

AUTOPHAGY

<https://doi.org/10.1080/15548627.2019.1590519>

BRIEF REPORT - BASIC SCIENCE

OPEN ACCESS 

## Targeted interplay between bacterial pathogens and host autophagy

Padhmanand Sudhakar<sup>a,b,c</sup>, Anne-Claire Jacomin<sup>d</sup>, Isabelle Hautefort<sup>a</sup>, Siva Samavedam<sup>d</sup>, Koorosh Fatemian<sup>d,e</sup>, Eszter Arif<sup>f,g</sup>, Leila Gul<sup>a</sup>, Amanda Demeter<sup>a,b,f</sup>, Emily Jones<sup>a,b</sup>, Tamas Korcsmaros <sup>a,b,\*</sup>, and Ioannis P. Nezis<sup>d\*</sup>

<sup>a</sup>Earlham Institute, Norwich Research Park, Norwich, UK; <sup>b</sup>Gut Health and Microbes Programme, Quadram Institute, Norwich Research Park, Norwich, UK; <sup>c</sup>Department of Chronic Diseases, Metabolism and Ageing, KU Leuven, Leuven, Belgium; <sup>d</sup>School of Life Sciences, University of Warwick, Coventry, UK; <sup>e</sup>Current affiliation: Exaelements LTD, Coventry, UK; <sup>f</sup>Department of Genetics, Eotvos Lorand University, Budapest, Hungary; <sup>g</sup>Synthetic and System Biology Unit, Institute of Biochemistry, Biological Research Centre of the Hungarian Academy of Sciences, Szeged, Hungary

### ABSTRACT

Due to the critical role played by autophagy in pathogen clearance, pathogens have developed diverse strategies to subvert it. Despite previous key findings of bacteria-autophagy interplay, a systems-level insight into selective targeting by the host and autophagy modulation by the pathogens is lacking. We predicted potential interactions between human autophagy proteins and effector proteins from 56 pathogenic bacterial species by identifying bacterial proteins predicted to have recognition motifs for selective autophagy receptors SQSTM1/p62, CALCOCO2/NDP52 and MAP1LC3/LC3. Using structure-based interaction prediction, we identified bacterial proteins capable to modify core autophagy components. Our analysis revealed that autophagy receptors in general potentially target mostly genus-specific proteins, and not those present in multiple genera. The complementarity between the predicted SQSTM1/p62 and CALCOCO2/NDP52 targets, which has been shown for *Salmonella*, *Listeria* and *Shigella*, could be observed across other pathogens. This complementarity potentially leaves the host more susceptible to chronic infections upon the mutation of autophagy receptors. Proteins derived from enterotoxigenic and non-toxigenic *Bacillus* outer membrane vesicles indicated that autophagy targets pathogenic proteins rather than non-pathogenic ones. We also observed a pathogen-specific pattern as to which autophagy phase could be modulated by specific genera. We found intriguing examples of bacterial proteins that could modulate autophagy, and in turn being targeted by autophagy as a host defense mechanism. We confirmed experimentally an interplay between a *Salmonella* protease, YhjJ and autophagy. Our comparative meta-analysis points out key commonalities and differences in how pathogens could affect autophagy and how autophagy potentially recognizes these pathogenic effectors.

**Abbreviations:** ATG5: autophagy related 5; CALCOCO2/NDP52: calcium binding and coiled-coil domain 2; GST: glutathione S-transferase; LIR: MAP1LC3/LC3-interacting region; MAP1LC3/LC3: microtubule associated protein 1 light chain 3 alpha; OMV: outer membrane vesicles; SQSTM1/p62: sequestosome 1; SCV: *Salmonella* containing vesicle; TECPR1: tectonin beta-propeller repeat containing 1; YhjJ: hypothetical zinc-protease.

### ARTICLE HISTORY

Received 15 May 2017  
Revised 21 February 2019  
Accepted 1 March 2019

### KEYWORDS

Autophagy; bacterial regulation of host; MAP1LC3/LC3; MAP1LC3/LC3-interacting region motif; microbiota; CALCOCO2/NDP52; SQSTM1/p62; pathogen recognition; interplay


## Introduction

Selective autophagy is a fundamental catabolic process where cytosolic material is specifically engulfed within double-membrane vesicles, known as autophagosomes, and targeted for degradation by lysosomes [1]. Xenophagy is a type of selective autophagy and refers to the selective autophagic degradation of invading bacteria and viruses, and is an important aspect of the host innate immune response to protect against infection [1–3]. Upon bacterial infection, autophagy is induced via multiple host factors and signaling pathways resulting in pathogen encapsulation within double-membrane structures called autophagosomes which ultimately fuse with lysosomes, causing the degradation of the enclosed bacteria [1,4–8].

Various proteins are involved in the autophagy process, with MAP1LC3/LC3 (microtubule associated protein 1 light chain 3 alpha) being among the most prominent [9,10]. MAP1LC3/LC3 is known to have a stable association with the autophagosome membrane, and can recruit cargo for degradation via a conserved LIR (LC3-Interacting Region) motif found in selective autophagy [11] receptor proteins such as SQSTM1/p62 (sequestosome 1), OPTN (optineurin) and CALCOCO2/NDP52 (calcium binding and coiled-coil domain 2) [12–14]. Besides their ability to bind autophagy receptors, MAP1LC3/LC3 proteins have also been shown to interact directly with proteins degraded by autophagy [15]. Pathogen-targeting by the autophagy machinery happens i) through direct protein-protein interactions between selective autophagy receptors and bacterial effector proteins

**CONTACT** Tamas Korcsmaros  [Tamas.Korcsmaros@earlham.ac.uk](mailto:Tamas.Korcsmaros@earlham.ac.uk)  Earlham Institute, Norwich Research Park, Norwich NR4 7UZ, UK

\*Tamas Korcsmaros and Ioannis P. Nezis these two authors are joint corresponding authors

 Supplemental data for this article can be accessed [here](#).

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

[16–19] and ii) also occurs in an ubiquitin-dependent manner wherein the autophagy receptor interacts with and recognizes ubiquitinated cargoes [7,20]. Independently of canonical autophagy, MAP1LC3/LC3 can also be recruited to single-membrane phagosomes in a process called LAP (MAP1LC3/LC3-associated phagocytosis). This process does not depend on ubiquitination of the invading bacteria, but contributes to its degradation by the lysosomes [21].

Conversely, autophagy is often hijacked or manipulated by pathogenic bacteria, such as *Salmonella*, *Shigella* and *Helicobacter spp* [2,3,22,23]. Bacterial pathogens can either modulate their membrane surface, or express effectors and toxins to interfere with autophagy, and hijack autophagy to increase intracellular replication and further invasion potential [2,3,22,23]. Bacteria are able to avoid the autophagic defenses of the host through a variety of methods that can broadly be categorized into 3 different approaches – evasion, inhibition, and subversion. Bacterial evasion of host autophagy means escaping autophagic clearance or avoiding autophagic detection inside the cell. For example, although ATG5 (autophagy related 5) and TECPR1 (tectonin beta-propeller repeat containing 1) target *Shigella* to the autophagosome by binding the bacterial IcsA/VirG protein (Outer membrane protein IcsA autotransporter) [24], *Shigella* secretes IcsB (Virulence protein IcsB) which competitively binds to IcsA and therefore helps the pathogen avoid detection by the host autophagic machinery [25]. Inhibition of autophagy refers to bacteria that are able to arrest autophagy to some degree by interfering with the core autophagy machinery. For example, *Legionella pneumophila* infection in mammalian cells results in cell-wide cessation of autophagy. This ability is conferred by RavZ (*Legionella pneumophila* effector protein), which is a cysteine protease that localizes to phagophores involved in the early stages of autophagosome formation and cleaves MAP1LC3/LC3 family proteins conjugated to lipids [26]. Some other bacteria are able to subvert host cell autophagy to aid their own survival. *Coxiella burnetii* trafficks within a vacuole along the endosome-lysosome pathway in order to replicate in a low pH environment [27,28].

Even though there have been documented reports of specific bacterial pathogen proteins interfering with host proteins during particular phases of autophagy, it is not apparent if pathogens have adopted clearly devised strategies to hijack autophagy, for example by preferentially modulating specific phases to suit their survival and pathological niches. Likewise, a systemic view of which bacterial effectors can potentially modulate autophagy is lacking. To the best of our knowledge, no integrative study has yet been carried out to investigate whether those bacterial proteins that directly affect the host (including modulating autophagy) are specifically targeted by autophagy for degradation. A limited number of computational approaches have been carried out to estimate the bacteria-host interaction networks [29–33]. Scheidel and colleagues for example simulated the effect of *in silico* knock-outs of specific autophagy receptors such as CALCOCO2/NDP52, OPTN and TBK1 (TANK binding kinase 1) on the

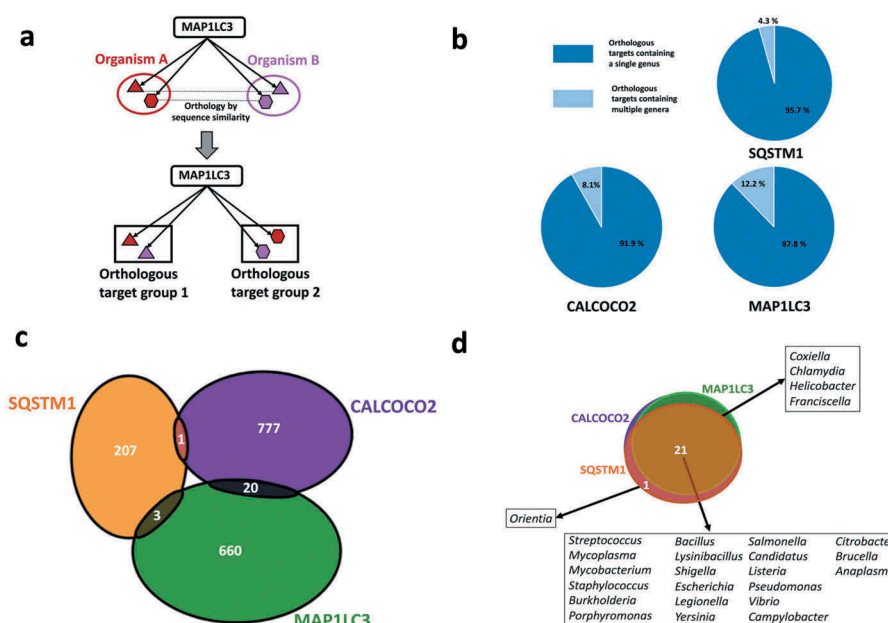
xenophagic capturing of individual pathogens such as *Salmonella* [29]. Budak and colleagues were able to identify key signaling proteins in the host-*Salmonella* interaction network by combining protein-protein interactions and phosphoproteomic data, demonstrating the efficacy of a computational analysis guided validation approach [30–33]. However, such studies were either confined to particular pathogens or were focused only on the unidirectional effects on the host without accounting for complex bidirectional interplays. Recently, Behrends and colleagues carried out a systemic analysis to map the ubiquitinome both in the host and in *Salmonella* to provide a global insight to host-pathogen interplay, specifically mediated by ubiquitination [34].

