

# Novel *Burkholderia mallei* Virulence Factors Linked to Specific Host-Pathogen Protein Interactions\*<sup>§</sup>

Vesna Memišević‡, Nela Zavaljevski‡, Rembert Pieper§, Seesandra V. Rajagopala§, Keehwan Kwon§, Katherine Townsend§, Chenggang Yu‡, Xueping Yu‡, David DeShazer¶, Jaques Reifman‡||, and Anders Wallqvist‡

*Burkholderia mallei* is an infectious intracellular pathogen whose virulence and resistance to antibiotics makes it a potential bioterrorism agent. Given its genetic origin as a commensal soil organism, it is equipped with an extensive and varied set of adapted mechanisms to cope with and modulate host-cell environments. One essential virulence mechanism constitutes the specialized secretion systems that are designed to penetrate host-cell membranes and insert pathogen proteins directly into the host cell's cytosol. However, the secretion systems' proteins and, in particular, their host targets are largely uncharacterized. Here, we used a combined *in silico*, *in vitro*, and *in vivo* approach to identify *B. mallei* proteins required for pathogenicity. We used bioinformatics tools, including orthology detection and *ab initio* predictions of secretion system proteins, as well as published experimental *Burkholderia* data to initially select a small number of proteins as putative virulence factors. We then used yeast two-hybrid assays against normalized whole human and whole murine proteome libraries to detect and identify interactions among each of these bacterial proteins and host proteins. Analysis of such interactions provided both verification of known virulence factors and identification of three new putative virulence proteins. We successfully created insertion mutants for each of these three proteins using the virulent *B. mallei* ATCC 23344 strain. We exposed BALB/c mice to mutant strains and the wild-type strain in an aerosol challenge model using lethal *B. mallei* doses. In each set of experiments, mice exposed to mutant strains survived for the 21-day duration of the experiment, whereas mice exposed to the wild-type strain rapidly died. Given their *in vivo* role in pathogenicity, and based on the yeast

two-hybrid interaction data, these results point to the importance of these pathogen proteins in modulating host ubiquitination pathways, phagosomal escape, and actin-cytoskeleton rearrangement processes. *Molecular & Cellular Proteomics* 12: 10.1074/mcp.M113.029041, 3036–3051, 2013.

Taken out of their natural rhizosphere environment, many members of the genus *Burkholderia* have proven adept in surviving and adapting to many diverse environments. Of particular interest are the two closely related pathogenic members: *Burkholderia mallei*, the causative agent of glanders, a disease primarily affecting horses but transmittable to humans; and *Burkholderia pseudomallei*, which is responsible for melioidosis, a human disease endemic to Southeast Asia and Northern Australia. Human infection of these opportunistic pathogens occurs through ingestion, inhalation, or skin abrasion. Given their considerable antibiotic resistance, ability to infect via aerosol, and absence of vaccines, these pathogens form both an emerging public health threat in their natural environment and a potential bioterrorism threat (1, 2). *B. mallei*, the focus of our study and the less characterized of the two species, is considered a deletion clone of *B. pseudomallei*, with >1000 genes lost in an adaptation to equine hosts. Thus, the genes retained in *B. mallei* share a high sequence similarity to genes common to *B. pseudomallei* (3), and many virulence factors are common for these two species (4). Some virulence factors present in *B. pseudomallei* were lost in *B. mallei*'s adaptation to its obligatory intracellular lifestyle in equids. However, the mechanisms and capabilities required for successful host colonization are retained in *B. mallei*, e.g. quorum sensing (5, 6), adhesion (7), capsular polysaccharide gene clusters and lipopolysaccharides (8), actin-based motility (9, 10), and a number of different secretion systems (11–14). Secreted proteins, known as “effector” proteins interact with and hijack critical human proteins and pathways, which allow the pathogen to survive and propagate in the host environment. These pathogen-induced interactions encompass alteration of host cell signaling (15), cytoskeletal modulations (16), ubiquitin modification (17), autophagy suppression (4), and apoptotic/pyroptotic con-

From the ‡Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, Maryland 21702; §J. Craig Venter Institute, Rockville, Maryland 20850; ¶Bacteriology Division, U.S. Army Medical Research Institute of Infectious Diseases, Fort Detrick, Maryland 21702

\* Author's Choice—Final version full access.

Received March 8, 2013, and in revised form, June 10, 2013

Published, MCP Papers in Press, June 23, 2013, DOI 10.1074/mcp.M113.029041

trol (18). An apparently distinctive capability of *B. pseudomallei* and *B. mallei* is their tendency to form multinucleated giant host-cell complexes as a means for the pathogen to spread using a direct cell-to-cell traversal mechanism (4, 9, 19). All of these capabilities are ultimately derived from natural adaptations and survival strategies evolved for living in the rhizosphere (20).

