulnerability can be predicted from the number of bacteria previously internalized. A. Distribution of the number of intracellular bacteria detected at 1.5 h pi in HeLa cells (average from 3 replicates). The infection efficiencies are clustered in 3 groups: low, medium and high infection, corresponding respectively to 1 to 2; 3 to 8 or more than 9 bacteria per cell. **B.** The vulnerability score is represented as a function of the number of intracellular bacteria resulting from the 1^{st} infection in HeLa cells. **C.** Probability of a cell to be highly infected during the 2^{nd} infection ($nI_2 \geq 9$) as a function of the number of intracellular bacteria being internalized during the 1^{st} infection in HeLa cells. **B** and **C** represent the data merged from all the experiments (delay of 1, 2 and 3 h before the second infection). Groups of infection efficiency are identical in **A**, **B** and **C**.

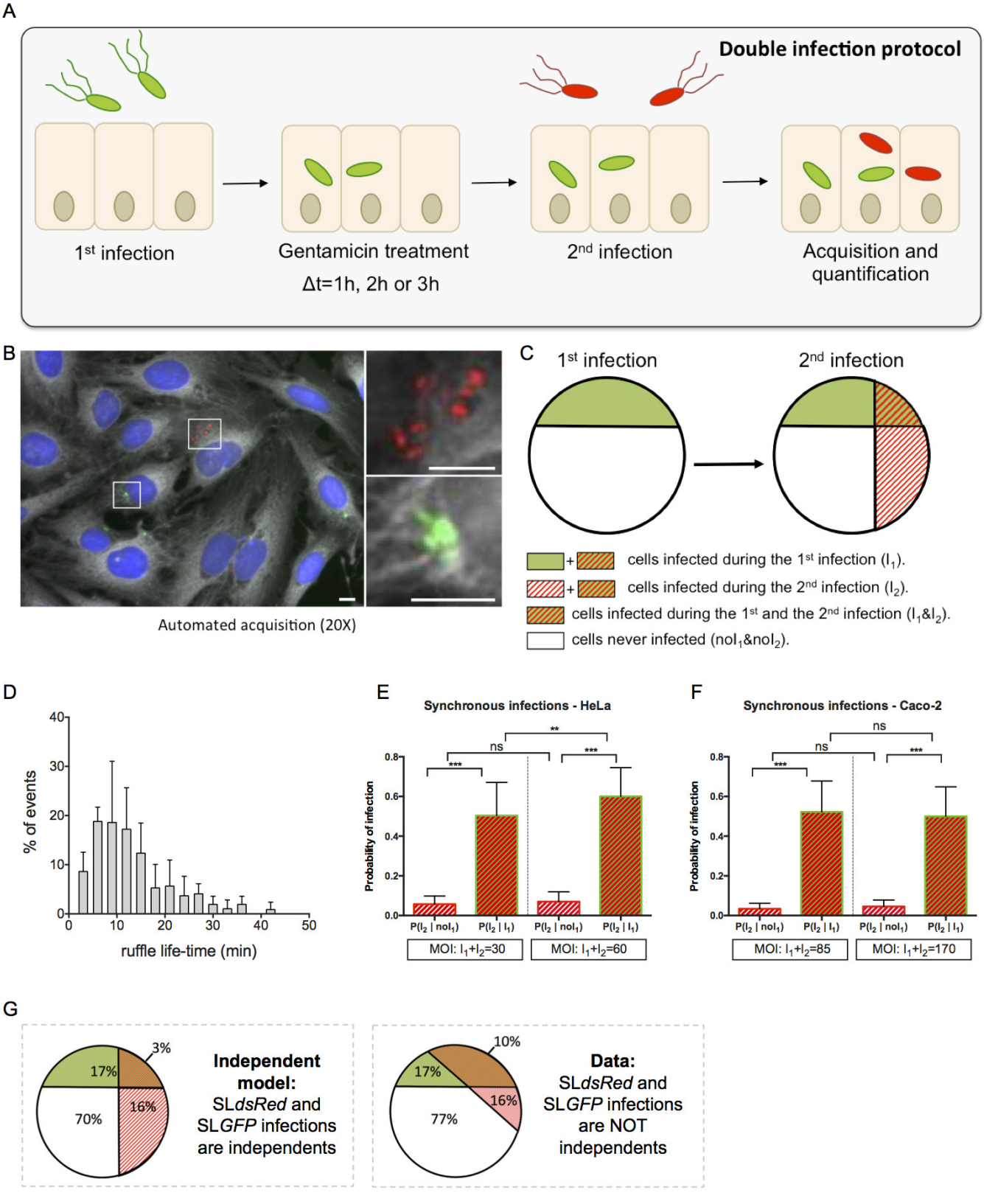
Fig.4. Cell vulnerability examined as an intrinsic or an induced property. A. Schemes of the two hypotheses for the origin of cell vulnerability. In the hypothesis 1, cell vulnerability is inherent: some cells (in orange) are more vulnerable towards infection than other cells (in yellow). In the hypothesis 2, cell vulnerability is induced by bacterial uptake: before infection cells are equal regarding their vulnerability (in yellow), but after infection the infected cells turn progressively more vulnerable (in orange). **B.** Graphic representation of the theoretical distribution of the different populations in the case of hypothesis 1 (left) or hypothesis 2 (right). **C.** Probability of infection during sequential infection of HeLa cells with 1, 2 and 3 h delays for control cells (I_{2Ctr}) and cells non infected during the 1^{st} infection (noI_1). P-values were obtained after unpaired t-test ($P(I_{2ctr})$ vs $P(I_2 \mid noI_1)$).

