Large-scale identification and translocation of type IV secretion substrates by *Coxiella burnetii*

Chen Chen[a,1], Simran Banga[b,1], Katja Mertens[a,1], Mary M. Weber[a], Ivana Gorbaslieva[a], Yunhao Tan[b], Zhao-Qing Luo[b,2], and James E. Samuel[a,2]

[a]Department of Microbial and Molecular Pathogenesis, Texas A&M Health Science Center, College Station, TX 77843; and [b]Department of Biological Sciences, Purdue University, West Lafayette, IN 47907

Edited by Thomas J. Silhavy, Princeton University, Princeton, NJ, and approved October 29, 2010 (received for review July 22, 2010)

Coxiella burnetii **is an obligate intracellular bacterial pathogen responsible for acute and chronic Q fever. This bacterium harbors a type IV secretion system (T4SS) highly similar to the Dot/Icm of *Legionella pneumophila* that is believed to be essential for its infectivity. Protein substrates of the *Coxiella* T4SS are predicted to facilitate the biogenesis of a phagosome permissive for its intracellular growth. However, due to the lack of genetic systems, protein transfer by the *C. burnetii* Dot/Icm has not been demonstrated. In this study, we report the identification of 32 substrates of the *C. burnetii* Dot/Icm system using a fluorescence-based β-lactamase (TEM1) translocation assay as well as the calmodulin-dependent adenylate cyclase (CyaA) assay in the surrogate host *L. pneumophila*. Notably, 26 identified T4SS substrates are hypothetical proteins without predicted function. Candidate secretion substrates were obtained by using (*i*) a genetic screen to identify *C. burnetii* proteins interacting with DotF, a component of the T4SS, and (*ii*) bioinformatic approaches to retrieve candidate genes that harbor characteristics associated with previously reported substrates of the Dot/Icm system from both *C. burnetii* and *L. pneumophila*. Moreover, we have developed a shuttle plasmid that allows the expression of recombinant proteins in *C. burnetii* as TEM fusion products. Using this system, we demonstrated that a Dot/Icm substrate identified with *L. pneumophila* was also translocated by *C. burnetii* in a process that requires its C terminus, providing direct genetic evidence of a functional T4SS in *C. burnetii*.**

effectors | protein translocation | transporter

Specialized protein transport systems play essential roles in the establishment of symbiotic or pathogenic interactions between microorganisms and their hosts. Among these, a conjugationrelated T4SS has been found in the highly adapted pathogen *Coxiella burnetii,* an obligate intracellular Gram-negative bacterium that replicates inside alveolar mononuclear phagocytes and causes acute and chronic Q fever in humans (1). Unlike other intracellular bacteria that use mechanisms to evade endocytic pathways, *C. burnetii* has a unique intracellular life cycle. After internalization into a host cell, *C. burnetii* establishes a parasitophorous vacuole (PV) that eventually fuses with compartments of the lysosomal network and expands to occupy the majority of the cytoplasmic space within the infected cell (2).

The putative T4SS in *C. burnetii* consists of 23 of the 26 Dot/Icm proteins found in *Legionella pneumophila,* the causative agent of Legionnaires' disease (3). The high similarity between these two transport systems has allowed the use of genetic tools available in *Legionella* to dissect the function of the *C. burnetii* secretion system. Some *C. burnetii dot/icm* genes are capable of complementing corresponding mutations in *L. pneumophila,* suggesting that the *Coxiella* T4SS is active (4). Various strategies have led to the identification of more than 150 protein substrates for the *L. pneumophila* T4SS (5–7). These proteins are predicted to modulate various host processes, including apoptosis, ubiquitination (8–10), lipid metabolism, and membrane trafficking (6, 11–13). By combining bioinformatics tools with the use of *Legionella* as a surrogate host to measure protein translocation, 11 *Coxiella* proteins containing the ankyrin repeat motif have been identified as substrates of its Dot/Icm transporter (11, 14).

The genome of *C. burnetii* and the potential number of genes that it encodes are significantly smaller than those of *L. pneumophila* (3). However, given the diverse challenges that the bacterium faces in the intracellular environment, it is likely that additional Dot/Icm protein substrates are present in the *C. burnetii* genome. To obtain a more complete inventory of substrates transferred by the Dot/Icm system of *C. burnetii*, we initiated a study to identify such proteins by dual strategies that combined genetic screenings and bioinformatics analyses. Because a subset of *Legionella* Dot/Icm substrates specifically interacts with DotF (5), an important component of the T4SS that localizes to the inner membrane of the bacterium (15). We hypothesized that similar interactions occur between *Coxiella* T4SS substrates and its DotF protein. In addition, in *L. pneumophila*, expression of some effectors is coregulated with some *dot/icm* genes, and the genetic elements of such regulatory circuits have been identified (16, 17). Finally, it has been shown that proteins with motifs and structural features specific for eukaryotic cells are more likely to be effectors (7, 14). Thus, we have used a bacterial two-hybrid screening and bioinformatics analyses to search for *Coxiella* proteins that fit one or more of these features. Our efforts with these strategies have led to the retrieval of 57 *C. burnetii* T4SS substrate candidates. Using independent translocation assays based on the Cya (18) or the β-lactamase (TEM1)–mediated FRET on CCF4-AM (6, 19), we demonstrated that 32 of these proteins are translocated into host cells by the *L. pneumophila* Dot/Icm system.

One of the biggest obstacles in the study of obligate intracellular pathogens such as *C. burnetii* is the inability to perform genetic manipulations, making it difficult to directly examine the function of identified virulence factors. In this study, we have generated a *C. burnetii* shuttle vector using the backbone of the IncQ group plasmid RSF1010, which can be stably maintained under selection after introduction by electroporation. Using this vector, we demonstrated that β-lactamase fusion proteins can be expressed in *C. burnetii*. Furthermore, we demonstrated that *C. burnetii* is capable of transferring substrates into host cells in a manner that requires the C-terminal portion of the proteins, providing genetic evidence for the first time that C. *burnetii* has a functional T4SS.