Research into the pathogenicity of *B. pseudomallei* and *B. mallei* has focused mostly on the bacterial aspect of infectivity, host immune responses, and disease progression, whereas the fundamental molecular host-pathogen interactions that mediate these activities have gathered less attention. Although the transcriptional immune response in the host has been investigated by microarray experiments (21–23), most of the proteins used by the pathogen to exert its influence are not known. A more comprehensive picture of how the pathogen interacts with host cells requires novel approaches that combine traditional bacteriology, high-throughput experiments, and computational methods. In this study, we used a combined computational and experimental approach to systematically identify *B. mallei* proteins and their host protein interactions. In particular, we focused on *B. mallei* proteins that could potentially be translocated into host cells via specific secretion systems. We then screened these bacterial proteins *in vitro* using yeast two-hybrid (Y2H)<sup>1</sup> assays against both human and murine whole proteomes. A computational analysis of these interactions provided likely functional targets for these proteins. Accordingly, we selected three *B. mallei* genes previously not tied to virulence, separately constructed insertion mutants for each of the genes, and infected BALB/c mice in an aerosol challenge model with the mutants. Whereas mice infected with the wild-type *B. mallei* ATCC 23344 strain rapidly succumbed to the infection and died, all animals separately infected with each of the three mutant strains survived the aerosol challenge for the duration of the experiment.

These results implicated three *B. mallei* proteins, not previously considered virulence factors, to an *in vivo* virulence phenotype. Two of these proteins are annotated as hypothetical proteins of unknown function, and we used the Y2H data to characterize their putative molecular functions. The apparent role of BMAA0728 relates to ubiquitination and phagosomal escape processes, whereas BMAA1865 activity appears directed toward host actin-cytoskeleton rearrangement processes. The third protein, BMAA0553, a putative phosphatase by sequence homology, is postulated to interfere with

host signaling pathways and actin cytoskeleton rearrangement. Based on our current understanding of *Burkholderia* spp. host infections, these findings shed complementary mechanistic light on *B. mallei* pathogenicity as it pertains to host cell invasion and survival. They also raise the possibility that *B. mallei* virulence factors are involved in more than one stage of the bacterial intracellular life cycle. These hypotheses and observations lay the foundation for future biochemical and immunological work.

#### EXPERIMENTAL PROCEDURES

**Comparative Genomics**—We used comparative genomics to 1) identify genes that discriminate pathogenic and nonpathogenic strains of *B. mallei* and 2) search for proteins using orthology with known virulence factors. In these computations, we used a recently published orthology detection method developed by our group, termed QuartetS (24, 25), which exploits evolutionary evidence to predict evolutionary related genes in a computationally efficient manner, enabling whole genome comparisons.

Functional annotations of genomes were taken from GenBank and Pathema (26, 27). For hypothetical or unannotated proteins, we generated function hypotheses using the algorithms implemented in our integrated protein function annotation software, PIPA (28), which combines multiple programs and databases to provide a unified terminology and consensus function annotation.

**Ab Initio Prediction of Type 3 Secretion System Proteins**—As secretion system proteins frequently diverged beyond recognition by orthology, specialized methods are required for their prediction. Recently, several groups have described computational methods to identify type 3 secretion system (T3SS) proteins based on protein sequence information (29–32), whereas algorithms for the prediction of type 6 secretion system (T6SS) proteins are not available. These *ab initio* methods for the prediction of T3SS proteins use machine learning techniques to train predictive models on sets of known secreted proteins in a range of bacteria. In this work, we used two such algorithms and reported their consensus predictions. The first method, EffectiveT3 (29), uses amino acid composition and di- and tri-peptide pattern frequency as inputs to a number of different machine learning algorithms, whereas the second method, developed by Lower and Schneider (30), uses a sliding window of protein sequence as input to neural networks and support vector machine algorithms. For the EffectiveT3 software (29), we used a probability cutoff of 0.995 for classification as a T3SS protein, whereas for the method described in Ref. (30) we used a 0.6 cutoff.

**Y2H Screens to Identify Host-Pathogen Protein-Protein Interactions (PPIs)**—*B. mallei* genes were cloned from *B. mallei* ATCC 23344 genomic DNA using the Gateway™ Entry system, pDONR™/Zeo (Invitrogen, Carlsbad, CA). In addition to clones of intact genes, 11 putative virulent factors were cloned in more than one truncated form (domain fragments). The oligonucleotides of gene-specific regions for polymerase chain reaction (PCR) amplification were designed using the “OligoCalc: Oligonucleotide Properties Calculator” Web tool (<http://www.basic.northwestern.edu/biotools/oligocalc.html>). To ensure the specificity of the PCRs, the gene-specific regions were designed at their salt-adjusted melting temperature near 60 °C. PCR primers incorporated the following forward and reverse recombination cloning sequences: *attB1* (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTC-3') and *attB2* (5'-GGGGACCACTTTGTACAAGAAAGCTGGGTC-3'), respectively. The oligonucleotides were purchased from Invitrogen. The selected open reading frame (ORF) sequences were amplified by PCR in a 96-well format. The PCR amplification was performed in a volume of 50 µl containing 2 ng of genomic DNA, 200

<sup>1</sup> The abbreviations used are: CFU, Colony-forming units; GO, Gene ontology; IACUC, Institutional Animal Care and Use Committee; IBC, Institutional Biosafety Committee; KEGG, Kyoto encyclopedia of genes and genomes; LB, Luria-Bertani; NCBI, National Center for Biotechnology Information; PPI, Protein-protein interaction; T3SS, Type 3 secretion system; T6SS, Type 6 secretion system; USAMRIID, United States Army Medical Research Institute for Infectious Diseases; Y2H, Yeast two-hybrid.

nM of each primer, 0.8 mM dNTPs, and Phusion DNA polymerase (New England Biolabs, Ipswich, MA). After denaturation at 98 °C for 1 min, the incubation conditions were performed as follows: 25 cycles of 30 s at 98 °C, 30 s at 55 °C, 1 min at 72 °C, and a final 10 min incubation at 72 °C.