Motivated by these unanswered questions and to complement existing approaches, we systematically investigated the interplay between potential bacterial proteins and host autophagy. Here, we describe putative interactions between bacterial proteins targeted by autophagy via human orthologs (such as MAP1LC3/LC3) of Atg8-family proteins and the selective autophagy receptors SQSTM1/p62 and CALCOCO2/NDP52 in an ubiquitin-independent manner. Importantly, we provide experimental validation for one of them, the *Salmonella* protease, YhjJ. To understand the potential modulatory effects of bacterial effectors on host autophagy, we also identified specific bacterial proteins which are capable of being post-translationally modified by autophagy proteins. By combining the 2 analyses, we were able to identify a potential subset of bacterial proteins which could not only target autophagy but could also be targeted by autophagy. Specifically identifying the proteins targeted by autophagy receptors will provide a starting point for further study into how bacteria are marked for degradation and therefore how autophagy is triggered. This in turn will enable studies to improve the understanding of how bacteria behave during the course of an infection, how host cells respond, and if any of these mechanisms can be exploited for medical treatments. Comparative studies on different pathogens could identify evolutionarily conserved and strain-specific targeting patterns that may guide therapeutic development strategies.

## Results and discussion

### Autophagy targets genus-specific proteins

The bacterial proteins we analyzed belonged to 56 different strains (Table S1) from 26 distinct genera of bacterial pathogens with known associations with host autophagy. We grouped the bacterial proteins targeted by each of the autophagy targeting proteins (SQSTM1/p62, CALCOCO2/NDP52 or MAP1LC3/LC3) into orthology based clusters (i.e., orthologous targets – Table S2) to comparatively analyze the targeting features of autophagy (Figure 1(a), Table S3). A large proportion of the orthologous clusters potentially targeted by the autophagy proteins contained just 1 bacterial genus (Figure 1(b)) indicating that autophagy could be able to recognize proteins that are genus-specific,



**Figure 1.** Genera and protein specificities of the autophagy receptors SQSTM1/p62, CALCOCO2/NDP52 and the autophagy adaptor protein MAP1LC3/LC3. (a) Definition of orthologous target groups. Orthologous targets are defined as the set of orthologous proteins which share sequence homology with each other and recognized as substrates by a particular autophagy targeting protein. (b) Number of orthologous target proteins of SQSTM1/p62, CALCOCO2/NDP52 and MAP1LC3/LC3 in single and multiple pathogen genera. (c) Comparison of bacterial proteins targeted by SQSTM1/p62, CALCOCO2/NDP52 and MAP1LC3/LC3 indicating that targeting of bacterial proteins by autophagy is mostly complementary. (d) Comparison of the studied bacterial genera targeted by SQSTM1/p62, CALCOCO2/NDP52 and MAP1LC3/LC3 showing a high overlap, which may promote efficient pathogen surveillance.

and not widespread among pathogens. For example, MAP1LC3/LC3 potentially targets ClfA (clumping factor A), which is uniquely found in *Staphylococcus aureus* strains in comparison to the other pathogens in our study. ClfAs are known to contribute to the virulence mechanisms of *S. aureus* by binding to human fibrinogen [35,36] and inhibiting phagocytosis [37]. This suggests autophagy has evolved to possibly target-specific proteins and functions that could be important for the invasive properties of specific pathogens. On the other hand, autophagy targeting proteins could also recognize conserved proteins from various genera of pathogens in a smaller fraction of orthologous groups (Figure 1(b)). An example for this more general targeting strategy are the MAP1LC3/LC3 targeted FlgL (flagellar hook protein) orthologs, present in multiple genera of pathogens. FlgL has a key role in motility, which is important for the systemic invasive properties of many bacterial pathogens [38].

### Complementarity of autophagy targeting proteins potentially increase infection risk upon mutation

Autophagy receptors SQSTM1/p62 and CALCOCO2/NDP52 act independently of each other as demonstrated by the almost extreme exclusivity of their targets (only 2 proteins are targeted by both SQSTM1/p62 and CALCOCO2/NDP52 (Figure 1(c)). This complementarity between SQSTM1/p62 and CALCOCO2/NDP52 has so far only been shown for proteins from *Salmonella*, *Listeria* and *Shigella* [4,39], whereas we substantially extended it for a very diverse set of pathogens, indicating that this mechanism could be in principle observed across genera. Furthermore, MAP1LC3/LC3 shared

a very small number of overlapping targets with SQSTM1/p62 and CALCOCO2/NDP52, suggesting that these 3 autophagy proteins elicit their action in targeting bacterial proteins in an almost independent, complementary manner (Figure 1(c)). These observations suggest the existence of possible selection pressures, which may have influenced the specific recognition features of the autophagy receptors.

While the 3 autophagy proteins target bacterial proteins with little overlap, 21 out of the 26 bacterial genera in our study were commonly targeted by all these 3 autophagy receptors (Figure 1(d)). The remaining five genera were targeted by any 2 of the 3 autophagy proteins (Figure 1(d)). Given the complete lack of overlap among the targets of the 3 autophagy proteins, and the relatively large overlap in terms of the genera, this observation shows that autophagy could potentially target different sets of bacterial proteins from a same pathogen. The redundancy in targeting a same genera is beneficial for robust pathogen recognition, while the complementarity in the specific bacterial proteins is advantageous to increase the surveillance coverage of autophagy.

Nonetheless, the extreme complementarity of the autophagy targeting proteins could make the host more susceptible to chronic disorders and infections if the gene encoding one of the autophagy targeting proteins becomes mutated, and the autophagy system is overloaded or suffers other malfunctions. For example, various studies have highlighted the association between mutations in CALCOCO2/NDP52 and Crohn disease [40] which is an inflammatory bowel disease characterized by an altered gut microbiome. Accordingly, there is an increased vulnerability to infections by foodborne pathogens such as *Salmonella* and

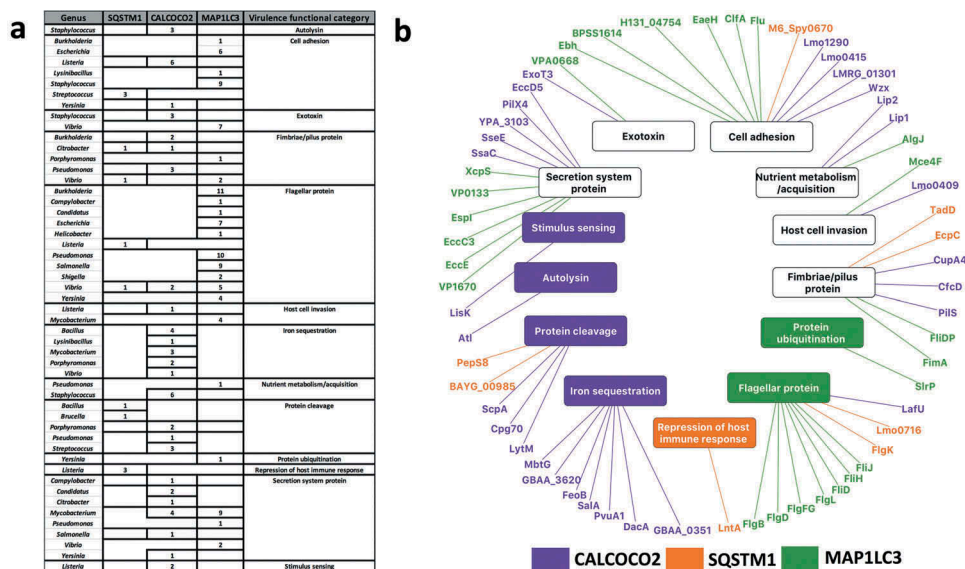
*Shigella* in the intestine of individuals with Crohn disease [4,39,40]. Given the complementarity among the autophagy targeting proteins, the compromised function of CALCOCO2/NDP52 could increase the risks associated with exposure to CALCOCO2/NDP52 targeted pathogens (e.g. *Salmonella*, *Shigella*, *Escherichia* etc). In addition, as mutations of several other autophagy genes, most importantly NOD2 (nucleotide binding oligomerization domain containing 2) [41] and ATG16L1 (autophagy related 16 like 1) [42,43] are associated with Crohn disease, the lack of proper CALCOCO2/NDP52 surveillance could potentially result in even less efficient bacterial degradation in this autophagy deficient background. Similarly, mutations of SQSTM1/p62 have been associated with Paget disease, a bone development related disease that often causes increased risk of infection, and with amyotrophic lateral sclerosis (ALS) causing episodes of lower respiratory tract infections [44,45]. Relevant to both ALS and our observation, the *Orientia* genus is one of the two genera whose proteins are uniquely recognized by SQSTM1/p62. *Orientia* has been detected in the lungs of patients with acute respiratory distress syndrome [46]. These observations could indicate that in the SQSTM1/p62 mutant background of ALS patients, the lack of *Orientia* spp. detection is capable of contributing to this pulmonary disease.

### Autophagy targets virulence factors

Using data from the Virulence Factor Database [47], we observed that autophagy recognizes more virulence factors than would be expected by chance considering all the potential target proteins: i.e. virulence factors were overrepresented (hypergeometric test; P-value<0.05) among the targets of CALCOCO2/NDP52 (0.048) and MAP1LC3/LC3

(2.25e-09) but not in the predicted targetome of SQSTM1/p62. By adding known functions to the targeted virulence factors, and grouping them into 13 distinct categories, we identified commonalities and specificities among the autophagy targeting proteins (Figure 2(a-b); Table S3). Figure 2(b) represents this specific targeting of virulence categories as a network with nodes representing the interacting members and edges the interactions. Out of the 14 functional categories, five specialized virulence factors such as autolysins, and iron sequestering proteins which were potentially recognized uniquely by a single autophagy targeting protein considered in this study (Figure 2(a)). Autolysins which were specifically targeted by CALCOCO2/NDP52 alone have been known to play a major role in infection and contribution to virulence by controlling the bacterial peptidoglycan structure, thereby evading detection by innate immune system receptors such as NOD2 [48]. Conversely, virulence factors responsible for more general functions such as nutrient acquisition and motility (e.g. flagella, fimbriae and pili proteins) were potentially recognized by multiple autophagy targeting proteins (Figure 2(b)). Since most intracellular pathogens rely on motility within the host cell to either establish infection, evade detection or to migrate towards host cell regions with their preferred nutrient source, it is reasonable to expect that general virulence functions such as motility are potentially targeted commonly by all 3 proteins to boost the first line of host defense. Thus, by fine-tuning the recognition of their targets to include generalized as well as specific virulence functions, selective autophagy provides the host an efficient and robust defense mechanism against many pathogens.