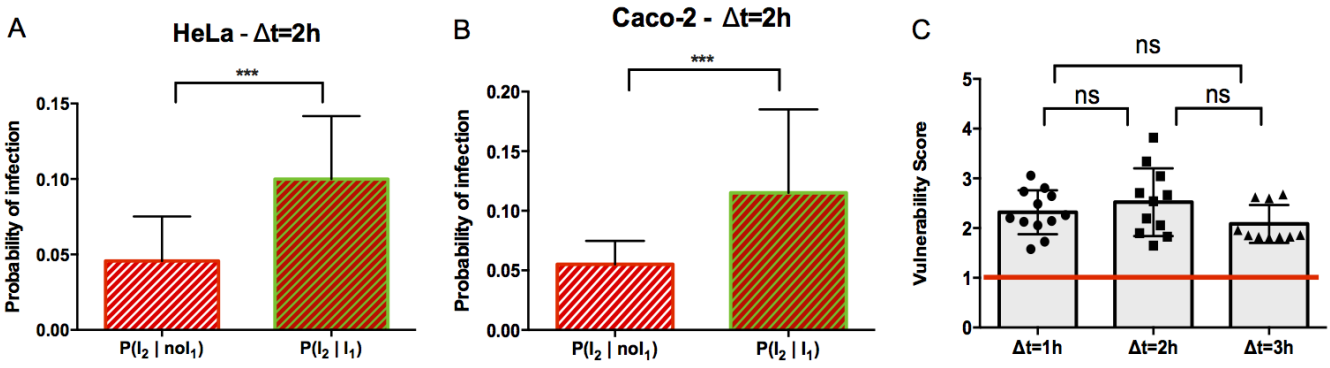
Fig.5. Single cell vulnerability to *Salmonella* infection is a combination of intrinsic and induced vulnerability. A. The depicted cellular parameters were determined for HeLa cells (upper panel) and Caco-2 cells (lower panel) as described in detail in *Materials and Methods*. An overlay of the distribution of some of these parameters in infected (red) or non-infected (blue) cells is shown. **B.** Quantification of the improvement of infection prediction by each cell parameters by subtracting the likelihood (in log) of the model including all parameters from a model ignoring one parameter. Results are averaged over 100 training/testing circles for each model. P-values were obtained after paired t-test. **C.** Fold change of the probability of infection as a function of the intrinsic vulnerability and of a previous infection. **D.** Increasing

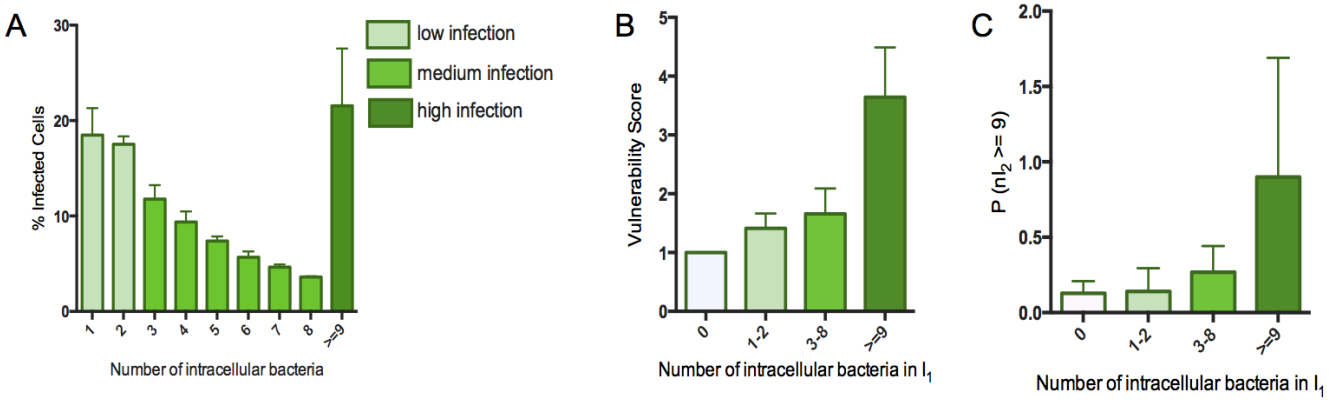
or decreasing of the probability of infection when the listed cell parameters increase their values.