The amplified ORFs were cloned into pDONR™/Zeo by recombination reactions using the BP Clonase™ II Enzyme Mix (Invitrogen). The reactions were performed with 150 ng of amplified inserts and 150 ng of pDONR™/Zeo in TE buffer (pH 8.0) and incubated overnight at 25 °C. The reaction mixtures were incubated with 1 μg proteinase K for 30 min at 37 °C. Two microliters of the reaction mixtures were transformed into chemically competent DH5α. Transformants were plated on Luria-Bertani (LB) agar-containing zeocin at a final concentration of 30 μg/ml and incubated overnight. Selected transformants were grown in LB medium with zeocin to prepare frozen cultures and for plasmid preparation. The cloned plasmids were validated by Sanger DNA sequencing.

Target genes in entry clones were subcloned into NH<sub>2</sub>-terminal and COOH-terminal DNA-binding domain (bait) Y2H vectors, pGBGT7g and pGBKCg (33), using Gateway LR reactions (34). LR reactions were performed using equal quantities of the entry clone plasmid and the destination vector of 35 fmol in a 10 μl volume and 2 μl (1 × unit) LR clonase II enzyme mix. Reactions were incubated for 16 h at 37 °C and terminated by an incubation with 2 μg proteinase K for 10 min at 37 °C. The destination clones were transformed into DH5α and plated on LB agar containing 30 μg/ml kanamycin. After bacterial transformation, miniprep plasmid DNAs of all the bait clones were transformed into Y2H yeast strain AH109 (*MAT-α*) (35), as previously described in Ref. (36).

Before the two-hybrid analyses, all bait yeast strains were examined for auto-activation, *i.e.* detectable bait-dependent reporter gene activation in the absence of any prey interaction partner. Weak to intermediate strength auto-activator baits can be used in two-hybrid array screens because the corresponding bait-prey interactions confer stronger signals than the auto-activation background. We used the *HIS3* reporter gene and suppressed auto-activation by adding different concentrations of 3-amino-1,2,4-triazole (3-AT), a competitive inhibitor of *HIS3*. All of the *B. mallei* baits were inspected for auto-activation on plates containing different concentrations of 3-AT. The lowest concentration of 3-AT that suppressed growth for the interaction screen was used (see below), because it avoided background growth while still detecting true interactions.

A haploid yeast strain expressing a single *B. mallei* protein as bait was used to separately perform protein interaction screens with human and mouse normalized universal two-hybrid cDNA libraries (catalog nos. 630480 and 630482, Clontech). The bait and prey (cDNA library) yeast culture was grown and mixed (at a 1:1 ratio) and plated on YEPDA agar plates (as recommended by the cDNA library provider, Clontech). YEPDA agar plates were incubated at 30 °C for 6 h or overnight at room temperature. During this process, both prey and bait plasmids were combined in the diploid yeast cells by yeast mating. The cells from the mating plates were collected and transferred onto interaction selection yeast synthetic medium with pre-defined concentrations of 3-AT (media lacking tryptophan, leucine, and histidine plus 3-AT), and plates were incubated at 30 °C for 4–6 days. The interaction selection plates that showed colony growth but no colonies on control plates (bait mated to empty prey vector) were identified as the two-hybrid positive yeast clones. Positive yeast colonies were selected either manually or using robotics and subjected to yeast colony PCR followed by DNA sequencing to identify the interacting preys (36).

**Mapping of Host Proteins to Pathways and Networks**—All computational analyses were performed in R, using Bioconductor/BioMart (37, 38) for GO and KEGG analyses and the “survival” package for

survival analysis (Terry Therneau, “A Package for Survival Analysis in S,” R package version 2.36–14, 2012). All networks were plotted using Cytoscape (39).

**Protein–Protein Interaction Data**—To identify conserved PPIs between the *B. mallei*-human and *B. mallei*-murine PPI data sets, we first identified human-murine orthologs of all human and murine proteins interacting with *B. mallei*, using the National Center for Biotechnology Information (NCBI) HomoloGene database of homologs (<http://www.ncbi.nlm.nih.gov/homologene>) (40, 41). Then, we denoted host-pathogen interactions in which human proteins interacted with the same *B. mallei* proteins as their murine orthologs as conserved interactions. For example, the interaction between *B. mallei* protein BMAA0553 and human protein NPC2 was conserved because BMAA0553 also interacted with murine protein Npc2, and NPC2 and Npc2 are defined as human-murine orthologs in the NCBI HomoloGene database.