We also checked if there is any underlying evolutionary selection pressure on the proteins capable of being targeted by autophagy and their functional relevance in terms of virulence



**Figure 2.** Virulence factor targeting features of SQSTM1/p62, CALCOCO2/NDP52 and MAP1LC3/LC3. (a) Tabular representation of the virulence factor classes targeted by the autophagy receptors and MAP1LC3/LC3. Only those bacterial genera where the targeted bacterial proteins are annotated as virulence factors are listed. (b) A network based view of the virulence factors (outer ring) and the corresponding virulence functions (inner ring) targeted by the autophagy proteins. We colored the virulence functions based on the associations between the autophagy targeting proteins and virulence factors. Whereas some of the virulence functions are targeted by only 1 particular autophagy targeting protein, other functions are targeted by multiple autophagy targeting proteins (in black).

and pathogenicity. To this end, we performed an additional analysis by comparing the sequences of non-pathogenic *E. coli* reference strain K12 to multiple *Salmonella* strain to identify single nucleotide polymorphic (SNP) mutations in the predicted LIR motifs in the *Salmonella* proteins. We identified 18 orthologous groups corresponding to *Salmonella* proteins harboring LIR motifs and having SNPs compared to *E. coli*. 14 of the 18 orthologous groups were associated with virulence functions such as protein cleavage, host cell invasion and iron sequestration (Table S4). This could indicate that the proteins capable of being targeted by autophagy are under selective evolutionary pressure driven by virulence and pathogenicity.

We also identified a key difference between the virulence factor targeting features of CALCOCO2/NDP52, MAP1LC3/LC3 and SQSTM1/p62. As CALCOCO2/NDP52 is already known to specifically potentially recognize ubiquitinated bacteria during xenophagy, and the presented analysis showed CALCOCO2/NDP52 and MAP1LC3/LC3 targets were enriched with virulence factors, it is possible that CALCOCO2/NDP52 and MAP1LC3/LC3 evolved specifically to target pathogens or pathogenic proteins for autophagic degradation [20,21]. On the other hand, SQSTM1/p62, conspicuous by the very small number of virulence factor related functions it targets, is mostly involved in other types of autophagy (such as the removal of internal ubiquitinated protein aggregates). This may suggest that whilst SQSTM1/p62 could target more generic bacterial proteins containing a target motif but not (yet) related to virulence, CALCOCO2/NDP52 and MAP1LC3/LC3 probably evolved to exploit the particular niche of bacterial proteins that SQSTM1/p62 not capable of interacting with, thereby attenuating and improving the host's ability to recognize and remove virulent molecules via autophagy.

To confirm that our observations were not biased by selecting pathogenic bacteria and only the proteins thereof which are accessible to the autophagy targeting machinery, we performed a control analysis with proteins found within bacterial outer membrane vesicles (OMVs) using the same workflow. OMVs are extracellular bodies [49] secreted by pathogenic [50–52] and non-pathogenic bacterial species including probiotics and commensals [53–55]. OMVs are engulfed by host cells upon endocytosis and due to this mode of action, the OMV contents including proteins can be spatially accessed by the host machinery [54]. We used a recent study [56], which listed the proteins from OMVs of pathogenic (enterotoxigenic) as well as non-pathogenic (nontoxigenic) strains of *Bacillus fragilis*. From the analysis, we found an overrepresentation of autophagy targeted proteins: from 140 enterotoxigenic OMV proteins 33 could potentially be targeted by the autophagy receptors SQSTM1/p62, CALCOCO2/NDP52 and the autophagy adaptor protein MAP1LC3/LC3 (hypergeometric test P-value 0.039) (Table S5). The results from this control analysis using pathogenic and non-pathogenic bacterial proteins known to be localized in the host cytosol and potentially affect the host processes including autophagy suggest that pathogenic proteins could be selectively targeted by autophagy to a greater extent than non-pathogenic proteins.

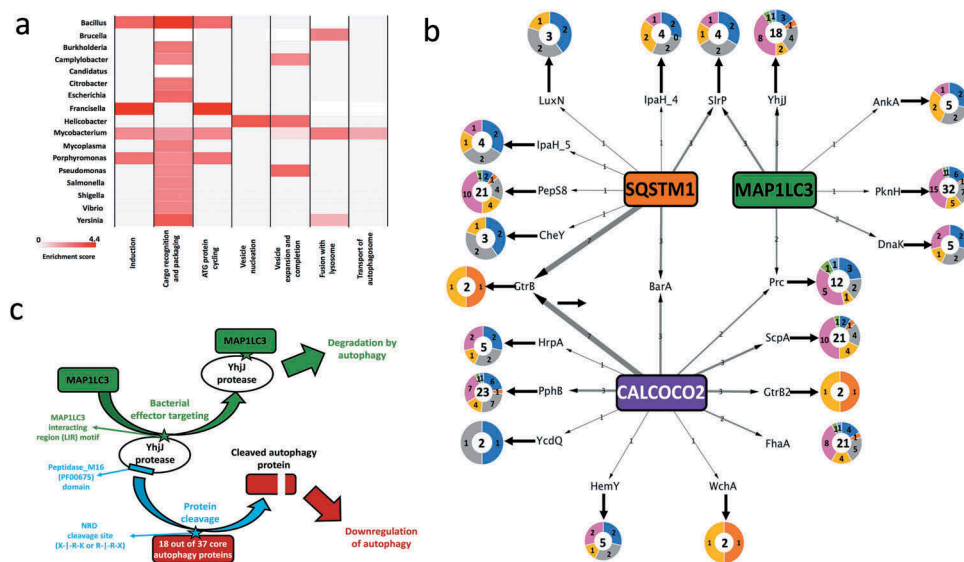
### Phase-specific regulation of autophagy by pathogens

The previous sections outlined how autophagy targets bacterial effector proteins. But this is only one side of the story as bacterial proteins from various pathogenic genera are also potentially able to modulate autophagy [2,3,22,23]. We observed genus-specific patterns as to which phases of autophagy (e.g., induction, autophagosome formation, fusion with lysosome) are potentially regulated by a given pathogen group (Figure 3(a)). For example, from the over-representation analysis of bacterial proteins predicted to modulate autophagy proteins (Table S6), we observed that *Bacillus spp.* have propensity to putatively modulate multiple early phases of autophagy ('Induction', 'Cargo recognition and packaging' and 'ATG protein cycling' phases) than the further downstream phases. This ties in well with previous observations that *Bacillus anthracis* inhibits autophagy initiation via a cyclic AMP concentration-promoting adenylyl cyclase [57]. Our analysis suggests that the *Bacillus* genus also potentially modulates the autophagy induction phase by deploying effectors such as extracellular proteases and serine-threonine kinases to modify the post-translational states of key autophagy proteins, and subvert host defenses.

We noted that some autophagy phases could uniquely be modulated by particular pathogens (e.g., the 'autophagosome-lysosome fusion' phase uniquely by *Mycobacterial*, *Bacillus* and *Yersinia* species only), while some phases were modulated by multiple pathogen genera (e.g., the 'Cargo recognition and packaging' phase) (Figure 3(b)). Even though it is known that *Mycobacterial* species such as *M. tuberculosis* H37v accumulate in non-acidic or non-degradative autophagosomes during infection [58], the actual mechanisms responsible for this modulatory relationship have not yet been explained. From our computational analysis, we observed that multiple *Mycobacterial* species could potentially interfere and block the fusion of the autophagosome and lysosome via the action of various phosphate-group removing hydrolases as well as peptidases on TECPR1, which facilitates autophagosome maturation and autophagosome-lysosome fusion [59]. In mice, TECPR1 is a binding partner for ATG5 and colocalizes with ATG5 at *Shigella*-containing phagophores [24]. TECPR1 activity was also specifically found to be necessary for selective autophagy targeting bacteria and not for the more generic stress induced autophagy [24]. In this context, it is intriguing that intracellular pathogens such as those from the *Mycobacterial* genus could putatively regulate selective autophagy proteins like TECPR1, which is needed for the maturation of the autophagosome and the autophagosome-lysosome fusion [59]. The modulation of TECPR1 as inferred by our analysis provides clues that explains the underlying mechanisms contributing to the *Mycobacterial* induced blockage of autophagosome-lysosome fusion.

### Autophagy strikes back

To investigate co-evolutionary, interplay strategies, we looked for the overlap between bacterial proteins predicted to targeted by autophagy and autophagy proteins potentially modulated by the same bacterial protein. We were able to identify



**Figure 3.** Pathogenic modulation of autophagy and the bi-directional interplay **(a)** Heat map showing the phase-specific regulation of autophagy by various bacterial genera. Hypergeometric distribution was used to determine the over-representation of proteins from each bacterial genus in our study and show those predicted (based on domain-domain and domain-motif interactions) to modulate proteins functioning in 1 or multiple phases of the core autophagy process. The significance score is determined as the  $-\log_{10}$  function of the corrected hypergeometric distribution based enrichment P-value. **(b)** Interplay between autophagy receptors and their target bacterial effectors, which regulate different phases of autophagy. The donut plots display the phase classification of the core autophagy proteins targeted by each bacterial effector protein, and the total number of host autophagy proteins targeted by the bacterial effector is indicated by the bold number within the donut plot. The thickness of the arrows from SQSTM1/p62, CALCOCO2/NDP52 and MAP1LC3/LC3 denote the number of orthologs of the targeted bacterial effector. **(c)** An example of the interplay between host autophagy and the protease Yhj from *Salmonella typhimurium* SL1344.