Results

Identification of Putative *C. burnetii* Dot/Icm Substrates by Bacterial Two-Hybrid Screening. *L. pneumophila* DotF, as a major component

Author contributions: C.C., S.B., K.M., Z.-Q.L., and J.E.S. designed research; C.C., S.B., K.M., M.M.W., I.G., and Y.T. performed research; C.C., S.B., K.M., M.M.W., I.G., Y.T., Z.-Q.L., and J.E.S. contributed new reagents/analytic tools; C.C., S.B., K.M., M.M.W., I.G., Y.T., Z.-Q.L., and J.E.S. analyzed data; and C.C. and Z.-Q.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]C.C., S.B., and K.M. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: jsamuel@medicine.tamhsc.edu or luoz@purdue.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1010485107/-/DCSupplemental.

of type IVb apparatus, has been successfully used as bait to identify at least eight *Legionella* T4SS substrates (5). Bioinformatics and experimental evidence indicates extensive homology between *L. pneumophila* and *C. burnetii* T4SS (3). Therefore, we used a bacterial two hybrid screen (20) to identify *C. burnetii* proteins that specifically interact with DotF. Fragments of *C. burnetii* DNA were inserted into pUT18 plasmid to construct a genomic library. After cotransforming the *Escherichia coli* strain BTH101 (20) expressing bait with plasmid DNA harboring a *C. burnetii* library, we used LB X-gal plates or synthetic medium with lactose as the sole carbon source to monitor Lac^+ phenotypes. Clones with interactions significantly higher than the background (Fig. 1*A*) were sequenced and analyzed. From ~100,000 colonies, we identified 25 distinct *orfs* that were fused in-frame with the *Bordetella pertussis* adenylate cyclase T18 fragment. Several genes were identified independently more than one time (Table S1). Notably, another inner membrane-anchored Dot/Icm component, IcmO, was identified multiple times, which suggests a specific interaction between DotF and IcmO. Of the 25 identified genes, 11 candidates were eliminated from further analysis based on their predicted function and conservation among Gram-negative bacteria (Table S1).

To determine whether the DotF-binding proteins can be delivered into host cells by the Dot/Icm system of the surrogate host *L. pneumophila*, we used two different methods, CyaA (18) and TEM1 translocation assays (6). We found that six of 14 proteins could be transferred into host cells in both assays in a Dot/Icm-dependent manner (Figs. 1*B* and 2). Furthermore, translocation of several selected proteins was also confirmed by SidCΔC100 fusion translocation assays (21) (Fig. 1 *C* and *D*). All proteins found to be positive in one assay were also positive in all assays listed, indicating consistency among methods. These results indicate that some *C. burnetii* Dot/Icm substrates interact with type IV component proteins during translocation. Using DotF as bait in a genome-wide preliminary screen led to a validation rate of 42.8% (6/14) for the T4SS substrates, which is severalfold higher than an estimate of total substrates predicted from the *C. burnetti* genome based on the number of effectors predicted from *L. pneumophila* (5–10%) (7).

Genome-Wide Bioinformatic Analysis to Identify Putative *C. burnetii* Dot/Icm Substrates. Based on recent studies, bioinformatic prediction is a promising and effective approach to identify putative bacterial effectors, particularly when a considerable number of validated effectors has been identified by other methods. By using appropriate criteria, computational methods can dramatically narrow the number of potential targets, resulting in a reduction in subsequent experimental validation. Based on the successful application of bioinformatics-guided approaches in identifying *L. pneumophila* Dot/Icm effector proteins (7), we used a similar strategy to search for *C. burnetii* Dot/Icm effector candidates. Several studies with *L. pneumophila* suggested that some Dot/Icm effectors have one or more paralogs present in the *L. pneumophila* genome, which are believed to confer functional redundancy within effector families (5, 22). It is reasonable to speculate that some effectors in *L. pneumophila* and *C. burnetii* have similar targets in the host cell during infection. We performed a genomic search for paralogs of the reported 134 *L. pneumophila* and nine *C. burnetii* T4SS substrates (Table S2). By using a cutoff of expected value (E) <0.01, we obtained 27 *C. burnetii* proteins with varying degrees of similarity ($E = 7e^{-3}$ to $1e^{-95}$) to known T4SS substrates. Of these 27 *orfs*, one Dot/Icm component protein IcmE and 11 ankyrin repeat family proteins, which have been characterized in previous studies (11, 14), were eliminated. The remaining 15 *orfs* were retained for experimental validation (Table S2).

Proteins containing the Fic (filamentation-induced by c-AMP) domain involved in posttranslational modification of target molecule by ampylation represent a recently identified class of virulence factors and have been found in *Vibrio parahaemolyticus*, *Histophilus somni* and *L. pneumophila* (23–25). Three *Coxiella* proteins that contain the core sequence (HPFx[D/E]GN[G/K]R) of the Fic motif (26) were retrieved and included in the candidate pool (Table S3).

Zusman et al. showed that the two-component regulatory system component, PmrA, may directly regulate Dot/Icm type IV expression