To evaluate the influence of the pathogen on host proteins, including the relationships among host proteins targeted by the pathogen’s proteins, we mapped the identified host proteins interacting with *B. mallei* proteins onto a human PPI network (42). The human PPI network consists of 81,226 physical PPIs among 13,382 human proteins. Of the 170 human proteins found to interact with BMAA0728, BMAA1865, or BMAA0553, 125 (~74%) were found in the PPI network. Currently available murine PPIs data are largely incomplete and sparse, so we decided not to map the murine proteins onto a murine PPI network. Instead, we used the NCBI HomoloGene database to identify human orthologs of murine proteins that interacted with *B. mallei* proteins and mapped the identified orthologs onto the human PPI network. Of the 191 murine proteins interacting with BMAA0728, BMAA1865, or BMAA0553, 166 (87%) had human orthologs listed in HomoloGene. Of the 336 human proteins (170 human proteins + 166 murine orthologs) interacting with three *B. mallei* proteins of interest, 244 (~73%) were found in the human PPI network.

To assign protein-domain annotation to *B. mallei*, human, and murine proteins, we used the high-confidence annotations from PFAM-A database (43).

**Gene Ontology (GO) Enrichment Analysis**—We used the hypergeometric distribution to assess the statistical significance of observing a given GO term enrichment in our data. For a given host (human or murine), we first counted the total number of proteins (*N*) annotated with any GO term. Then, for each GO term, *g*, we counted the total number of proteins (*m*) annotated with *g*. Next, we calculated the probability (*p*) of observing the same, or higher, enrichment in term *g* by chance, as follows:

$$p = 1 - \sum_{i=0}^{k-1} \frac{\binom{m}{i} \binom{N-m}{n-i}}{\binom{N}{n}} \quad (\text{Eq. 1})$$

where *n* represents the number of host proteins interacting with *B. mallei* proteins annotated with any GO term and *k* denotes the number of host proteins interacting with *B. mallei* proteins annotated with *g*. For the total number of human proteins, we used the set of all human and murine proteins available in Bioconductor and BioMart (37, 38) and annotated with at least one GO term, excluding root GO terms.

In this analysis, we only considered GO annotations that were enriched at a false discovery rate control level of 0.2 or below, after accounting for multiple hypothesis testing using the Benjamini and Hochberg multiple test correction (44).

**Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathways Enrichment Analysis**—KEGG pathways are not independent of each other, and a large number of pathways share a statistically significant number of genes (45). Thus, analyses that search for statistically significantly enriched pathways may identify multiple pathways as



significant simply because they contain a large number of shared genes (e.g. the long-term depression pathway shares ~60% of its genes with the vascular smooth muscle contraction pathway) and not necessarily because of shared functionality. As these nonpathway-specific genes are not necessarily a good representation of a specific biological process, we focused on the identified pathways that are enriched with genes that mainly belong to that pathway or a very small number of other pathways. Thus, to identify gene-specific KEGG pathways, we used the hypergeometric distribution with adjusted parameters to account for the specificity of each gene. For a given host, for each gene  $i$  in KEGG, we counted the number of pathways in which the gene appears and denoted this number as  $k_i$ . We then defined the probability of each gene  $i$  being specific ( $p_i$ ) as follows:

$$p_i = \frac{1}{k_i} \quad (\text{Eq. 2})$$

Next, we defined the adjusted population size of all host proteins in KEGG ( $A_N$ ) as follows:

$$A_N = \sum_j^{N_{KEGG}} p_j \quad (\text{Eq. 3})$$

where the summation runs over all  $N_{KEGG}$  host proteins that appear in at least one KEGG pathway. We then defined the adjusted number of genes in a pathway  $P$  ( $A_m$ ) as:

$$A_m = \sum_j^{n_P} p_j \quad (\text{Eq. 4})$$

where the summation runs over all  $n_P$  host proteins that appear in pathway  $P$ . The adjusted total number of interacting host proteins found in KEGG pathways ( $A_n$ ) was defined as:

$$A_n = \sum_j^{n_H} p_j \quad (\text{Eq. 5})$$

where the summation runs over all  $n_H$  host proteins that interact with any *B. mallei* protein and appears in any pathway. Finally, the adjusted number of host proteins interacting with *B. mallei* proteins in pathway  $P$  ( $A_k$ ) was defined as:

$$A_k = \sum_j^{n_{HP}} p_j \quad (\text{Eq. 6})$$

where the summation runs over all  $n_{HP}$  interacting host proteins in pathway  $P$ .

Because the adjusted numbers  $A_N$ ,  $A_m$ ,  $A_n$ , and  $A_k$  are real numbers, we converted each of them to its next smallest integer,  $N$ ,  $m$ ,  $n$ , and  $k$ , respectively, and calculated the probability of observing the same or higher specific enrichment in pathway  $P$  purely by chance using Eq. 1. For the total set of human proteins, we used all unique human proteins that belong to at least one KEGG pathway.