67 such bacterial proteins (out of the 2119 bacterial proteins predicted to modulate autophagy) (Figure 3(b); Table S7). The low number of such special proteins is expected given the complex and co-evolutionary processes required to meet the strict selection criteria we applied. In support of this, we found that the 67 proteins were derived from 18 bacterial genera (out of the 26 examined), showing a fairly widespread strategy with a notable level of conservation (the 67 proteins were grouped into 40 orthologous groups). Most of the bacterial-autophagy interplays involved the recognizing function of either CALCOCO2/NDP52 or MAP1LC3/LC3. Analyzing the functions of the interplay-related bacterial proteins, compared to all the proteins targeted by autophagy in our study, indicated that the former has unique characteristics. Some of the interplay-related bacterial proteins have functions specifically related to proteolysis (YhjJ, ScpA [C5a peptidase]), phosphorylation (PknH [Serine/threonine-protein kinase PknH], PphB [Serine/threonine-protein phosphatase 2]), and post-translational effects such as transfer of ubiquitin (SlrP [E3 ubiquitin-protein ligase]) moieties, signifying that they are capable of successfully rewiring the host autophagy mechanisms. Due to their potential modulatory effects on autophagy, the interplay-related proteins could in turn be selectively targeted by autophagy for degradation as a counter-measure. 23 of the 40 orthologous groups represented proteins with proteolytic and post-translational activity such as phosphorylation and ubiquitination which can interfere with the activity of autophagy proteins, and thereby promote autophagy inhibition as shown previously, for example with the MAP1LC3/LC3 degrading function of the *Legionella* effector protease RavZ [26]. We also observed that as many as 7 of the orthologous groups involved in the interplay could potentially

modulate a substantial portion of the autophagy machinery (i.e., at least 16 of the 37 core autophagy proteins). Thus, by targeting the bacterial effectors, which are capable of modulating a major part of the autophagy machinery, autophagy protects itself from these effectors as well as strengthening the overall surveillance in the host.

A relevant new example of a bacterial effector involved in the bacteria-autophagy interplay is YhjJ, which is a zinc-protease from the food-borne pathogen *Salmonella enterica* serovar Typhimurium SL1344 and was found to be expressed under conditions when *Salmonella* infects host cells [60]. YhjJ has an M16 peptidase domain (PFAM ID: PF00675), which can recognize the N-Arg dibasic (NRD) convertase cleavage site (X-|R-K or R-|R-X). We found this NRD cleavage site to be present in multiple core autophagy proteins distributed across all the seven phases of autophagy. Compared to the proteins in the whole human proteome harboring the NRD convertase cleavage site, we found a significant enrichment of core autophagy proteins possessing this cleavage site (hypergeometric distribution P-value  $7.7E-03$ ). The significant enrichment means that YhjJ could have a specificity with its M16 peptidase domain in acting on autophagy proteins. The M16 peptidase domain of YhjJ is similar to the M16 peptidase domain of insulysin and nardilysin and other metallopeptidases found in different bacteria and in eukaryotic parasite species to have protease activity on host proteins [61,62]. Nardilysin in *Helicobacter felis* for example regulates gastric inflammation upon infection by cleaving host proteins on their NRD cleavage sites confirming the physiological role of this interaction [63]. In agreement with the potential harmful role of the *Salmonella* YhjJ on the host, we found with the *in silico* interaction prediction approach that MAP1LC3B/LC3B

can bind to YhjJ, potentially as a self-defense mechanism (Figure 3(c)).

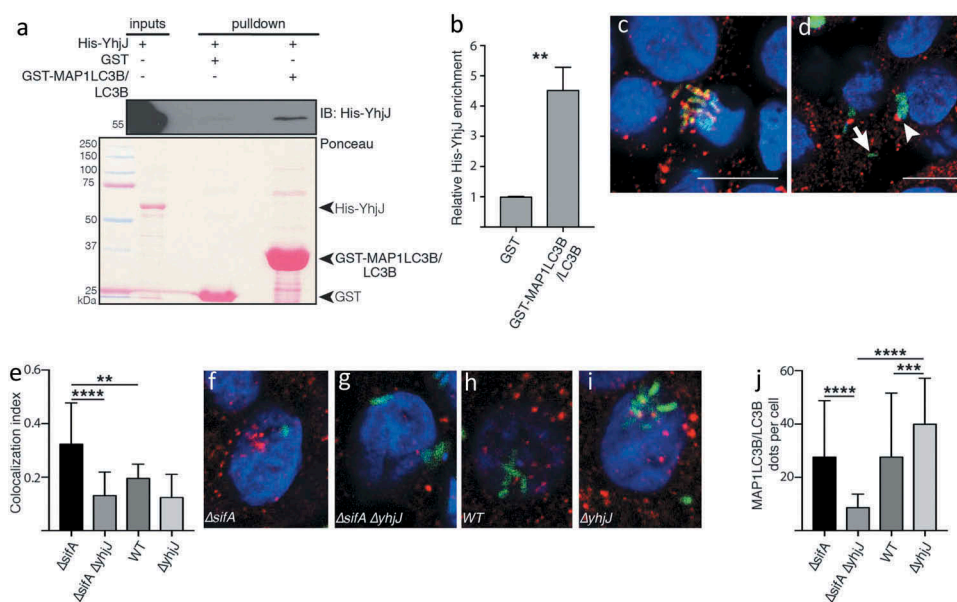
To validate whether the *Salmonella* YhjJ protease via its LIR motif (Figure S1) indeed interacts with MAP1LC3B/LC3B, we performed a GST affinity-isolation assay between recombinant GST-MAP1LC3B/LC3B and 6xHis-YhjJ from *Salmonella* Typhimurium expressed in *E. coli*. We observed that 6xHis-YhjJ is significantly enriched in the affinity isolation with GST-MAP1LC3B/LC3B compare to GST alone (Figure 4(a,b); Table S8), thus confirming that MAP1LC3B/LC3B binds YhjJ. To confirm the functional importance of YhjJ protease in autophagy modulation, we infected HT-29 human epithelial cells with various bacterial strains derived from *S. Typhimurium* SL1344 for 6 h. As expected, the deletion of the *sifA* gene increased the frequency at which *S. Typhimurium* escapes the intracellular *Salmonella*-containing vacuole (SCV), hence raising the likelihood of the secreted YhjJ to encounter the autophagy machinery and reflecting what normally happens to up to 20% intracellular WT *Salmonella* cells during infection of epithelial cells [64,65]. In the absence of YhjJ in the cytosolic *Salmonella* strain (TK0024,  $\Delta sifA\Delta yhjJ$ ), significantly fewer *Salmonella* cells colocalized with MAP1LC3B/LC3B compared with a single  $\Delta sifA$  mutant (TK0019), suggesting that YhjJ is affecting autophagy (Figure 4(c-e); Table S8). Furthermore, the reduction of MAP1LC3B/LC3B recruitment to the bacteria was correlated with a significant decrease in MAP1LC3B/LC3B-positive structures per cell (Figure 4(f-j); Table S8), and we also observed an increase in SQSTM1/p62-positive structures (Figure S2) confirming YhjJ interferes with autophagy regulation when *Salmonella* is in the host cell cytosol.

Similar example is the C5a-peptidase (ScpA), whose orthologs are present in 3 different strains of *Streptococcus pyogenes*, and known to be required to establish infection and promote virulence in mice [66]. Using domain-domain and domain-motif predictions, we predicted that the C5a-peptidase *via* its Subtilase peptidase-domain (PF0082) could potentially cleave several core autophagy proteins including ATG5, which has been reported to be essential for host cells to defend themselves from *Streptococcus* infections [67]. We also predicted that C5a-peptidase could putatively be targeted by CALCOCO2/NDP52, suggesting that the host autophagy machinery could target this particular bacterial protein involved in autophagy modulation (Figure S3).

These and several other interplays (listed in Table S7) could serve as potential examples for the complex evolutionary arms race between virulence-associated traits of bacterial pathogens and host defense mechanisms, in particular autophagy based bacterial targeting.

## Conclusion

We present here an interaction network analysis on the complex interplay between bacterial proteins and the autophagy machinery, highlighting potential ubiquitin independent bi-directional connections. We showed that host autophagy proteins could be capable to recognize genus-specific bacterial proteins, indicating that they could potentially have evolved to target specific bacterial proteins. The 3 autophagy targeting proteins we examined have very little overlap for their bacterial protein targets but they often overlap in targeting the same bacterial genus. Still, there are a number of bacterial



**Figure 4.** *Salmonella* YhjJ protease interacts with MAP1LC3B/LC3B (a) GST affinity-isolation assay between recombinant GST-MAP1LC3B/LC3B and His-YhjJ. Upper panel: immunoblot against 6xhistidine-tagged YhjJ; lower panel: Ponceau S staining. (b) Quantification of the enrichment of His-YhjJ based on 3 independent replicates. (c, d) Illustration of events of complete co-localization (c) adjacent localization (d, arrowhead) or no co-localization (d, arrow) between GFP-tagged *S. Typhimurium* (green) and MAP1LC3B/LC3B (red). Nuclei are stained with DAPI (blue). Scale bars: 10  $\mu$ m. (e) Quantification of the 'co-localization index' (see Materials and Methods section). (f-i) Representative single HT-29 cell pictures from cells infected with  $\Delta sifA$  (f)  $\Delta sifA\Delta yhjJ$  (g) wild type (h) or  $\Delta yhjJ$  *S. Typhimurium* (i). (j) Quantification of the number of MAP1LC3B/LC3B dots per individual cell. Bar charts show mean  $\pm$  s.d. Statistical significance was determined using Students' t-test (b) or one-way ANOVA (e, j), \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ . For a full description of the statistics, refer to Table S8.