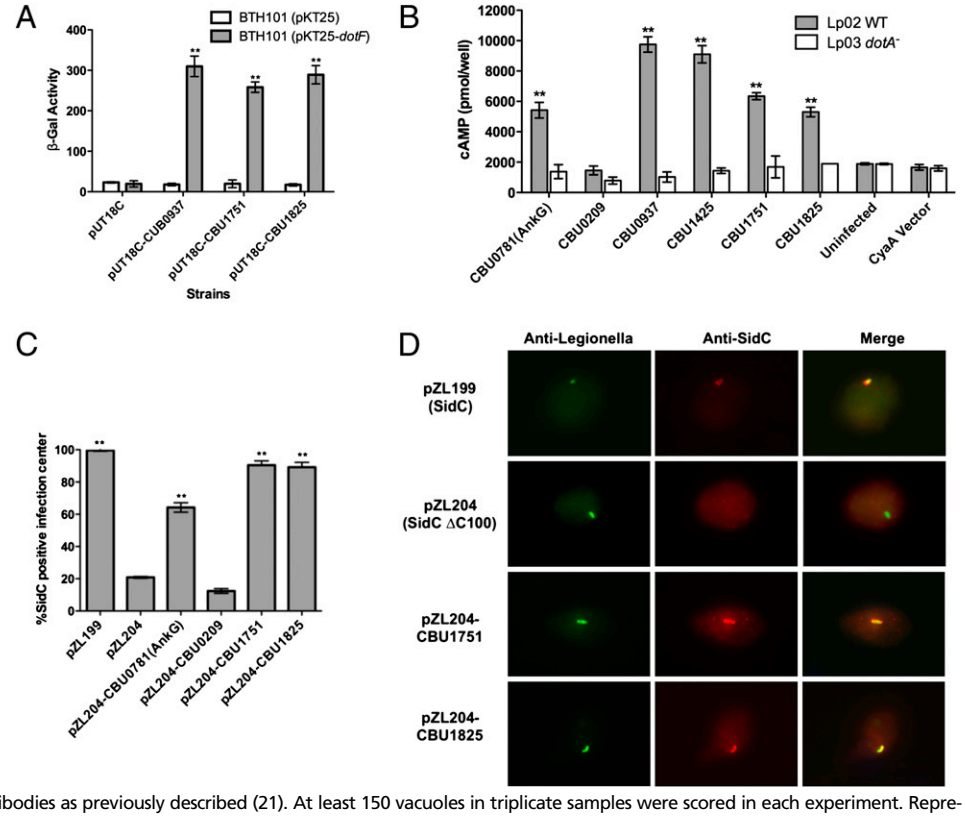

Fig. 1. Several DotF interacting proteins are translocated substrates of the *L. pneumophila* Dot/Icm system. (*A*) Interactions between *C. burnetii* DotF and substrates of its T4SS assayed by the bacterial two-hybrid system. *E. coli* strain BTH101 containing the indicated plasmids were grown to exponential phase with 100 nM IPTG at 28 °C in LB broth. An appropriate volume was withdrawn for β-galactosidase assay. β-Galactosidase activity was expressed as Miller units. Experiments were performed three independent times. Data shown are from one representative experiment. $^{**}P < 0.01$. (*B*) Wild-type (wt) *L. pneumophila* strain Lp02 (gray bar) and the *dotA* mutant strain Lp03 (white bar) harboring the CyaA fusion proteins were used to infect differentiated U937 cells at MOI 30, and the cAMP level of infected cells was determined. A previously validated effector protein *ankG* (CBU0781) (11, 14) was used as positive control. Each bar represents the average cAMP value obtained from triplicate wells ± SD. (*C* and *D*) DotF interacting proteins mediated translocation of the transfer deficient SidCΔC100 into macrophages. Mouse macrophages were infected with *L. pneumophila* expressing SidC (pZL199), SidCΔC100 (pZL204), SidCΔC100::CBU1825 (pZL204-CBU1825), SidCΔC100::CBU0781 (pZL204-CBU0781), SidCΔC100::CBU0209 (pZL204-CBU0209), or SidCΔC100::CBU1751 (pZL204-CBU1751) for 1 h; *L. pneumophila* and SidC were differentially labeled with distinctive antibodies as previously described (21). At least 150 vacuoles in triplicate samples were scored in each experiment. Representative images from individual assay wells are shown in *D*. Bacteria are labeled in green, and SidC is stained in red. $^{**}P < 0.01$ to pZL204.

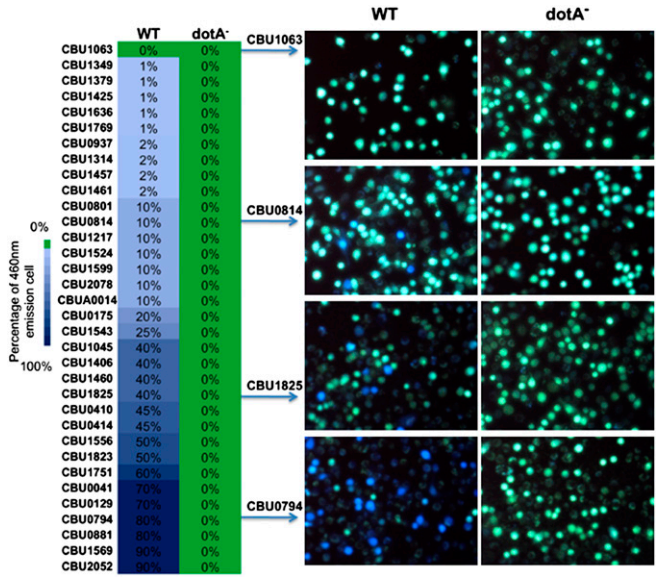

Fig. 2. Dot/Icm-dependent translocation of *C. burnetii* proteins into U937 cells by *L. pneumophila.* WT *L. pneumophila* strain Lp02 and *dotA* mutant strain Lp03-expressing TEM1 fusion proteins were used to infect differentiated U937 cells at an MOI of 30. Infected cells were loaded with CCF4/AM and translocation was determined by counting the percentage of cells with 450 nm emission (with cleaved CCF4/AM). At least 400 events in triplicate samples were scored for each experiment. Representative images from individual assay wells are shown on the right. The results shown represent the average of two or three experiments.

in *L. pneumophila* and *C. burnetii* (17). In addition, they predicted 68 *C. burnetii orfs* that contain putative PmrA response elements by the presence of the consensus sequence TTAA-N6-TTAA. We examined the −200 to 0 bp of all *C. burnetii* Dot/Icm component genes and of experimentally verified *C. burnetii* Dot/Icm substrate genes (11, 14) for PmrA consensus sequences. Along with five *dot/icm* genes (*icmD, icmQ, icmV, icmW,* and *dotD*), several genes encoding ankyrin repeat domain-containing proteins (14) (AnkB, AnkJ, and AnkM), one gene encoding a DotF-binding protein (CBU1751), and five of 15 effector paralogs (Table S3) were found to contain similar PmrA responding elements (Fig. S1*A*). Interestingly, one nucleotide from the putative PmrA responding element upstream of *ankJ* and *ankM* was distinct from the consensus sequence suggested by Zusman et al. (17). This discrepancy is consistent with the notion that PmrA responding elements do not have strict sequence requirements in certain positions (27). We postulated that the *C. burnetii* PmrA-dependent regulon might contain additional unidentified targets, and that some of these targets could be substrates of *C. burnetii* T4SS. To screen all *orfs* potentially regulated by PmrA, we first constructed a PmrA-binding motif model by using sequence information of PmrA-responsive *C. burnetii dot/icm* genes and verified Dot/Icm substrate genes (Fig. S1*B*). Using thresholds described in *Materials and Methods,* we performed a genome-wide search of *C. burnetii* with this PmrA-binding motif model and obtained 126 *orfs* (Table S4).

Bacterial pathogens can acquire virulence factors such as effectors by convergent evolution or by horizontal gene transfer (6, 28, 29). Some of the genes presumably acquired from eukaryotic cells still contain structural features typical for eukaryotic proteins. In agreement with this notion, *C. burnetii* encodes an abundance of eukaryotic-like proteins in its genome (14). These proteins more likely function in eukaryotic cells by targeting various host processes to allow successful infection, making them strong candidates for *Coxiella* T4SS effectors. Indeed, among the 126 putative PmrA-regulated genes, 35 encode proteins with eukaryotic domains implicated in protein–protein interactions (Table S3). After exclusion of all ankyrin domain-containing proteins, which have been previously analyzed for Dot/Icm translocation (11, 14), 24 novel *orfs* were retained for experimental validation (Table S3).