**Bacterial Strains, Plasmids, and Growth Conditions**—Table I shows the plasmids and bacterial strains used in this study. *E. coli* was grown at 37 °C on LB (Lennox) agar or in LB broth. *B. mallei* was grown at 37 °C on LB agar with 4% glycerol or in LB broth with 4% glycerol. All bacterial strains were grown in broth for ~18 h with constant agitation at 250 revolutions/min. When appropriate, antibiotics were added at the following concentrations: 25 μg kanamycin per milliliter for *E. coli* and 15 μg kanamycin and 25 μg polymyxin B per milliliter for *B. mallei*. Phosphate-buffered saline (PBS) was used

to make serial dilutions of saturated bacterial cultures, and the number of cfu present in the starting culture was determined by spreading 100 μl aliquots onto agar media and incubating for 24–48 h. A 20 mg/ml stock solution of the chromogenic indicator 5-bromo-4-chloro-3-indolyl-b-D-galactoside was prepared in *N,N*-dimethylformamide, and 40 μl were spread onto the surface of the plate medium for blue/white screening in *E. coli* TOP10. All manipulations with *B. mallei* were carried out in a class II microbiological safety cabinet located in a designated biosafety level 3 laboratory.

**PCR of Internal Gene Fragments**—Table I also shows the PCR primers used to amplify internal gene fragments of BMAA0728, BMAA1865, and BMAA0553. PCR amplifications were performed in a final reaction volume of 100 μl containing 1× TaqPCR Master Mix (Qiagen), 1 μM oligodeoxyribonucleotide primers, and ~200 ng of *B. mallei* ATCC 23344 genomic DNA. Bacterial genomic DNA was prepared using a previously described protocol (49). PCR cycling was performed using a PTC-150 MiniCycler with a Hot Bonnet accessory (MJ Research, Inc.) and heated to 97 °C for 5 min. This was followed by 30 cycles of a three-temperature cycling protocol (97 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min) and 1 cycle at 72 °C for 10 min. PCR products were sized and isolated using agarose gel electrophoresis, cloned using the pCR2.1-TOPO TA Cloning Kit (Invitrogen), and transformed into chemically competent *E. coli* TOP10.

**DNA Manipulation and Plasmid Conjugation**—The restriction enzymes *EcoRI* and *NotI*, Antarctic phosphatase, and T4 DNA ligase were purchased from Roche Molecular Biochemicals and used according to the manufacturer's instructions. DNA fragments used in cloning procedures were excised from agarose gels and purified with a GeneClean III kit (Q BIOgene). Plasmids were purified from overnight cultures using Wizard Plus SV Minipreps (Promega). Plasmids pEXKm5-0728, pEXKm5-1865, and pEXKm5-0553 (Table I) were electroporated into *E. coli* S17-1 (12.25 kV/cm) and conjugated with *B. mallei* ATCC 23344 for 8 h, as previously described (8). Polymyxin B was used to counter-select *E. coli* S17-1 donor strains.

**Animal Experiments**—Female BALB/c mice were obtained from Charles River Laboratories (National Cancer Institute) and were 6–8 week old at the time of use. Animals were provided with rodent feed and water *ad libitum* and maintained on a 12:12-hour light-dark cycle. Whole body exposures to aerosols of *B. mallei* were performed as previously described by Roy *et al.* (50). Research was conducted in compliance with the Animal Welfare Act and other federal statutes and regulations relating to animals and experiments involving animals and adhered to principles stated in the National Institutes of Health *Guide for the Care and Use of Laboratory Animals* (<http://www.nap.edu/readingroom/books/labrats/chaps.html>). The United States Army Medical Research Institute for Infectious Diseases (USAMRIID) Institutional Animal Care and Use Committee (IACUC) reviewed and approved the animal protocol entitled "Evaluation of *Burkholderia mallei* Secretion Inhibitors and Identification of *B. mallei* and *B. pseudomallei* Virulence Determinants in Mice and Hamsters." The animal protocol number assigned to this protocol by the USAMRIID IACUC was AP-09-045. The facility where this research was conducted is fully accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International.

**Survival Analysis**—We used the Kaplan-Meier method (51) to generate survival curves from the number of surviving mice as a function of time. To this end, Kaplan-Meier survival curves can be interpreted as the probability of animals surviving a given length of time. This probability is defined as the cumulative probability of surviving  $k$  time periods  $S(k)$ , as follows:

$$S(k) = p_1 \cdot p_2 \cdot p_3 \cdots p_k \quad (\text{Eq. 7})$$

where  $p_i$  represents the proportion of animals surviving period  $i$ . Here,  $p_i$  was defined as follows:

TABLE I

Primers, plasmids, and strains used in the mouse aerosol challenge model experiment. *Km<sup>r</sup>*, kanamycin resistant. *Km<sup>s</sup>*, kanamycin sensitive. *Pm<sup>r</sup>*, polymyxin B resistant. *Pm<sup>s</sup>*, polymyxin B sensitive