pathogens that are probably recognized only by a specific autophagy protein. For the cases of these interactions, if the given autophagy protein is mutated, the host could be more susceptible to chronic disorders associated with infection (such as in Crohn disease or ALS). As virulence factors were shown to be mostly targeted by CALCOCO2/NDP52 and MAP1LC3/LC3 but not SQSTM1/p62, this suggests that SQSTM1/p62 preferably and potentially targets generic bacterial proteins, and that CALCOCO2/NDP52 and MAP1LC3/LC3 evolved to target those specific pathogenic virulence proteins that SQSTM1/p62 is unable to recognize. To offset any biases caused by our selection criteria focusing only on pathogens and the proteins thereof selected on the basis of their spatial localization and access to the host machinery, our control analysis performed using outer membrane vesicle-derived proteins from enterotoxigenic (pathogenic) and non-toxigenic (non-pathogenic) strains of *Bacillus fragilis* suggest that autophagy has a greater propensity to target pathogenic proteins than non-pathogenic ones. Pathogens are known to modulate various phases of autophagy as a defense mechanism and for their own survival. Interestingly, some autophagy phases are modulated specifically by only a few pathogens, while other phases, such as ‘cargo recognition and packaging’ are modulated by multiple genera, highlighting both the genus-specific relevance of certain autophagy phases and the general importance of core autophagy processes, respectively. We propose this mechanism as a potential evolutionary arms-race since some of the bacterial proteins that could potentially modulate autophagy were also targeted by autophagy proteins as a self-defense mechanism for the host. In vitro testing with the *Salmonella* Typhimurium YhjJ zinc protease, known to be produced and secreted during infection, validated the robustness of our computational pipeline. We confirmed that MAP1LC3B/LC3B can bind in vitro YhjJ. In the wt background, upon infection, most *Salmonella* reside in SCVs. However, a small fraction of *Salmonella* escape from the SCVs to the host cell cytosol where they hyperproliferate [68] and disseminate to new hosts. We speculate that this cytosolic fraction of *Salmonella* secretes YhjJ protease that potentially cleaves ATG proteins (as we predicted) and inhibits autophagy. Therefore, in the wild-type background,  $\Delta yhjJ$  mutants would exhibit increased autophagy. On the other hand, in  $\Delta sifA$  mutants, most of *Salmonella* are cytosolic and this could trigger autophagy massively. In this case, we speculate that YhjJ via its interaction with MAP1LC3B/LC3B, may facilitate autophagic degradation of *Salmonella*, therefore reduced anti-*Salmonella* autophagy is observed in *sifA/yhjJ* double mutants. Therefore, this dual proposed role of YhjJ in autophagy may rely on the proportion of *Salmonella* bacteria in the cytosol. Overall, our systems-level analysis has highlighted the complex interplay between host autophagy and bacteria to inspire future experimental studies to elucidate the detailed molecular mechanisms of autophagy in the pathogenesis of bacterial infections.

Recognition of bacterial cargoes by the various autophagy receptors involves a number of complex interactions based on various post-translational signatures, the most common of which is ubiquitination. The bi-directional ubiquitome has been characterized on a systemic basis within the context of

infections by particular pathogens such as *Salmonella* [34]. We complemented this approach with a systemic and computational analysis of ubiquitin independent mechanisms driving bacterial proteins targeted by autophagy receptors via recognition motifs, and autophagy proteins modulated by interactions with pathogen proteins. As we were able to correlate systemic patterns with rational biological explanations, we could demonstrate which interactions are unlikely to occur, at least through direct protein-protein interactions. Thus, in this context, although our approach does not cover ubiquitination mediated autophagy or autophagy receptors other than SQSTM1/p62 and CALCOCO2/NDP52, the identified putative targets in our study provide a novel insight for the range of substrates for the 2 autophagy receptors studied herein. We have to note however, that these proteins and others not identified as putative targets could still either be possible targets of ubiquitination enzymes, which mark their substrates for ubiquitin-dependent autophagic clearance mechanisms, or for other autophagy receptors like NBR1 [69].

Because autophagy is a major upstream modulator of the inflammasome, the findings we present here could shed new light on how pathogens indirectly evade the immune system by subverting autophagy. Our observations also provide plausible mechanistic explanations for increased infection susceptibility related with disorders such as Crohn disease. Validation results emanating from this study will potentially boost the development of novel pharmacological agents such as immunomodulators and autophagy modulating drugs which will be especially important in the current era of antibiotic resistance [70].

## Materials and methods

### *Proteomes and inferring protein localizations*

We analyzed 56 pathogenic bacterial species (Table S1) from 26 different genera with known associations to autophagy based on literature evidence. The proteomes of the selected organisms were retrieved from Uniprot (accessed May 05 2016). The bacterial proteins were filtered for their spatial accessibility (for the autophagy receptors to target the bacterial proteins and for them in turn to be able to physically interact with the core autophagy proteins via their corresponding motifs) by exploiting the cellular localization profiles. Only secreted, extracellular, outer-membrane and surface proteins were included by integrating protein annotation information from the Gene Ontology cellular localization terms, already existing cellular localization profiles and trans-membrane and signal-peptide predictions retrieved from Uniprot [71]. In addition to the annotation based profiling, de novo localization prediction was performed using the PSORTdb standalone prediction software [72] in order to assign localization for proteins without existing localization annotations. For all further analysis, only the bacterial proteins shortlisted after the spatial filtering process were used. Information on virulence factor annotation and protein classification into the various virulence factor categories was retrieved from the VFDB resource [47].



### Protein-protein interaction predictions

The consensus binding motifs of the autophagy receptors SQSTM1/p62 [73] and CALCOCO2/NDP52 [74] were used to scan the bacterial proteins in order to predict their putative targets. The presence of functional LIR motifs in bacterial proteins was examined using the iLIR software (<https://ilir.warwick.ac.uk>). The predictions for MAP1LC3/LC3 interacting proteins is based on the presence of putative functional canonical LIR motifs [75,76]. The predicted LIR motif-containing proteins have been shown to interact with all Atg8-family proteins in human (MAP1LC3A/LC3A, MAP1LC3B/LC3B, MAP1LC3/LC3C, GABARAP, GABARAPL1, GABARAPL2). However, since MAP1LC3C/LC3C has been shown to interact with non-canonical LIR motifs [12], iLIR predictions do not apply for MAP1LC3C/LC3C or any interaction that depends on a non-canonical LIR motif. iLIR uses structural filtering based on disordered region inferences to filter the LIR motifs thus yielding the putative MAP1LC3/LC3 targets. These ‘disordered regions’ are stretches of residues in the protein sequence which do not correspond to a well-defined three-dimensional structure but are still functional. These regions are generally characterized by the lack of hydrophobic amino acids which make up hydrophobic cores and structured domains and they expose short 3–10 residue stretches of amino acids called motifs which mediate the interaction with other protein domains. By virtue of these properties, disordered regions actively participate in a range of protein functions and are often subjected to post-translational modifications, thus broadening the number of functional states in which the protein can exist within the cell [77].

Domain-domain and domain-motif prediction tools as described elsewhere [78,79] were used to identify bacterial proteins which have the potential to interact with and modulate host core autophagy proteins which also includes the selective autophagy receptor SQSTM1/p62 and the autophagy adaptor protein MAP1LC3/LC3. These methods enable the detection of potential post-translational modifications mediated between domains and motifs present in the microbial and host proteins. The reference set of all known and validated domain-domain and domain-motif pairs were retrieved from the DOMINE [80] and ELM [81] databases. In brief, if a bacterial protein harbors a domain which is already known to interact (from the gold set) with a domain or motif within a host autophagy protein, then that particular bacterial protein is considered to post-translationally modulate the host autophagy protein. The bacterial proteins described were derived from the earlier spatial filtering process. The list of core autophagy proteins was retrieved from the Autophagy Regulatory Network resource (<http://autophagyregulation.org>) [82].

Information about the stage-wise classification of autophagy proteins was retrieved from [83]. Orthology grouping was performed using InParanoid [84].

The 2 sets of predictions (bacterial proteins which are capable of being targeted by host autophagy receptors or MAP1LC3 and the bacterial proteins which could potentially modulate the activity of host autophagy proteins) described above were annotated based on the structural features such as disordered regions and globular domains [77] of the proteins involved. We then utilized information on disordered regions to highlight the probable motifs that could be modified and accessed by the domains of other proteins. Disordered region predictions for the proteins of interest were generated using the standalone version of IUPRED [85,86]. We used the InterProScan tool [87] to obtain *de novo* domain annotation for the proteins of interest. All custom codes for the analysis were written in Python.

Interactions for the experimental validation were short-listed based on several criteria: (1) the bacterial protein should be involved in bi-directional interactions – in other words, the bacterial protein needs to be a potential target for 1 of the autophagy receptors and/or MAP1LC3/LC3 and be predicted to interact and post-translationally modify host core autophagy proteins; (2) in the case of MAP1LC3/LC3 targets, the bacterial protein needs to have an iLIR score of at least 18 as a higher iLIR score indicates increased significance of the predicted target; (3) the corresponding motifs (involved in the bi-directional interactions) to be within disordered regions; (4) experimental evidence for expression of the bacterial protein under infection conditions; (5) experimental evidence for the pathogenic protein to be in the bacterial secretome.

### Cloning *yhjJ* in pET28a(+)

The *yhjJ* gene, encoding an M16 zinc protease, from *Salmonella* Typhimurium was produced by PCR amplification from gDNA from *S. Typhimurium* strain SL1344. To clone *yhjJ* into pET28a(+) plasmid (Novagen, Merck Millipore, 69,864), we designed the *yhjJ\_F* and *yhjJ\_R* forward and reverse PCR primers (Table 1.docx). The forward primer incorporates the *Bam*HI restriction site, while the reverse primer includes *Not*I site (underlined). The reverse primer also includes a TAA stop codon before the restriction site. The PCR product was cloned into pET28a(+) vector using *Bam*HI and *Not*I enzymes and subsequent ligation. The resulting pET28-YhjJ plasmid was checked by sequencing. The resulting plasmid constructs were cloned into NEB® 10-beta Competent *E. coli* cells (New England Biolabs, C3019).