Dot/Icm-Dependent Translocation of *C. burnetii* Proteins Predicted by Bioinformatics Analyses. To determine whether any of the bioinformatically predicted *C. burnetii* proteins are substrates of the Dot/Icm system, we tested 15 effector paralogs, three Fic motif-containing proteins, and 24 putative PmrA-regulated eukaryotic domain-containing proteins using the TEM1 translocation assay in *L. pneumophila*. We fused each of the 42 *orfs* to the C-terminal end of TEM1 protein. *L. pneumophila* strains expressing these fusions were used to challenge differentiated U937 cells. Three hours after uptake, CCF4-AM was loaded into infected monolayers. In these experiments, samples infected by strains harboring fusions of the *L. pneumophila* RalF displayed 95% blue cells, whereas no blue cells were detected in infections receiving the strain carrying the empty plasmid. We considered a protein to be a substrate of the Dot/Icm system if more than 1% of U937 cells emit blue fluorescence after 3 h incubation with WT *L. pneumophila* (Lp02) but not with the DotA-deficient mutant (Lp03) expressing the same fusion protein. These experiments revealed that 26 of the 42 candidates are translocated by *L. pneumophila* in a Dot/Icm-dependent manner (Fig. 2). Among these, nine proteins are from the 15 Dot/Icm effector paralogs and 16 are from the 24 putative PmrA-regulated eukaryotic domain-containing proteins (Tables S3 and S5). CBU2078, one of the three Fic motif-containing proteins, can also be translocated (Table S5). We also examined the steady-state protein levels of all fusions in *L. pneumophila* by immunoblot using a β-lactamase specific antibody. In most strains, the fusion proteins were expressed at a comparable level, indicating that the negative translocation results in infections using these strains were not caused by low-level expression of the fusion proteins (Fig. S2*A*). These results also indicate that there is no clear correlation between transfer efficiency and levels of the fusion protein in the bacterial cells. Thus, a large fraction (26/42) of the bioinformatically predicted *C. burnetii* proteins are Dot/Icm substrates. The success of identifying substrates using the putative PmrA-binding element suggests that expression of some of these genes may regulated by this two-component system.

***C. burnetii* Is Capable of Translocating into Host Cells Substrates Identified with the *L. pneumophila* Heterologous Expression System.** The fact that some *dot/icm* genes from *C. burnetii* were able to complement corresponding *L. pneumophila* mutants suggests that the T4SS in *Coxiella* is active (4). However, there has been no direct experimental evidence demonstrating the activity of the *C. burnetii* Dot/Icm system. To address this problem, we first developed a system for expressing recombinant proteins in *C. burnetii.* Vectors derived from the IncQ plasmid RSF1010 are capable of replicating in a wide range of bacterial species and were the basis of the first vectors used in the genetic studies of *L. pneumophila* pathogenesis (30). Thus, we modified pJB908 (18), a derivative of RSF1010 (30), by inserting an mCherry-Cm^R selection marker cassette (mChcat cassette). The resultant plasmid, pKM230 (Fig. 3*A*), was used to transform *C. burnetii*. Transformed bacterial cells were then used to infect mammalian cells. Approximately 3–4 wk after infection, mCherry-expressing bacteria became visible in all cultures (Fig. 3*B*). We rescued the shuttle vector by transformation of *E. coli* strain TOP10 with total *C. burnetii* DNA isolated from infected cells. Furthermore, the hybridization pattern obtained with the mCherry-specific probe confirmed the presence of pKM230 in *C. burnetii* (Fig. 3*C*). Finally, with the recently developed semisolid ACCM agarose plates (31), we were able to isolate and propagate *C. burnetii* from single colonies (Fig. 3*D*). These results demonstrate that we have successfully developed a plasmid that can stably replicate in *C. burnetii*.

The success in the development of a shuttle vector (pKM230) allowed us to examine whether *C. burnetii* is able to translocate one or more Dot/Icm substrates identified with *L. pneumophila* heterologous expression. We constructed another shuttle vector, pCBTEM, which allows use of the TEM translocation assay in both *L. pneumophila* and *C. burnetii* using origins of pXDC61 (6) and pKM230. The *orf* of CBU1825 was inserted into the 3′ end of the β-lactamase gene to generate CBU1825-pCBTEM. After transformation into *C. burnetii*,

single colonies were isolated from semisolid ACCM agarose plates and propagated in liquid ACCM for 1–2 mo. After induction with 1 mM IPTG for 3 d to induce the expression of the fusion protein, bacteria were purified from ACCM. By immunoblotting using anti–β-lactamase antibodies, we showed that the fusion protein of predicted molecular mass was produced in *C. burnetii* (Fig. 4*A*).

To test protein translocation by *C. burnetii,* we infected THP-1 cells with an MOI of 30 for 24 h and then loaded the samples with CCF4/AM for 8 h at room temperature. Whereas infections with bacteria harboring the empty vector pCBTEM did not cause any cell to emit blue fluorescence, ≈37% blue fluorescent cells were detected in samples infected with a *C. burnetii* strain expressing the TEM-CBU1825 fusion, indicating that CBU1825 was capable of promoting the transfer of the fusion protein into the cytoplasm of THP1 cells (Fig. 4 *B* and *C*). To verify the Dot/Icm substrates identified with *L. pneumophila,* five more *orfs* (CBU0129, CBU0881, CBU1569, CBU1823, and CBU2052) were tested in *C. burnetii*. All of them were capable of transferring the fusion protein into the cytoplasm of THP1 cells (Fig. S2*B*). These results indicate that during infection, *C. burnetii* is able to translocate protein substrates into host cells via the Dot/Icm T4SS.