Primer, plasmid, or strain	Relevant characteristics	Source or reference
<b>Primers</b>		
0728-up	5'-GCGGCCGCACACATCCGGCAGCAACAG-3'	This study
0728-dn	5'-GCGGCCGCTCTCCAAAAGAATCCACAAC-3'	This study
1865-up	5'-GAATTCATCGACGAAATCCCGA-3'	This study
1865-dn	5'-GAATTCGAGCTTGGCGATCTCGAC-3'	This study
0553-up	5'-GAATTCGAAGCCGTCGAGCTTTC-3'	This study
0553-dn	5'-GAATTCGAATAGTTCGCGCTGC-3'	This study
<b>Plasmids</b>		
pCR2.1-TOPO	3,931-bp TA vector; pMB1 <i>oriR</i> ; <i>Km<sup>r</sup></i>	Invitrogen
pCR2.1-0728	pCR2.1-TOPO containing 298-bp PCR product generated with 0728-up and 0728-dn	This study
pCR2.1-1865	pCR2.1-TOPO containing 297-bp PCR product generated with 1865-up and 1865-dn	This study
pCR2.1-0553	pCR2.1-TOPO containing 535-bp PCR product generated with 0553-up and 0553-dn	This study
pEXKm5	Mobilizable <i>Km<sup>r</sup></i> suicide vector	(46)
pEXKm5-0728	pEXKm5 derivative containing <i>NotI</i> insert from pCR2.1-0728	This study
pEXKm5-1865	pEXKm5 derivative containing <i>EcoRI</i> insert from pCR2.1-1865	This study
pEXKm5-0553	pEXKm5 derivative containing <i>EcoRI</i> insert from pCR2.1-0553	This study
<b><i>E. coli</i></b>		
TOP10	General cloning and blue/white screening	Invitrogen
S17-1	Mobilizing strain with transfer genes of RP4 integrated on chromosome; <i>Km<sup>s</sup></i> , <i>Pm<sup>s</sup></i>	(47)
<b><i>B. mallei</i></b>		
ATCC 23344	Type strain, isolated in 1944 from a human case of glanders; <i>Km<sup>s</sup></i> , <i>Pm<sup>r</sup></i>	(48)
DDA0728	ATCC 23344 derivative with pEXKm5-0728 integrated into the chromosome; <i>BMAA0728<sup>-</sup></i>	This study
DDA1865	ATCC 23344 derivative with pEXKm5-1865 integrated into the chromosome; <i>BMAA1865<sup>-</sup></i>	This study
DDA0553	ATCC 23344 derivative with pEXKm5-0553 integrated into the chromosome; <i>BMAA0553<sup>-</sup></i>	This study

$$p_i = \frac{r_i - d_i}{r_i} \quad (\text{Eq. 8})$$

where  $r_i$  and  $d_i$  denote the number of alive mice at the beginning of time period  $i$  and the number of deaths within period  $i$ , respectively.

To compare murine survival curves, we used the following log-rank test:

$$\chi^2(\text{logRank}) = \frac{\left(O_1 - \sum_{i=1}^k \frac{d_i}{r_i} r_{1i}\right)^2}{\sum_{i=1}^k \frac{d_i}{r_i} r_{1i}} + \frac{\left(O_2 - \sum_{i=1}^k \frac{d_i}{r_i} r_{2i}\right)^2}{\sum_{i=1}^k \frac{d_i}{r_i} r_{2i}} \quad (\text{Eq. 9})$$

where 1 and 2 represent the set of mice infected with *B. mallei* mutant (set 1) and the set of mice infected with wild-type *B. mallei* (set 2),  $r_{1i}$  and  $r_{2i}$  represent the number of alive mice from sets 1 and 2 at time period  $i$ , respectively, and  $O_1$  and  $O_2$  represent the total number of observed events in each of the two sets, respectively.

**Biosafety and Biosecurity**—We do not anticipate that this manuscript provides knowledge, products, or technologies that could be directly misapplied by others to pose a threat to public health and safety, agricultural crops and other plants, animals, the environment, or materiel. The Institutional Biosafety Committee (IBC) that approved

this work is composed of members of the USAMRIID research staff, Commander's office, and qualified representatives from external institutions and is tasked to provide local, institutional oversight of research utilizing recombinant DNA. The USAMRIID IBC was established under the United States (U.S.) National Institutes of Health Guidelines for Research Involving Recombinant DNA Molecules.

RESULTS

Figure 1 outlines the components and steps of our combined *in silico*, *in vitro*, and *in vivo* effort to identify *B. mallei* virulence factors via a systematic analysis of host-pathogen protein interactions. In the first step, we used a set of bioinformatics tools for *in silico* identification of putative virulence factors, including orthology detection based on pathogen proteins experimentally linked to pathogenicity in other species and *ab initio* prediction methods. We complemented these methods with literature review of published experimental data, resulting in a pruned list of 49 proteins. In the second step, we screened the selected proteins *in vitro* for potential interaction partners, using Y2H assays against normalized human and murine whole proteome libraries, and identified host-pathogen protein-protein interactions (PPIs). In the third