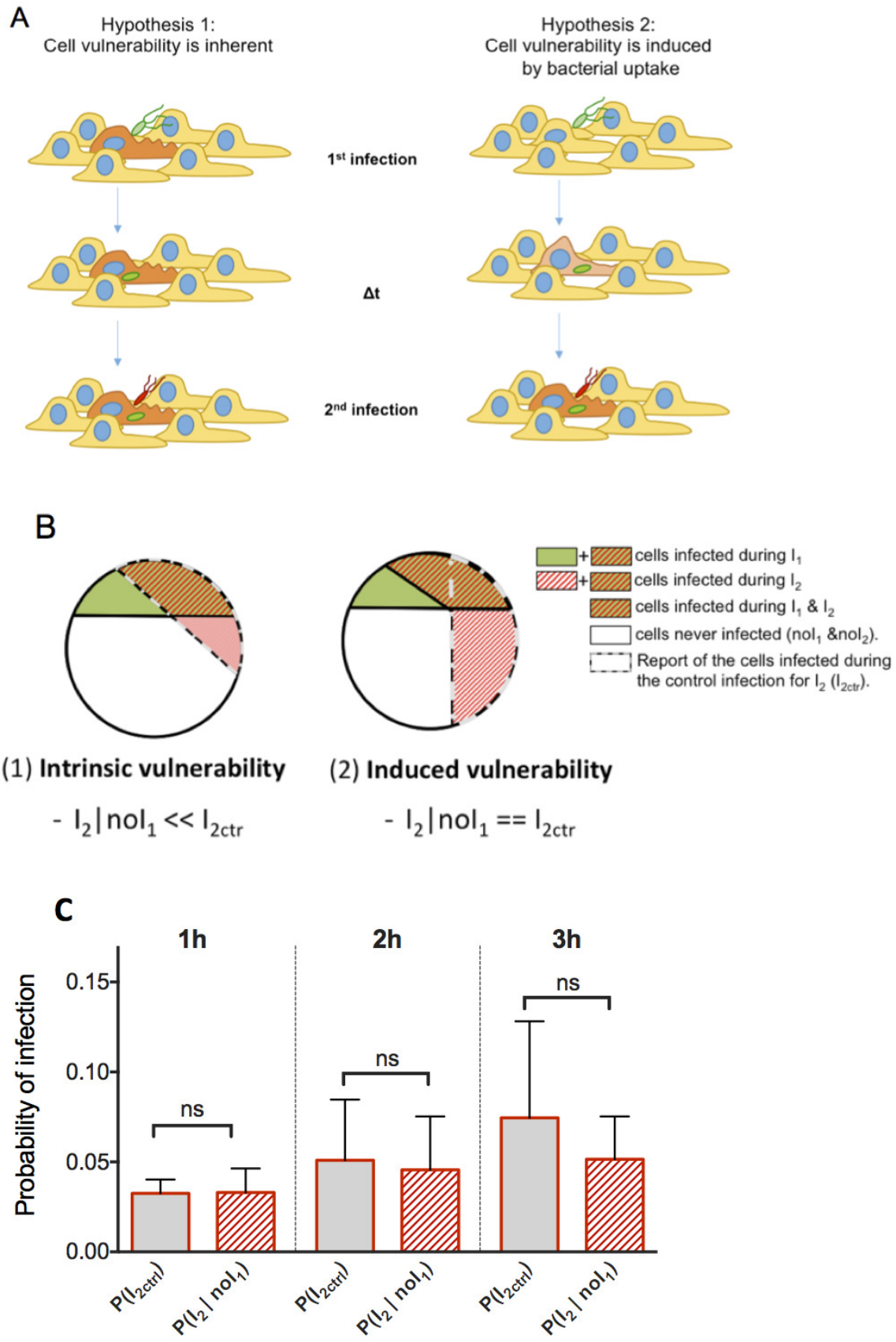
Fig.6. Comparison of model-predicted vulnerability of single-cell with measured-infection. A. Model-predicted probability of infection displayed on reproduced original images of HeLa cells (left panel). Colors are adapted for maximum contrast between lowest (deep red) and highest (white) probability of infection. Measured infections from experiments are shown (top-right panel). **B.** Estimation of the reliability of the two (HeLa and Caco-2) models developed when tested on a total of 100 infected cells and 100 non-infected cells.