C-Terminal End of Substrate Is Critical for Protein Translocation by *C. burnetii*. Although the exact features important for Dot/Icm-dependent protein translocation are unknown, signal motif essential for protein translocation by the Dot/Icm system of *L. pneumophila* have been mapped to the C termini of its substrates (32). Voth et al. suggested that the C-terminal portion of *C. burnetii* Dot/Icm substrates is also required for efficient translocation by *L. pneumophila* (14). To determine the importance of the C terminus of *C. burnetii* Dot/Icm substrates in their translocation, we created a CBU1825 mutant that lacked the coding capacity for the last 20 amino acids of the protein and tested its translocation in *L. pneumophila* and *C. burnetii*. The β-lactamase fusion protein of this mutant was stably expressed in *C. burnetii* (Fig. 4*A* 1825ΔC20-pCBTEM). However, this mutant was unable to promote the delivery of the fusion protein into host cells, as no blue fluorescence cells were detected (Fig. 4 *B* and *C*). Similarly, no translocation was observed in macrophages infected by *L. pneumophila* expressing this fusion (Fig. S2*C*). To confirm that the translocation signal is located at C terminus of CBU1825, we created a truncated CBU1825 that encode for the last 50 amino acids of the protein. As we expected, this truncated CBU1825C50 was well expressed in *C. burnetii* (Fig. 4*A* 1825C50-pCBTEM) and partially restored the translocation capacity of the full length CBU1825 in both *C. burnetii* and *L. pneumophila* (Fig. 4 *B* and *C* and Fig. S2*C*). These results indicate that, similar to other T4SS, the C-terminal end of these *C. burnetii* Dot/Icm substrates contains a translocation signal, further supporting the hypothesis of T4SS-dependent translocation of β-lactamase fusion proteins in *C. burnetii*.

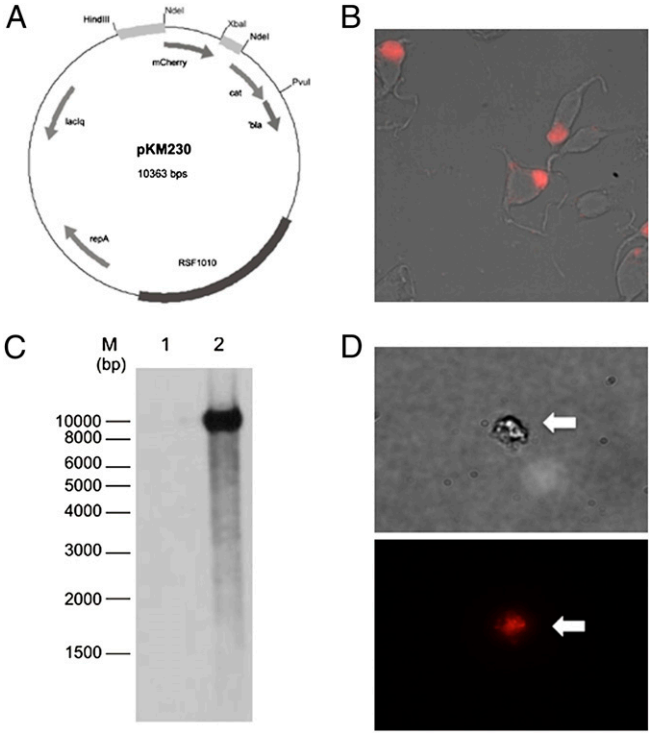

Fig. 3. Development of a stably expressing RSF1010-based shuttle vector. (*A*) Schematic map of the constructed RSF1010 shuttle vector pKM230. (*B*) Live image of L929 cells infected by pKM230-transformed *C. burnetii*. Fluorescence images for detection of mCherry expression (594nm) and DIC (60×) were taken separately and merged. (*C*) Southern blot analysis of *C. burnetii* carrying pKM230. Lane 1, genomic *C. burnetii* DNA, Lane 2, genomic DNA from *C. burnetii* transformed with pKM230. (*D*) *C. burnetii* pKM230 colony formation on solid ACCM medium. Colonies were visible by microscopy (20×) after 3 d of inoculation (*Upper*), and pKM230-derived mCherry expression was visualized by fluorescence microscopy (*Lower*). Arrows indicate single colony formation.

***C. burnetii* Dot/Icm Substrates Localize to Distinct Subcellular Regions When Expressed Ectopically in Mammalian Cells.** Upon being translocated into host cells, bacterial effectors often target specific host organelles to fulfill their physiological functions (6, 11, 14). To gain insights into these identified *C. burnetii* effector proteins, we constructed GFP fusions of these proteins and expressed each in Vero cells by transient transfection. Whereas many of these fusions appeared to be cytosolic in patterns indistinguishable from that of GFP alone, 12 fusions displayed distinct distribution patterns. For example, CBU1314 exclusively localized to nuclei, which was evidenced by colocalization with the DNA labeling dye Hoechst (Fig. S3). On the other hand, CBU1543, CBU1825, and CBU0801 displayed punctuate structures that associated with undefined vesicular compartments. CBU1556 was excluded from nuclei and showed filament structure throughout the cytoplasm (Fig. S3). Localization of these proteins to distinct host organelles adds strong support to the notion that many of these proteins are bona fide effectors that function to manipulate host processes in specific subcellular sites.

Discussion

It is well established that effector proteins of intracellular bacterial pathogens play essential roles in infection and the development of disease (33, 34). Functional analysis of effectors holds great promise in the study of pathogenesis, especially in the elucidation of host cellular pathways, which are either essential or co-opted by the pathogen. However, the lack of genetic tools for obligate intracellular pathogens has made the identification of their effectors a daunting task. In this study, by combining genetic screening and bioinformatic analysis, we obtained a large collection of candidate protein substrates for the *C. burnetii* Dot/Icm T4SS. The use of *L. pneumophila* as a surrogate host led to the identification of 32 proteins that were translocated into host cells in a Dot/Icm-dependent manner. Interestingly, two of the identified Dot/Icm substrates (CBU1314 and CBU1823) are suggested to localize in the cytoplasm of *C. burnetii* infected Vero cells (by mass spectrometry) (35). Although the biological significance of the interaction between protein substrates and components of the transport system is largely undefined, such interactions have been very useful in identifying novel substrates (5, 18, 36). The observation that some *C. burnetii* Dot/Icm substrates bind to its own DotF further indicates the high conservation between these two IVB secretion systems.