**Table 1.** Oligonucleotides used in this study.

Name	Sequence	Use in this study
3578delF	5'-GCTGTCTTTTATTACCAGGATTGTTGATCAGGGTTACGtgtaggctggagctgcttc-3' <sup>a</sup>	Construction of gene deletion mutant
3578delR	5'-GCCCGGTGGCGCTGCGCTTACCGGGCCGACAGGCGGCAGCcatatgaatcctccttag-3' <sup>a</sup>	Construction of gene deletion mutant
sifA_RedF2	5'-ATTATGTAGTCATTTTACTCCAGTATAAGTGAGATTAATcatatgaatcctccttag-3' <sup>a</sup>	Construction of gene deletion mutant
sifA_RedR2	5'-TAAACCTGAACGTGACGTCTGAGAAAGCGTCGCTGATTgtgagctggagctgcttc-3' <sup>a</sup>	Construction of gene deletion mutant

<sup>a</sup> Uppercase sequences indicate homology with the flanking regions of the target gene

### Protein expression and GST affinity isolation

pET28a-YhjJ, pGEX-4T1 (GE Healthcare Life Sciences, 28954549) and pGEX-4T1-MAP1LC3B/LC3B (gift from Dr Rob Layfield, School of Life Sciences, University of Nottingham) were transformed into Rosetta2(DE3) pLacI competent cells (Novagen, Merck Millipore 71404) that were grown on plates supplemented with kanamycin and chloramphenicol (pET28a-YhjJ) or ampicillin and chloramphenicol (pGEX-4T1 and pGEX-4T1-MAP1LC3B/LC3B). For expression, overnight cultures were diluted 1:100 in fresh LB medium, grown to OD<sub>600</sub> = 0.6 at 37°C, and induced with 0.5 mM Isopropyl β-D-1-thiogalactopyranoside. Cells were harvested by centrifugation after 3 h at 37°C (pET28a-YhjJ) or 16 h at 2°C (pGEX-4T1 and pGEX-4T1-MAP1LC3B/LC3B), and lysed by sonication in Classic Lysis Buffer (CLB; 25 mM Tris, pH 7.4, 100 mM NaCl, 2 mM EDTA, 0.05% beta-mercaptoethanol supplemented with a protease inhibitor cocktail (Roche, 5892791001). Following clarification of the lysates by centrifugation at 14310.4 g for 20 min at 4°C (Beckman centrifuge, JA-20 rotor), GST-fusion proteins were purified on glutathione-Sepharose 4 Fast Flow beads (GE Healthcare, 17-5132-01) for 30 min at 4°C. Beads were rinsed with at least 10 beads volumes of High Salt Wash Buffer (25 mM Tris, pH 7.4, 500 mM NaCl, 2 mM EDTA) and Low Salt Wash Buffer (25 mM Tris, pH 7.4, 50 mM NaCl, 2 mM EDTA) before incubation 6xHis-YhjJ for 2 h at 4°C. The beads were washed 4 times with CLB supplemented with 10 mM imidazole and boiled in 2X Laemmli loading buffer for 5 min.

Cell lysates and GST affinity-isolation eluates were subjected to SDS-PAGE (10% polyacrylamide gel) followed by transfer onto nitrocellulose membrane for 60 min at 100 V. Membranes were stained with Ponceau Red (Sigma-Aldrich, P3504) before blocking with 5% non-fat milk in TBST (0.1% Tween-20 [Sigma-Aldrich, P9416] in TBS (Tris-buffered saline; 20 mM Tris and 150 mM NaCl) for 1 h. Primary antibody anti-His tag (Abcam, ab18184; 1:5,000) was incubated for 2 h in TBST, followed by incubation with HRP-coupled secondary goat anti-rabbit antibody (Thermo Scientific, NE171565) for 45 min in TBST+1% non-fat milk. All antibodies incubations were performed at room temperature with gentle agitation. The membrane were developed by chemiluminescence (Amersham ECL Reagents; GE Healthcare Life Sciences, RPN2134).

**Table 2.** Bacterial strains used in this study.

Strains	Description	Source or Reference
SL1344	4/74 <i>hisG rpsL</i>	Ref [88]
JH3009	SL1344 $\phi(ssaG^-gfp^+)$ , Cm <sup>R</sup>	Ref [89]
TK0016	JH3009 $\phi(ssaG^-gfp^+)$ , $\Delta yhjJ$ , Km <sup>R</sup> , Cm <sup>R*</sup>	This study
TK0017	SL1344 $\Delta yhjJ$ , Km <sup>R</sup>	This study
TK0018	SL1344 $\Delta sifA$ , Km <sup>R</sup>	This study
TK0019	JH3009 $\phi(ssaG^-gfp^+)$ , $\Delta sifA$ , Km <sup>R</sup> , Cm <sup>R*</sup>	This study
TK0021	SL1344 $\Delta sifA$ , Km <sup>S</sup>	This study
TK0024	JH3009 $\phi(ssaG^-gfp^+)$ , $\Delta sifA \Delta yhjJ$ , Km <sup>R</sup> , Cm <sup>R*</sup>	This study

\*  $\phi$  indicates transcriptional fusion. The  $\phi(ssaG^-gfp^+)$  fusion is inserted in the *putPA* chromosomal locus as described before [89].

### Bacterial strains and growth conditions

The *S. typhimurium* SL1344 (initially obtained from Catherine Lee [88]) and JH3009 strains used in this study were kindly provided by Jay Hinton [89]. All strains are listed below in Table 2. Strains were grown in 25 ml of LB broth [90] at 37°C, unless stated otherwise. Cultures were shaken in 250-ml conical flasks at 2.23 g. For invasion assays, a 1 in 100 dilution of a 5 ml overnight bacterial culture was grown in 25 ml of LBS (LB containing a total of 0.3 M NaCl) in 250-ml conical flasks until an optical density at 600 nm (A<sub>600</sub>) of 1.2. Antibiotics were added as required at the following final concentrations (ampicillin, 100 mg ml<sup>-1</sup>; kanamycin, 50 mg ml<sup>-1</sup>; chloramphenicol, 10 mg ml<sup>-1</sup>).

### Bacterial gene deletion and replacement

The *yhjJ* and *sifA* gene coding sequences were independently deleted from the *S. Typhimurium* strain SL1344 chromosome. Briefly, for each deletion a kanamycin resistance determinant was amplified by PCR from the pKD4 plasmid template using primers listed in Table 1 [91]. Each of these primers includes at its 5' ends a 40 base-long extension showing homology with the flanking regions of the respective target gene. The generated PCR products were used to replace the structural target genes on *S. Typhimurium* chromosome using the Lambda Red recombination system [91]. The TK0017 and TK0018 recombinant strains generated were selected for antibiotic resistance, verified by analytical PCR and transduced back into a clean SL1344 background using P22 phage transduction to avoid possible non-intended recombination events [92] (Table 2).

The kanamycin resistance cassette replacing the *sifA* gene in TK0018 was excised from the chromosome using the yeast F<sub>1</sub> recombinase expressed from the thermosensitive replicon pCP20 [93] generating the strain TK0021. The pCP20 replicon was subsequently removed from TK0021 after culture at 40°C in non-selective medium. The *yhjJ*, Km<sup>R</sup> deletion and the  $\phi(ssaG^-gfp^+)$ , Cm<sup>R</sup> transcriptional fusion were transduced into TK0021 by P22 phage transduction, generating the strain TK0024 used in this study (Table 2).

### Epithelial cell invasion assay

HT-29 intestinal epithelial cells (ATCC<sup>®</sup>, HTB-38<sup>™</sup>) were cultivated in complete medium composed of DMEM (Lonza, BE12-614F) supplemented with 10% heat-inactivated fetal bovine serum (Labtech, FCS-SA) and 2 mM L-glutamine (Lonza, BE17-605E) in a humidified incubator at 37°C, 5% CO<sub>2</sub>. Between 2 and 5 × 10<sup>5</sup> cells were seeded on glass coverslips (VWR, 631-0149; 13 mm diameter; Thickness No1) in 24-well plates. On the day of the invasion assay, cells were washed twice in non-supplemented DMEM and infected with bacterial suspensions in DMEM prepared from the LBS subculture of the *Salmonella* strains (see Bacterial strains and growth conditions) at a multiplicity of infection (MOI) of 10 bacterial cells per mammalian cell. Infected cells were incubated at 37°C, 5% CO<sub>2</sub> for 30 min. The *Salmonella*-containing medium was then replaced with complete medium containing

30 µg/ml gentamicin for 30 min to kill all *Salmonella* cells remaining extracellular. The gentamicin concentration was subsequently reduced to 5 µg/ml for the rest of the experiment. At 6 h post-infection, the medium was removed, cells were washed twice in Dubelcco's phosphate-buffered saline (DPBS; Sigma Aldrich, D8537), fixed in 4% paraformaldehyde, at room temperature for 20 min, and washed twice for 5 min at room temperature in PBS prior to immunofluorescence labelling.

### Immunocytochemistry

For the MAP1LC3B/LC3B immunostaining, cells were first treated with 50 mM NH<sub>4</sub>Cl in PBS for 10 min, permeabilized in methanol for 5 min and washed 3 times in PBS for 5 min. Blocking was performed in 1% bovine serum albumin (BSA) Fraction V (Sigma-Aldrich, 05479) in PBS for 30 min at room temperature. The rabbit anti- MAP1LC3B/LC3B antibody (Abcam, ab48394) and FITC-conjugated anti-GFP (Abcam, ab6662) applied were used at 1:2000 and at 1:200 in PBS containing 1% BSA fraction V (Sigma Aldrich, 05479), overnight at 4°C, respectively. For SQSTM1/p62 immunostaining, fixed cells were permeabilized and blocked in 1% BSA, 0.1% saponin (Fluka, 84,510) in PBS for 30 min at room temperature. The rabbit anti-SQSTM1/p62 antibody (Abcam, ab91526) was used at 1:6000 in PBS containing 1% BSA, 0.1% Saponin, overnight at 4°C. All primary antibodies were washed 3 times in either 1% BSA, PBS (MAP1LC3B/LC3B) or 1% BSA, 0.1% saponin, PBS (SQSTM1/p62). Alexa Fluor 594-conjugated anti-rabbit secondary IgG (Abcam, ab91526) was applied at 1:1000 dilution and DAPI stain at 1:2000 in the respective buffer for each primary antibody for 1 h at room temperature. Coverslips were then washed 3 times in respective buffers, once in water and mounted using Fluoromount-G anti-fading compound (ThermoFisher Scientific, 00-4958-02). Coverslips were left to set, sealed using nail varnish and stored at -4°C until observation. Slides were imaged on a Zeiss LSM710 microscope, using a 100x Apochromat (100x/1.4 Oil DIC plan Apo) oil immersion objective. Imaged areas were chosen randomly based solely on the DAPI staining to avoid unconscious-bias.

### Image analysis

For the semi-automated image analysis for SQSTM1/p62 and MAP1LC3/LC3 stainings, ImageJ/Fiji was used for image analysis. Individual cells, identified by the DAPI staining of their nuclei, were segmented, and objects on the edges and segmentation artefacts (small objects) were removed. Individual cell masks (ROIs) were saved for each image. For SQSTM1/p62 staining, an automated macro was used to collect the mean signal intensity data for all 3 channels and area data was gathered for every single segmented cell and saved into an Excel spreadsheet. For the analysis of MAP1LC3/LC3 puncta per cell, only the cells containing at least 1 bacteria were used (manual curation). An automated macro was used to identify the MAP1LC3/LC3 puncta (autophagosomal structures). Information

related to the number of dots per cells were gathered for every single segmented cell and saved into an Excel spreadsheet. Macros are available upon request.