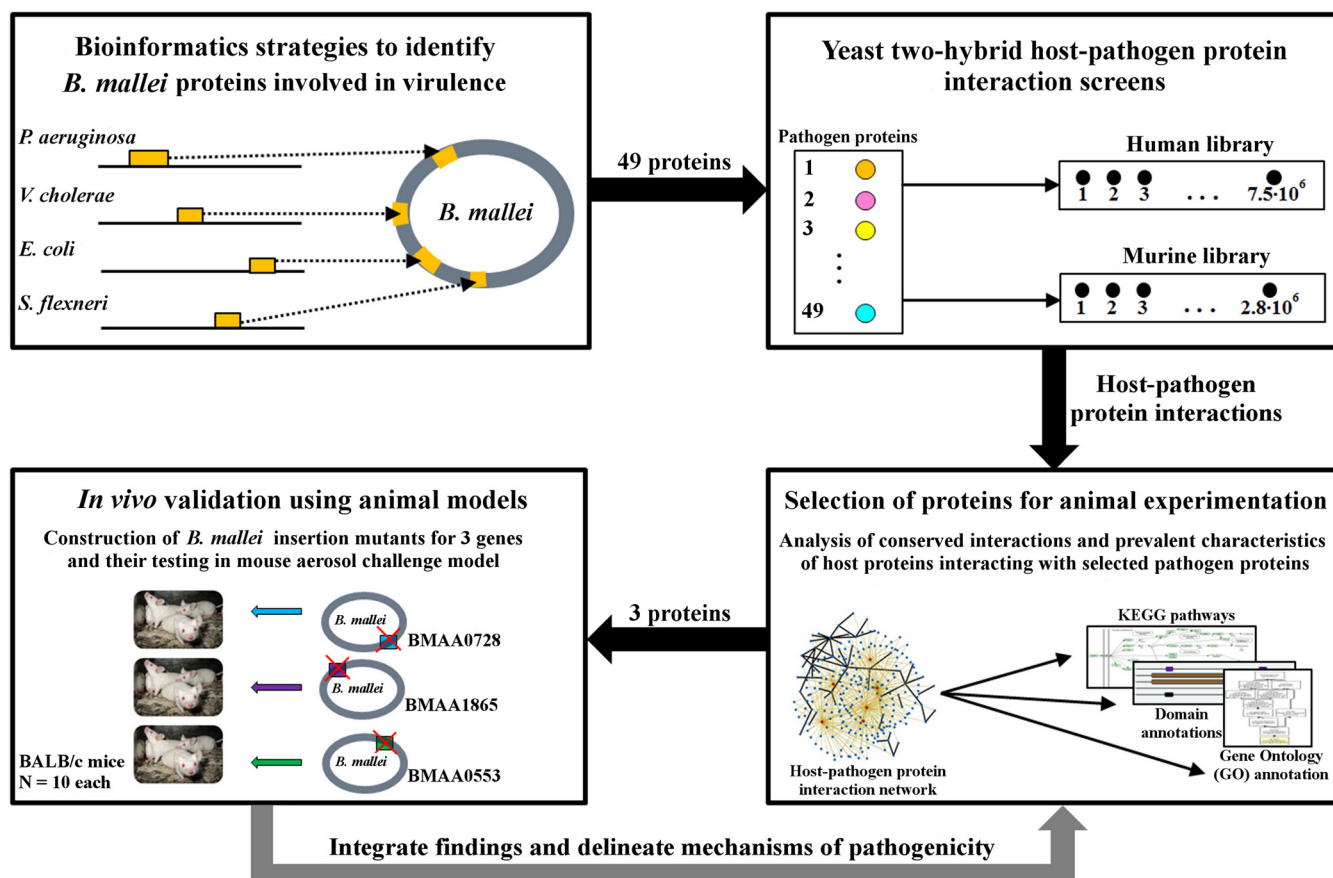


FIG. 1. **The combined *in silico*, *in vitro*, and *in vivo* approach.** We used bioinformatics tools to identify 49 proteins putatively involved in virulence. These proteins were then screened for interaction partners using yeast two-hybrid (Y2H) assays against whole human and whole murine proteome libraries. We mapped the identified host-pathogen protein-protein interactions (PPIs) to host pathways and host PPI networks. The host proteins interacting with *Burkholderia mallei* proteins were characterized with respect to their functions and possible role in pathogenicity. This analysis resulted in the identification of a three promising virulence factor candidates. We verified that mutants lacking these proteins showed attenuated virulence in a mouse aerosol challenge model. Finally, we used all obtained data to delineate mechanisms of pathogenicity and generated hypotheses about the potential roles of these three proteins in *B. mallei* invasion of the host cell.

step, we mapped interactions conserved in both hosts to canonical pathways and human PPI networks and proposed three *B. mallei* proteins (encoded by BMAA0728, BMAA1865, and BMAA0553), as likely to interfere with host processes and pathways associated with bacterial pathogenicity. In the fourth step, we constructed *B. mallei* insertion mutant strains for each of the three genes and verified their biological relevance *in vivo* using a BALB/c murine aerosol infection model, where we ascertained and compared virulence attenuation of the mutant strains with the *B. mallei* ATCC 23344 wild-type strain. The results of these studies are detailed below.

***In Silico* Identification of Putative Virulent Factors**—To create an initial set of possible virulence factors, we deployed three independent bioinformatics analyses based on the identification of: 1) genes present in pathogenic and absent in nonpathogenic strains of *B. mallei*, 2) orthologous genes that have been associated with secretion systems in other bacteria, and 3) type 3 secretion system (T3SS)-associated secreted proteins, using amino acid pattern-based *ab initio* methods. This extensive set

of proteins was then examined within the context of literature searches and published microarray experimental data to weigh, rank, and group the collected evidence into a final, pruned set of *in silico* predicted virulence factors.