Our choice for "bait" (DotF) in the bacterial two-hybrid screen was based primarily on studies with *Legionella* (5). However, a study by Vincent et al. (15, 37) indicated that other Dot/Icm component proteins, such as T4SS chaperones IcmS and IcmW, may be equally promising baits for this type of screen. Our bioinformatics approach achieved significant success in predicting *C. burnetii* type IV substrates, as 26 of 42 predicted candidates were confirmed as Dot/Icm substrates (Table S3). The high accuracy of this predictive approach suggested that a significant number of the substrates' homologs (Table S6) and putative PmrA-regulated genes (Table S2) that were not experimentally tested may also be Dot/Icm substrates. Considering that the number of predicted genes in *C. burnetii* is less than

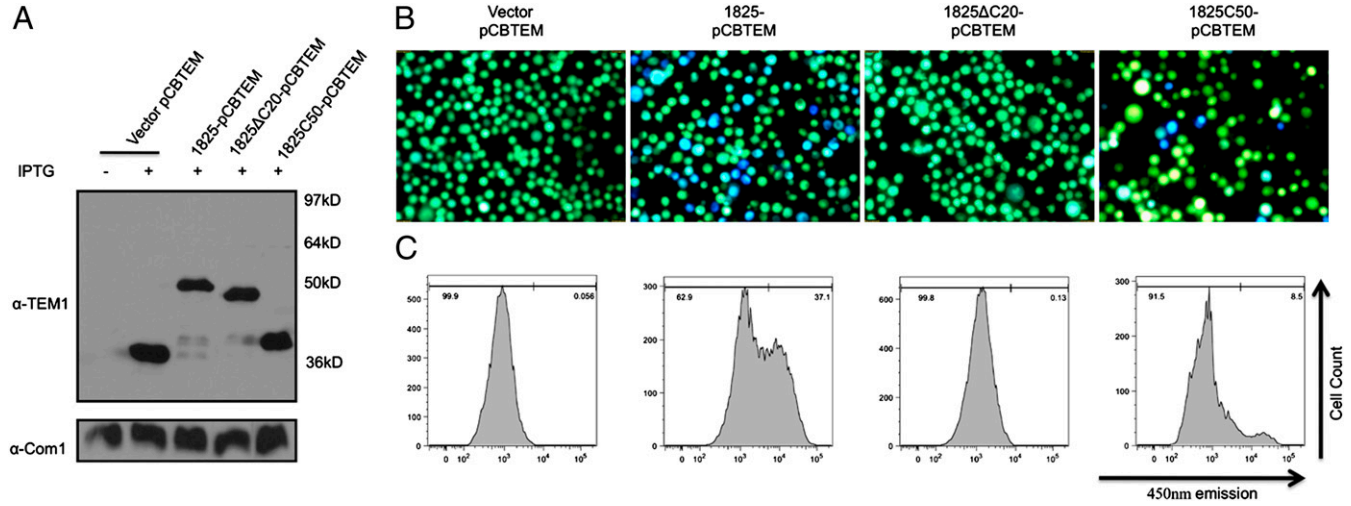

Fig. 4. *C. burnetii* has a functional secretion system and the C-terminal end of the substrate is important for efficient translocation. (*A*) TEM1 fusion proteins were induced by IPTG in pCBTEM plasmids-transformed *C. burnetii* NMII strain. ACCM propagated *C. burnetii* variants were harvested and purified. Expression of TEM1 fusion proteins was analyzed by Western blot. *C burnetii* surface antigen Com1 (CBU1910) was probed as loading control. Lane 1, *C. burnetii*/pCBTEM without IPTG. Lane 2, *C. burnetii*/pCBTEM with 1mM IPTG. Lane 3, *C. burnetii*/pCBTEM-CBU1825 with 1mM IPTG. Lane 4, *C. burnetii*/pCBTEM-CBU1825ΔC20 with 1 mM IPTG. Lane 5, *C. burnetii*/pCBTEM-CBU1825C50 with 1 mM IPTG. (*B* and *C*) *C. burnetii* NMII variants harboring the TEM1 fusion proteins were used to infect THP-1 cells at MOI 30, and TEM assay was performed as described in *Materials and Methods*. Representative image from individual assays are shown on left. The percentage of cells with an elevated 450 nm emission was measured by flow cytometry. Results shown represent the average of three to four independent experiments.

70% of that of the *L. pneumophila,* it is surprising that *C. burnetii* may have more putative PmrA-regulated genes than *L. pneumophila* (17) (126 vs. 35). Including previously described ankrin repeats containing proteins (11, 14), there are 31 *C. burnetii*-confirmed Dot/Icm substrates (Table S5) and five Dot/Icm proteins (Fig. 3) that contain a putative PmrA consensus sequence, which represent more than 25% of predicted PmrA-regulated genes in *C. burnetii*. These observations suggest that the PmrA-PmrB two-component regulatory system may play a critical role in regulating the physiology and pathgenesis of *C. burnetii,* probably by controlling the assembly and function of the T4SS. Validation of each putative PmrA-regulated gene will be an important addition to our understanding of the regulation of pathogenic factors by this organism.

The recent progress in the study of pathogenic mechanisms used by *C. burnetii*, particularly the recent development of a complex medium for axenic culture of this bacterium and relevant in vitro and in vivo models of Q fever, promise a rapid expansion in the analysis of T4SS and its effectors during infections. Because of the paucity of genetic tools, the activity of the *C. burnetii* Dot/Icm system has not been directly examined and some important *L. pneumophila dot/icm* genes are absent in the *C. burnetii* genome (3), the fundamental question of their role in infection has remained untested. Our demonstration that the reporter protein β-lactamase fused with a Dot/Icm substrates was translocated into host cells directly supports the functional assessment of the *C. burnetii* T4SS. Although the fusion protein was stably expressed after adding the CCF4-AM FRET substrate, it took considerably longer for *C. burnetii*-infected samples to develop blue fluorescence than those infected with *Legionella*. Many reasons, including the intrinsic properties of the transporters, efficiency of phagocytosis, and infection conditions, could account for these differences. Nevertheless, the observation that both pathogens were able to transfer the fusion protein in a manner that required the C-terminal end of the substrate demonstrates that the *C. burnetii* Dot/Icm system is active. With the development of mutagenesis systems for *C. burnetii* with the *Himar1* transposon (38), gene inactivation will allow phenotypic characterization of specific mutants in this organism. It will be important to determine whether all of the Dot/Icm proteins are essential for substrate translocation.

Like other highly adapted intracellular pathogens, many of the *C. burnetii* Dot/Icm substrates contain motifs or structural features specific for interaction with proteins of eukaryote origin. Indeed, the presence of the ankyrin repeat domain was used to identify the first *C. burnetii* Dot/Icm substrates (11). Accumulating evidence suggests high-level conservation between the T4SS of *L. pneumophila* and *C. burnetii* in structure and function in substrate translocation. That only few substrates from these organisms display detectable similarity further reflects their very different intracellular lifestyles and the different hosts the bacteria associate with in the natural environments. For examples, whereas *C. burnetii* needs to effectively evade immune recognition by its hosts, *L. pneumophila* needs to be highly adaptive to diverse amoebae hosts. Future work aiming at revealing the biochemical and cell biological functions of these substrates will validate the necessity of such differences.