### Quantification of LC3-Salmonella colocalization

For the quantification of colocalization between *Salmonella* and MAP1LC3/LC3, we used a co-localization index (CI) calculated using the following formula:

$$CI = \frac{1 \times n(\text{coloc}) + 0.5 \times n(\text{adjacent}) + 0 \times n(\text{noLC3})}{n(\text{total})}$$

As such, n(coloc) corresponds to the number of bacteria that colocalize completely with LC3 and was attributed the value 1, n(adjacent) corresponds to the number of bacteria that are immediately adjacent/touching MAP1LC3/LC3 structure and was given the value 0.5, and n(no LC3) correspond to the bacteria that are neither co-localizing or close to MAP1LC3/LC3 structures and was given a null value.

### Statistical analyses

Statistical analyses were done with Prism7/8 software (GraphPad). D'Agostino & Pearson normality test was used to test the normality of the sample distribution. Two-tailed t-test was used for the comparison of 2 groups. To compare 3 or more groups, one-way ANOVA was used. For the comparison of more than 3 groups, the multiple comparisons were corrected using Sidak's or Dunn's Multiple comparison test. Dunn's test was done on data that do not meet the assumption of parametric test (normal distribution).

### Acknowledgments

Authors are grateful for the discussions with the members of the Korcsmaros and Nezis groups. IN (Warwick, UK) was supported by BBSRC grants BB/L006324/1 and BB/P007856/1, TK by a fellowship tin computational biology at Earlham Institute (Norwich, UK) in partnership with the Quadram Institute (Norwich, UK), and strategically supported by Biotechnological and Biosciences Research Council, UK (BB/J004529/1 and BB/P016774/1). LG was supported by the BBSRC Norwich Research Park Biosciences Doctoral Training Partnership (grant BB/M011216/1).

### Disclosure statement

No potential conflict of interest was reported by the authors.

### Funding

This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) [BB/P007856/1]; Biotechnology and Biological Sciences Research Council (BBSRC) [BB/P016774/1]; Biotechnology and Biological Sciences Research Council (BBSRC) [BB/J004529/1]; Biotechnology and Biological Sciences Research Council (BBSRC) [BB/L006324/1].

### ORCID

Tamas Korcsmaros  <http://orcid.org/0000-0003-1717-996X>

## References

- [1] Levine B, Mizushima N, Virgin HW. Autophagy in immunity and inflammation. *Nature*. 2011 Jan;469:323–335.
- [2] Randow F, Youle RJ. Self and nonself: how autophagy targets mitochondria and bacteria. *Cell Host Microbe*. 2014 Apr;15:403–411.
- [3] Gomes LC, Dikic I. Autophagy in antimicrobial immunity. *Mol Cell*. 2014 Apr;54:224–233.
- [4] Mostowy S, Cossart P. Bacterial autophagy: restriction or promotion of bacterial replication? *Trends Cell Biol*. 2012 Jun;22:283–291.
- [5] Virgin HW, Levine B. Autophagy genes in immunity. *Nat Immunol*. 2009 May;10:461–470.
- [6] Heath RJ, Goel G, Baxt LA, et al. RNF166 determines recruitment of adaptor proteins during antibacterial autophagy. *Cell Rep*. 2016 Nov;17:2183–2194.
- [7] Tumbarello DA, Manna PT, Allen M, et al. The autophagy receptor TAX1BP1 and the molecular motor myosin VI are required for clearance of salmonella typhimurium by autophagy. *PLoS Pathog*. 2015 Oct;11:e1005174.
- [8] Verlhac P, Grégoire IP, Azocar O, et al. Autophagy receptor NDP52 regulates pathogen-containing autophagosome maturation. *Cell Host Microbe*. 2015 Apr;17:515–525.
- [9] Kabeya Y, Mizushima N, Ueno T, et al. LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing. *Embo J*. 2000 Nov;19:5720–5728.
- [10] Birgisdottir ÁB, Lamark T, Johansen T. The LIR motif - crucial for selective autophagy. *J Cell Sci*. 2013 Aug;126:3237–3247.
- [11] Khaminets A, Behl C, Dikic I. Ubiquitin-dependent and independent signals in selective autophagy. *Trends Cell Biol*. 2016 Jan;26:6–16.
- [12] von Muhlinen N, Akutsu M, Ravenhill BJ, et al. LC3C, bound selectively by a noncanonical LIR motif in NDP52, is required for antibacterial autophagy. *Mol Cell*. 2012 Nov;48:329–342.
- [13] Wild P, Farhan H, McEwan DG, et al. Phosphorylation of the autophagy receptor optineurin restricts Salmonella growth. *Science*. 2011 Jul;333:228–233.
- [14] Pankiv S, Clausen TH, Lamark T, et al. p62/SQSTM1 binds directly to Atg8/LC3 to facilitate degradation of ubiquitinated protein aggregates by autophagy. *J Biol Chem*. 2007 Aug;282:24131–24145.
- [15] Lam GY, Cemma M, Muise AM, et al. Host and bacterial factors that regulate LC3 recruitment to *Listeria monocytogenes* during the early stages of macrophage infection. *Autophagy*. 2013 Jul;9:985–995.
- [16] Mandell MA, Jain A, Arko-Mensah J, et al. TRIM proteins regulate autophagy and can target autophagic substrates by direct recognition. *Dev Cell*. 2014 Aug;30:394–409.
- [17] Veljanovski V, Batoko H. Selective autophagy of non-ubiquitylated targets in plants: looking for cognate receptor/adaptor proteins. *Front Plant Sci*. 2014 Jun;5:308.
- [18] Kim B-W, Kwon DH, Song HK. Structure biology of selective autophagy receptors. *BMB Rep*. 2016;49:73–80.
- [19] Kwon DH, Song HK. A structural view of xenophagy, a battle between host and microbes. *Mol. Cells*. 2018 Jan;41:27–34.
- [20] von Muhlinen N, Thurston T, Ryzhakov G, et al. NDP52, a novel autophagy receptor for ubiquitin-decorated cytosolic bacteria. *Autophagy*. 2010 Feb;6:288–289.
- [21] Schille S, Crauwels P, Bohn R, et al. LC3-associated phagocytosis in microbial pathogenesis. *Int J Med Microbiol*. 2017 Nov. pii: S1438-4221(17)30283-7.
- [22] Huang J, Brumell JH. Bacteria-autophagy interplay: a battle for survival. *Nat Rev Microbiol*. 2014 Feb;12:101–114.
- [23] Choy A, Roy CR. Autophagy and bacterial infection: an evolving arms race. *Trends Microbiol*. 2013 Sep;21:451–456.
- [24] Ogawa M, Yoshikawa Y, Kobayashi T, et al. A Tecpr1-dependent selective autophagy pathway targets bacterial pathogens. *Cell Host Microbe*. 2011 May;9:376–389.
- [25] Ogawa M, Yoshimori T, Suzuki T, et al. Escape of intracellular *Shigella* from autophagy. *Science*. 2005 Feb;307:727–731.
- [26] Choy A, Dancourt J, Mugo B, et al. The Legionella effector RavZ inhibits host autophagy through irreversible Atg8 deconjugation. *Science*. 2012 Nov;338:1072–1076.
- [27] Heinzen RA, Scidmore MA, Rockey DD, et al. Differential interaction with endocytic and exocytic pathways distinguish parasitophorous vacuoles of *Coxiella burnetii* and *Chlamydia trachomatis*. *Infect Immun*. 1996 Mar;64:796–809.
- [28] Winchell CG, Graham JG, Kurten RC, et al. *Coxiella burnetii* type IV secretion-dependent recruitment of macrophage autophagosomes. *Infect Immun*. 2014 Jun;82:2229–2238.
- [29] Scheidel J, Amstein L, Ackermann J, et al. In silico knockout studies of xenophagic capturing of salmonella. *PLoS Comput Biol*. 2016 Dec;12:e1005200.
- [30] Cui T, Li W, Liu L, et al. Uncovering new pathogen-host protein-protein interactions by pairwise structure similarity. *PLoS ONE*. 2016 Jan;11:e0147612.
- [31] Dyer MD, Murali TM, Sobral BW. Computational prediction of host-pathogen protein-protein interactions. *Bioinformatics*. 2007 Jul;23:i159–66.
- [32] Garamszegi S, Franzosa EA, Xia Y. Signatures of pleiotropy, economy and convergent evolution in a domain-resolved map of human-virus protein-protein interaction networks. *PLoS Pathog*. 2013 Dec;9:e1003778.
- [33] Budak G, Eren Ozsoy O, Aydin Son Y, et al. Reconstruction of the temporal signaling network in *Salmonella*-infected human cells. *Front Microbiol*. 2015 Jul;6:730.
- [34] Fiskin E, Bionda T, Dikic I, et al. Global analysis of host and bacterial ubiquitinome in response to salmonella typhimurium infection. *Mol Cell*. 2016 Jun;62:967–981.
- [35] Hartford OM, Wann ER, Höök M, et al. Identification of residues in the *Staphylococcus aureus* fibrinogen-binding MSCRAMM clumping factor A (ClfA) that are important for ligand binding. *J Biol Chem*. 2001 Jan;276:2466–2473.
- [36] Ganesh VK, Rivera JJ, Smeds E, et al. A structural model of the *Staphylococcus aureus* ClfA-fibrinogen interaction opens new avenues for the design of anti-staphylococcal therapeutics. *PLoS Pathog*. 2008 Nov;4:e1000226.
- [37] Higgins J, Loughman A, van Kessel KPM, et al. Clumping factor A of *Staphylococcus aureus* inhibits phagocytosis by human polymorphonuclear leucocytes. *FEMS Microbiol Lett*. 2006 May;258:290–296.
- [38] Haiko J, Westerlund-Wikström B. The role of the bacterial flagellum in adhesion and virulence. *Biology (Basel)*. 2013 Oct;2:1242–1267.
- [39] Cemma M, Kim PK, Brumell JH. The ubiquitin-binding adaptor proteins p62/SQSTM1 and NDP52 are recruited independently to bacteria-associated microdomains to target Salmonella to the autophagy pathway. *Autophagy*. 2011 Mar;7:341–345.
- [40] Ellinghaus D, Zhang H, Zeissig S, et al. Association between variants of PRDM1 and NDP52 and Crohn's disease, based on exome sequencing and functional studies. *Gastroenterology*. 2013 Aug;145:339–347.
- [41] Vignal C, Singer E, Peyrin-Biroulet L, et al. How NOD2 mutations predispose to Crohn's disease? *Microbes Infect*. 2007 Apr;9:658–663.
- [42] Hampe J, Franke A, Rosenstiel P, et al. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet*. 2007 Feb;39:207–211.
- [43] Rioux JD, Xavier RJ, Taylor KD, et al. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet*. 2007 May;39:596–604.
- [44] Galson DL, Roodman GD. Pathobiology of paget's disease of bone. *J Bone Metab*. 2014 May;21:85–98.
- [45] Servera E, Sancho J, Bañuls P, et al. Bulbar impairment score predicts noninvasive volume-cycled ventilation failure during an