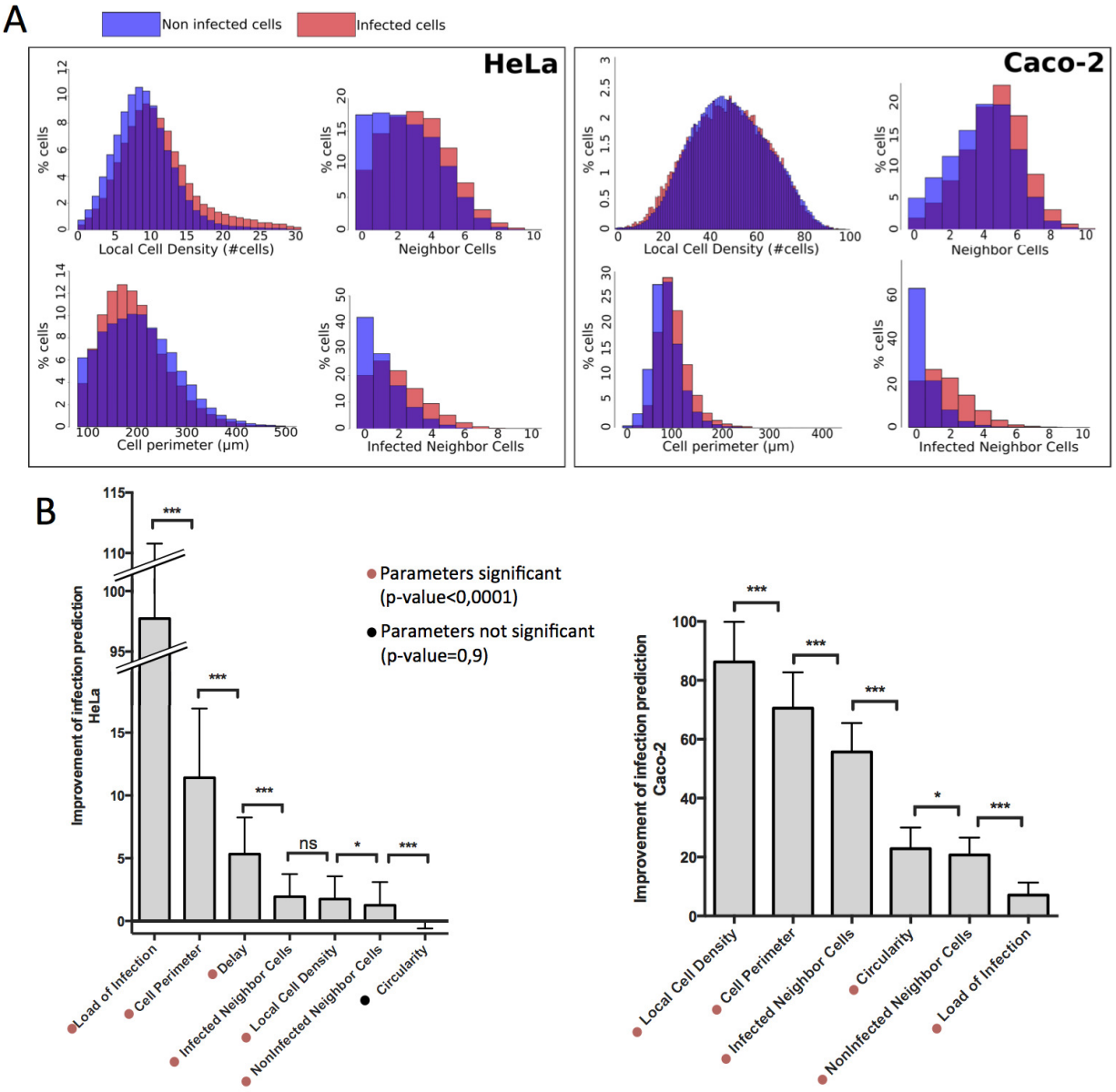
Fig.7. Probability of infection as a function of single cell cholesterol level. A. B. Variation of the probability of *Salmonella* infection at different levels of host cholesterol measured by FACS as described in detail in the text. Cholesterol levels were binned in five categories at 20% steps from lowest to highest levels over the total cell population, each category contains 20% of the whole cells. **A.** Results obtained for HeLa cells (n =3). **B.** Results obtained for Caco-2 cells (n =3).