Among the substrates identified in this study, 23 contain eukaryotic-like domains implicated in protein–protein interaction (Table S5). For example, 17 substrates harbor one or more coiled-coil domains (Table S5), which are present in components of large protein complexes. CBU1457 contains several tetratricopeptide repeats (TPRs), which are 3–16 repeated motifs of 34 amino acids known to mediate protein–protein interactions. Moreover, the eukaryotic-like serine/threonine protein kinase (STPK) domain is present in CBU0175 and CBU1397 (Table S5). These proteins may phosphorylate host targets to interfere with signal transduction pathways. The identification of three proteins (CBU0814, CBU1217, and CBUA0014) with an F-box motif suggests that *C. burnetii* actively modulates host ubiquitination pathways, which is a strategy shared by many bacterial pathogens, including *L. pneumophila* (8–10), *Salmonella enterica Typhimurium* (39), and *Shigella* spp (40). Interestingly, one Fic domain containing protein, which is involved in posttranslational modification of target molecules by ampylation (25, 41), has also been confirmed as a Dot/Icm substrate.

C. *burnetii* may have acquired these substrate genes by horizontal gene transfer (42). In support of this notion, 14 of these 32 genes (Table S5) have GC contents that are significantly (more that 10%) different from the average GC content of *C. burnetii* (42.7%). Despite the high level similarity between the Dot/Icm systems of *C. burnetii* and *L. pneumophila*, few substrates from these two organisms exhibit significant homology (six of them are homologs, Table S2). Such low similarity may result from the very different hosts (mammals vs. primarily amoebae), in which the two pathogens have coevolved. Alternatively, these effectors may have been acquired by convergent evolution. Thus, although no homology was detected, some of the proteins may have similar functional domains. In addition, similar to the finding in *L. pneumophila* type IV substrates (5, 22), several *C. burnetii* Dot/Icm substrates have more than one homolog in its

genome (Table S6), which suggests that functional redundancy of T4SS substrates may also exist in *C. burnetii*. Furthermore, except for one protein (CBUA0014) encoded by a plasmid carried gene, only 18 substrates are fully conserved among *C. burnetii* isolates (Table S5), a phenomenon in contrast to *L. pneumophila,* in which T4SS substrates are highly conserved among different isolates of the same species (43). The majority of missing homologs were presented as pseudogenes in different isolates. There are 11 and 13 homologs missing in chronic disease prototype isolates, K and G, respectively, and 10 homologs are missing in the ancestral isolate, Dugway. This supports the hypothesis that several conserved Dot/Icm substrates play a key role in bacterial survival and that some group -pecific T4SS substrates may participate in pathotype-specific virulence.

At present, it is difficult to determine which host pathways may be hijacked by specific Dot/Icm substrate proteins. Nevertheless, specialized protein transfer systems and their substrates are important in manipulating the cell biology of host cell. These T4SS effectors may direct *C. burnetii* invasion, trafficking in the host cell, or modulation and evasion of host cell functions. Future functional analysis of effectors identified in this study should provide molecular mechanisms that account for the unique intracellular life cycle of this important pathogen.

Materials and Methods

All *L. pneumophila* strains used in this study were derivatives of the Philadelphia 1 strain Lp02 (44). Bacteria were grown and maintained on CYE medium as previously described (45). When necessary, chloramphenicol (5 μg/mL) and kanamycin (20 μg/mL) were added to the cultures. *C. burnetii* strain Nine Mile phase II RSA 439 (clone 4) was propagated in liquid acidified citrate cysteine medium (ACCM) or on semisolid ACCM agarose plates under 2.5% oxygen as previously described (31). The bacterial strains used in this study are listed in Table S7. Details of plasmid construction, bacteria purification and transformation, bacterial infection and CyaA/TEM/SidC translocation assays, immunofluorescence assays, flow cytometry analysis, and bioinformatic analysis are presented in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Dr. Howard Shuman (Columbia University, New York, NY) for the β-lactamase vector pXDC61, Dr. Joseph Vogel (Washington University, Seattle, WA) for the *Legionella* shuttle vector pJB908, and Wenhan Zhu (Purdue University, West Lafayette, IN) for constructing pXDC61m. We thank Dr. Laura Hendrix for critical review of the manuscript, and Hui Zhi and Mingzhe Zhang for performing BLAST search and data collection. This study was supported by National Institutes of Health/National Institute of Allergy and Infectious Diseases Grants R56AI090142-01 (to J.E.S./Z.Q.L.), R01AI057156 (to J.E.S.), and R01AI069344, K02AI085403, and R03AI073326 (to Z.Q.L.).