- acute lower respiratory tract infection in ALS. *J Neurol Sci.* 2015 Nov;358:87–91.
- [46] Angelakis E, Patrick G, Pelsoni JM, et al. *Orientia tsutsugamushi* in lung of patient with acute respiratory distress syndrome, France, 2013. *Emerging Infect. Dis.* 2015 Feb;21:373–375.
- [47] Chen L, Zheng D, Liu B, et al. VFDB 2016: hierarchical and refined dataset for big data analysis–10 years on. *Nucleic Acids Res.* 2016 Jan;44:D694–7.
- [48] Atilano ML, Pereira PM, Vaz F, et al. Bacterial autolysins trim cell surface peptidoglycan to prevent detection by the drosophila innate immune system. *Elife.* 2014 Apr;3:e02277.
- [49] Jan AT. Outer membrane vesicles (OMVs) of gram-negative bacteria: A perspective update. *Front Microbiol.* 2017 Jun;8:1053.
- [50] Bai J, Kim SI, Ryu S, et al. Identification and characterization of outer membrane vesicle-associated proteins in *Salmonella enterica* serovar Typhimurium. *Infect Immun.* 2014 Oct;82:4001–4010.
- [51] Galka F, Wai SN, Kusch H, et al. Proteomic characterization of the whole secretome of *Legionella pneumophila* and functional analysis of outer membrane vesicles. *Infect Immun.* 2008 May;76:1825–1836.
- [52] Liu Q, Yi J, Liang K, et al. *Salmonella Choleraesuis* outer membrane vesicles: proteomics and immunogenicity. *J Basic Microbiol.* 2017 Jul;852–861.
- [53] Fábrega M-J, Rodríguez-Nogales A, Garrido-Mesa J, et al. Intestinal anti-inflammatory effects of outer membrane vesicles from *Escherichia coli* Nissle 1917 in DSS-experimental colitis in mice. *Front Microbiol.* 2017 Jul;8:1274.
- [54] Cañas M-A, Giménez R, Fábrega M-J, et al. Outer membrane vesicles from the probiotic *Escherichia coli* Nissle 1917 and the commensal ECOR12 enter intestinal epithelial cells via clathrin-dependent endocytosis and elicit differential effects on DNA damage. *PLoS ONE.* 2016 Aug;11:e0160374.
- [55] Aguilera L, Toloza L, Giménez R, et al. Proteomic analysis of outer membrane vesicles from the probiotic strain *Escherichia coli* Nissle 1917. *Proteomics.* 2014 Feb;14:222–229.
- [56] Zakharzhevskaya NB, Vanyushkina AA, Altukhov IA, et al. Outer membrane vesicles secreted by pathogenic and nonpathogenic *Bacteroides fragilis* represent different metabolic activities. *Sci Rep.* 2017 Jul;7:5008.
- [57] Leppla SH. Anthrax toxin edema factor: a bacterial adenylate cyclase that increases cyclic AMP concentrations of eukaryotic cells. *Proc Natl Acad Sci USA.* 1982 May;79:3162–3166.
- [58] Romagnoli A, Etna MP, Giacomini E, et al. ESX-1 dependent impairment of autophagic flux by *Mycobacterium tuberculosis* in human dendritic cells. *Autophagy.* 2012 Sep;8:1357–1370.
- [59] Chen D, Zhong Q. A tethering coherent protein in autophagosome maturation. *Autophagy.* 2012 Jun;8:985–986.
- [60] Michaux C, Holmqvist E, Vasicek E, et al. RNA target profiles direct the discovery of virulence functions for the cold-shock proteins CspC and CspE. *Proc Natl Acad Sci USA.* 2017 Jun;114:6824–6829.
- [61] Eggleston KK, Duffin KL, Goldberg DE. Identification and characterization of falcilysin, a metallopeptidase involved in hemoglobin catabolism within the malaria parasite *Plasmodium falciparum*. *J Biol Chem.* 1999 Nov;274:32411–32417.
- [62] Dabonné S, Moallic C, Sine J-P, et al. Cloning, expression and characterization of a 46.5-kDa metallopeptidase from *Bacillus halodurans* H4 sharing properties with the ptililysin family. *Biochim Biophys Acta.* 2005 Aug;1725:136–143.
- [63] Kimura Y, Ikuta K, Kimura T, et al. Nardilysin regulates inflammation, metaplasia, and tumors in murine stomach. *Sci Rep.* 2017 Feb;7:43052.
- [64] Castanheira S, García-Del Portillo F. *Salmonella* populations inside host cells. *Front Cell Infect Microbiol.* 2017 Oct;7:432.
- [65] Knodler LA, Nair V, Steele-Mortimer O. Quantitative assessment of cytosolic *Salmonella* in epithelial cells. *PLoS ONE.* 2014 Jan;9:e84681.
- [66] Ji Y, McLandsborough L, Kondagunta A, et al. C5a peptidase alters clearance and trafficking of group A streptococci by infected mice. *Infect Immun.* 1996 Feb;64:503–510.
- [67] Nakagawa I, Amano A, Mizushima N, et al. Autophagy defends cells against invading group A *Streptococcus*. *Science.* 2004 Nov;306:1037–1040.
- [68] Knodler LA, Vallance BA, Celli J, et al. Dissemination of invasive *Salmonella* via bacterial-induced extrusion of mucosal epithelia. *Proc Natl Acad Sci USA.* 2010 Oct;107:17733–17738.
- [69] Kirkin V, Lamark T, Sou Y-S, et al. A role for NBR1 in autophagosomal degradation of ubiquitinated substrates. *Mol Cell.* 2009 Feb;33:505–516.
- [70] Brown D. Antibiotic resistance breakers: can repurposed drugs fill the antibiotic discovery void? *Nat Rev Drug Discov.* 2015 Dec;14:821–832.
- [71] The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 2017 Jan;45:D158–D169.
- [72] Peabody MA, Laird MR, Vlasschaert C, et al. PSORTdb: expanding the bacteria and archaea protein subcellular localization database to better reflect diversity in cell envelope structures. *Nucleic Acids Res.* 2016 Jan;44:D663–8.
- [73] Jadhav TS, Wooten MW, Wooten MC. Mining the TRAF6/p62 interactome for a selective ubiquitination motif. *BMC Proc.* 2011 May;5(Suppl 2):S4.
- [74] Kim B-W, Hong SB, Kim JH, et al. Structural basis for recognition of autophagic receptor NDP52 by the sugar receptor galectin-8. *Nat Commun.* 2013;4:1613.
- [75] Jacomin A-C, Samavedam S, Promponas V, et al. iLIR database: A web resource for LIR motif-containing proteins in eukaryotes. *Autophagy.* 2016 Oct;12:1945–1953.
- [76] Kalvari I, Tsompanis S, Mulakkal NC, et al. iLIR: A web resource for prediction of Atg8-family interacting proteins. *Autophagy.* 2014 May;10:913–925.
- [77] van der Lee R, Buljan M, Lang B, et al. Classification of intrinsically disordered regions and proteins. *Chem Rev.* 2014 Jul;114:6589–6631.
- [78] Korcsmaros T, Dunai ZA, Vellai T, et al. Teaching the bioinformatics of signaling networks: an integrated approach to facilitate multi-disciplinary learning. *Brief. Bioinformatics.* 2013 Sep;14:618–632.
- [79] Bader S, Kühner S, Gavin A-C. Interaction networks for systems biology. *FEBS Lett.* 2008 Apr;582:1220–1224.
- [80] Raghavachari B, Tasneem A, Przytycka TM, et al. DOMINE: a database of protein domain interactions. *Nucleic Acids Res.* 2008 Jan;36:D656–61.
- [81] Dinkel H, Van Roey K, Michael S, et al. The eukaryotic linear motif resource ELM: 10 years and counting. *Nucleic Acids Res.* 2014 Jan;42:D259–66.
- [82] Türei D, Földvári-Nagy L, Fazekas D, et al. Autophagy regulatory network - a systems-level bioinformatics resource for studying the mechanism and regulation of autophagy. *Autophagy.* 2015;11:155–165.
- [83] Kubisch J, Türei D, Földvári-Nagy L, et al. Complex regulation of autophagy in cancer - integrated approaches to discover the networks that hold a double-edged sword. *Semin Cancer Biol.* 2013 Aug;23:252–261.
- [84] Remm M, Storm CE, Sonnhammer EL. Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J Mol Biol.* 2001 Dec;314:1041–1052.
- [85] Mészáros B, Simon I, Dosztányi Z. Prediction of protein binding regions in disordered proteins. *PLoS Comput Biol.* 2009 May;5:e1000376.
- [86] Dosztányi Z, Csizmók V, Tompa P, et al. The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J Mol Biol.* 2005 Apr;347:827–839.
- [87] Jones P, Binns D, Chang H-Y, et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics.* 2014 May;30:1236–1240.

- [88] Hoiseith SK, Stocker BAD. Aromatic-dependent *Salmonella typhimurium* are non-virulent and effective as live vaccines. *Nature*. 1981;291(5812):238–239.
- [89] Hautefort I, Proenca MJ, Hinton JCD Single-copy green fluorescent protein gene fusions allow accurate measurement of *Salmonella* gene expression in vitro and during infection of mammalian cells. *Appl. Environ. Microbiol. Dec* 2003;69 (12):7480–7491.
- [90] Sambrook J, Russell D. *Molecular cloning: A laboratory manual*. 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2001.
- [91] Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Pnas*. 2000 June;6:6640–6645.
- [92] Gemski P, Stocker BAD. Transduction by bacteriophage P22 in nonsmooth mutants of *Salmonella typhimurium*. *Genet Mol Biol*. 1967 May;93(5):1588–1597.
- [93] Cherepanov PP, Wackernagel W. Gene disruption in *Escherichia coli*: *tcR* and *KmR* cassettes with the option of F<sub>1</sub>p-catalyzed excision of the antibiotic-resistance determinant. *Gene*. 1995 May;158(1):9–14.