C

Fold change of the probability of infection

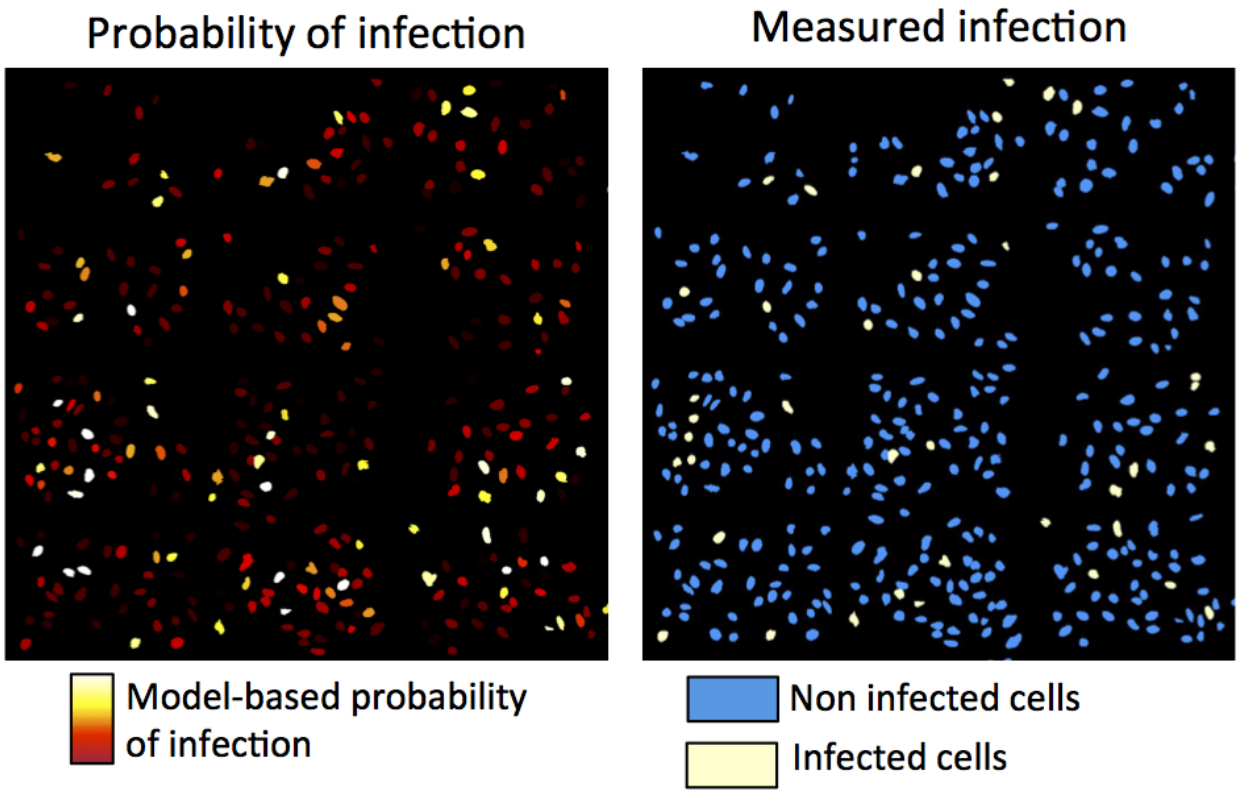
Inherent vulnerability	1st infection	HeLa	Caco-2
low	-	1	1
low	+	2.2	1.3
high	-	1.6	2.6
high	+	3.3	3.3

D

Change in P(inf) when increase in the parameter values

Cell parameters	HeLa	Caco-2
Cell size	↗	↗
Circularity	≈	↘
Local cell density	↗	↘
Number of neighbor cells	↗	↗
Number of infected neighbor cells	↗	↗
Load of infection	↗	↗

A



B

Model predictions	HeLa	Caco-2
True	62%	66%
False	38%	34%

A

B