1. Maurin M, Raoult D (1999) Q fever. *Clin Microbiol Rev* 12:518–553.
2. Heinzen RA, Scidmore MA, Rockey DD, Hackstadt T (1996) Differential interaction with endocytic and exocytic pathways distinguish parasitophorous vacuoles of Coxiella burnetii and Chlamydia trachomatis. *Infect Immun* 64:796–809.
3. Seshadri R, et al. (2003) Complete genome sequence of the Q-fever pathogen Coxiella burnetii. *Proc Natl Acad Sci USA* 100:5455–5460.
4. Vogel JP (2004) Turning a tiger into a house cat: Using Legionella pneumophila to study Coxiella burnetii. *Trends Microbiol* 12:103–105.
5. Luo ZQ, Isberg RR (2004) Multiple substrates of the Legionella pneumophila Dot/Icm system identified by interbacterial protein transfer. *Proc Natl Acad Sci USA* 101:841–846.
6. de Felipe KS, et al. (2008) Legionella eukaryotic-like type IV substrates interfere with organelle trafficking. *PLoS Pathog* 4:e1000117.
7. Burstein D, et al. (2009) Genome-scale identification of Legionella pneumophila effectors using a machine learning approach. *PLoS Pathog* 5:e1000508.
8. Price CT, et al. (2009) Molecular mimicry by an F-box effector of Legionella pneumophila hijacks a conserved polyubiquitination machinery within macrophages and protozoa. *PLoS Pathog* 5:e1000704.
9. Lomma M, et al. (2010) The Legionella pneumophila F-box protein Lpp2082 (AnkB) modulates ubiquitination of the host protein parvin B and promotes intracellular replication. *Cell Microbiol* 12:1272–1291.
10. Ensminger AW, Isberg RR (2010) E3 ubiquitin ligase activity and targeting of BAT3 by multiple Legionella pneumophila translocated substrates. *Infect Immun* 78: 3905–3919.
11. Pan X, Lührmann A, Satoh A, Laskowski-Arce MA, Roy CR (2008) Ankyrin repeat proteins comprise a diverse family of bacterial type IV effectors. *Science* 320:1651–1654.
12. Ge J, et al. (2009) A Legionella type IV effector activates the NF-kappaB pathway by phosphorylating the IkappaB family of inhibitors. *Proc Natl Acad Sci USA* 106: 13725–13730.
13. Banga S, et al. (2007) Legionella pneumophila inhibits macrophage apoptosis by targeting pro-death members of the Bcl2 protein family. *Proc Natl Acad Sci USA* 104: 5121–5126.
14. Voth DE, et al. (2009) The Coxiella burnetii ankyrin repeat domain-containing protein family is heterogeneous, with C-terminal truncations that influence Dot/Icm-mediated secretion. *J Bacteriol* 191:4232–4242.
15. Vincent CD, et al. (2006) Identification of the core transmembrane complex of the Legionella Dot/Icm type IV secretion system. *Mol Microbiol* 62:1278–1291.
16. Rasis M, Segal G (2009) The LetA-RsmYZ-CsrA regulatory cascade, together with RpoS and PmrA, post-transcriptionally regulates stationary phase activation of Legionella pneumophila Icm/Dot effectors. *Mol Microbiol* 72:995–1010.
17. Zusman T, et al. (2007) The response regulator PmrA is a major regulator of the icm/dot type IV secretion system in Legionella pneumophila and Coxiella burnetii. *Mol Microbiol* 63:1508–1523.
18. Bardill JP, Miller JL, Vogel JP (2005) IcmS-dependent translocation of SdeA into macrophages by the Legionella pneumophila type IV secretion system. *Mol Microbiol* 56:90–103.
19. Mills E, Baruch K, Charpentier X, Kobi S, Rosenshine I (2008) Real-time analysis of effector translocation by the type III secretion system of enteropathogenic Escherichia coli. *Cell Host Microbe* 3:104–113.
20. Karimova G, Pidoux J, Ullmann A, Ladant D (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci USA* 95:5752–5756.
21. VanRheenen SM, Luo ZQ, O'Connor T, Isberg RR (2006) Members of a Legionella pneumophila family of proteins with ExoU (phospholipase A) active sites are translocated to target cells. *Infect Immun* 74:3597–3606.
22. Ninio S, Celli J, Roy CR (2009) A Legionella pneumophila effector protein encoded in a region of genomic plasticity binds to Dot/Icm-modified vacuoles. *PLoS Pathog* 5: e1000278.
23. Luong P, et al. (2010) Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain. *J Biol Chem* 285:20155–20163.
24. Worby CA, et al. (2009) The fic domain: Regulation of cell signaling by adenylylation. *Mol Cell* 34:93–103.
25. Roy CR, Mukherjee S (2009) Bacterial FIC proteins AMP up infection. *Sci Signal* 2:pe14.
26. Kinch LN, Yarbrough ML, Orth K, Grishin NV (2009) Fido, a novel AMPylation domain common to fic, doc, and AvrB. *PLoS ONE* 4:e5818.
27. Marchal K, et al. (2004) In silico identification and experimental validation of PmrAB targets in Salmonella typhimurium by regulatory motif detection. *Genome Biol* 5:R9.
28. Cazalet C, et al. (2004) Evidence in the Legionella pneumophila genome for exploitation of host cell functions and high genome plasticity. *Nat Genet* 36:1165–1173.
29. Cazalet C, et al. (2010) Analysis of the Legionella longbeachae genome and transcriptome uncovers unique strategies to cause Legionnaires' disease. *PLoS Genet* 6:e1000851.
30. Marra A, Shuman HA (1989) Isolation of a Legionella pneumophila restriction mutant with increased ability to act as a recipient in heterospecific matings. *J Bacteriol* 171: 2238–2240.
31. Omsland A, et al. (2009) Host cell-free growth of the Q fever bacterium Coxiella burnetii. *Proc Natl Acad Sci USA* 106:4430–4434.
32. Nagai H, et al. (2005) A C-terminal translocation signal required for Dot/Icm-dependent delivery of the Legionella RalF protein to host cells. *Proc Natl Acad Sci USA* 102:826–831.
33. Voth DE, Heinzen RA (2009) Coxiella type IV secretion and cellular microbiology. *Curr Opin Microbiol* 12:74–80.
34. Ninio S, Roy CR (2007) Effector proteins translocated by Legionella pneumophila: Strength in numbers. *Trends Microbiol* 15:372–380.
35. Samoilis G, et al. (2010) Proteomic screening for possible effector molecules secreted by the obligate intracellular pathogen Coxiella burnetii. *J Proteome Res* 9:1619–1626.
36. Cambronne ED, Roy CR (2007) The Legionella pneumophila IcmSW complex interacts with multiple Dot/Icm effectors to facilitate type IV translocation. *PLoS Pathog* 3:e188.
37. Vincent CD, Vogel JP (2006) The Legionella pneumophila IcmS-LvgA protein complex is important for Dot/Icm-dependent intracellular growth. *Mol Microbiol* 61:596–613.
38. Beare PA, et al. (2009) Characterization of a Coxiella burnetii ftsZ mutant generated by Himar1 transposon mutagenesis. *J Bacteriol* 191:1369–1381.
39. Zhang Y, Higashide W, Dai S, Sherman DM, Zhou D (2005) Recognition and ubiquitination of Salmonella type III effector SopA by a ubiquitin E3 ligase, HsRMA1. *J Biol Chem* 280:38682–38688.
40. Singer AU, et al. (2008) Structure of the Shigella T3SS effector IpaH defines a new class of E3 ubiquitin ligases. *Nat Struct Mol Biol* 15:1293–1301.
41. Yarbrough ML, et al. (2009) AMPylation of Rho GTPases by Vibrio VopS disrupts effector binding and downstream signaling. *Science* 323:269–272.
42. Beare PA, et al. (2009) Comparative genomics reveal extensive transposon-mediated genomic plasticity and diversity among potential effector proteins within the genus Coxiella. *Infect Immun* 77:642–656.
43. Cazalet C, et al. (2008) Multigenome analysis identifies a worldwide distributed epidemic Legionella pneumophila clone that emerged within a highly diverse species. *Genome Res* 18:431–441.
44. Berger KH, Isberg RR (1993) Two distinct defects in intracellular growth complemented by a single genetic locus in Legionella pneumophila. *Mol Microbiol* 7:7–19.
45. Conover GM, Derré I, Vogel JP, Isberg RR (2003) The Legionella pneumophila LidA protein: A translocated substrate of the Dot/Icm system associated with maintenance of bacterial integrity. *Mol Microbiol* 48:305–321.