In the first analysis, we collected genomic data for *B. mallei* and related bacterial genomes using Pathema (27) for finished and draft *B. mallei* genomes and GenBank (26) for all other finished genomes. We used QuartetS (25) to identify orthologous proteins in the seven strains of *B. mallei* reported by Losada *et al.* (52). Table II shows the four pathogenic and three nonpathogenic/attenuated strains used in this study. We identified orthologous pairs of genes between each two strains and identified 78 genes that were either present in all pathogenic strains or overrepresented in the pathogenic strains. Thirty-two of these orthologs were present in our *B. mallei* ATCC 23344 target strain and had no orthologs among the nonpathogenic strains. Even though this set should be enriched with proteins associated with pathogenicity, the bulk of these 32 genes were annotated with hypothetical func-

TABLE II  
Strains of *Burkholderia mallei* used to identify proteins overrepresented in the pathogenic strains

Genome	Size, Mb	PATRIC coding sequences <sup>a</sup>	RefSeq coding sequences <sup>b</sup>	Finished	Pathogenic
<i>B. mallei</i> ATCC 23344	5.83	5724	5024	Yes	Yes
<i>B. mallei</i> NCTC 10229	5.74	5591	5510	Yes	Yes
<i>B. mallei</i> ATCC 10399	5.88	5850	5276	No	Yes
<i>B. mallei</i> PRL-20	5.62	5614	5456	No	Yes
<i>B. mallei</i> SAVP1	5.23	5095	5189	Yes	No
<i>B. mallei</i> 2002721280	5.51	5590	4847	No	No
<i>B. mallei</i> NCTC 10247	5.84	5717	5852	Yes	Attenuated

<sup>a</sup> This column shows the number of coding sequences for each of the genomes that are available in the PATRIC database (53).

<sup>b</sup> This column shows the number of coding sequences available in the National Center for Biotechnology Information Reference Sequence database (54).

tions or with functions that could not readily be associated with pathogenicity (supplemental Table S1) and, hence, their actual role in *Burkholderia* pathogenicity remains to be determined.

In the second analysis, we examined gene clusters associated with bacterial secretion systems (55). In particular, previous experimental work in animal models has indicated that the T3SS (14) and the type 6 secretion system (T6SS) (11, 13) are involved in virulence of *B. mallei* ATCC 23344. Recent publications have also elucidated the role of T3SS (56) and T6SS (11) in the virulence of *B. pseudomallei*. We used available literature and database information to compile a list of known and predicted proteins associated with bacterial secretion. Then, we used QuartetS to identify 24 orthologs of these proteins present in the *B. mallei* ATCC 23344 strain. supplemental Table S2 gives the gene locus, name, secretion system, and evidence used to identify this set of proteins.

Finally, in the third analysis, we used two T3SS protein prediction methods to identify *B. mallei* proteins that share characteristics with known T3SS proteins and, hence, represent potential virulence factors (29, 30). We selected proteins that were common to both prediction methods and located on the second chromosome, as the known *B. mallei* T3SS is located on this chromosome. This analysis identified 37 proteins (supplemental Table S3).

To select proteins for experimental evaluation, we first merged the lists obtained in three independent analyses above, and revised the merged list using published microarray experimental data. In particular, we compiled the available microarray and animal model experiments for *B. mallei* and the closely related bacteria *B. pseudomallei*, *B. cenocepacia*, and *B. thailandensis* (13, 22, 48, 57–59). We identified proteins that were overexpressed during infections and in disease states and, hence, considered relevant for pathogenicity. Next, we used QuartetS to detect orthologs of the overexpressed proteins that are present in the merged list of *B. mallei* proteins. This procedure provided additional evidence for pathogenicity among of our identified proteins. Finally, we individually assessed and classified each of the proteins and created a final list of 49 potential virulence factors for experimental evaluation. The final list was divided into four groups

based on the assessed reliability of evidences. Table III shows the 49 proteins, 18 of which we classified into the most reliable group (*group 1*). All identified proteins and their associated evidence for involvement in pathogenicity are further detailed in supplemental Table S4.

*Y2H Screening Results and Data Analysis*—Out of the 49 selected proteins, 43 of the corresponding genes (or gene fragments) were successfully cloned, prepared, and tested in Y2H assays against both human and murine whole proteome libraries. The raw interaction data set consisted of 637 PPIs between human and *B. mallei* proteins and 900 PPIs between murine and *B. mallei* proteins. We filtered these data sets by removing protein interactions between *B. mallei* and “sticky” host proteins known to be indiscriminate binders (F. Schwarz and M. Koegl, unpublished data). The filtered data consisted of 600 unique interactions between human and *B. mallei* proteins and 846 unique interactions between murine and *B. mallei* proteins (supplementary Data set S1 and supplementary Data set S2). Fig. 2 shows the host-pathogen protein interactions and Table III lists the number of Y2H interactions we detected for each tested protein. These 1446 PPIs involved 26 unique *B. mallei* proteins, 409 unique human proteins, and 574 unique murine proteins. Nineteen of the 26 *B. mallei* proteins interacted with proteins of both hosts, whereas two *B. mallei* proteins interacted only with human proteins and five *B. mallei* proteins interacted only with murine proteins. Furthermore, we found 33 conserved PPIs between these two data sets, *i.e.* interactions in which the human proteins interacted with the same *B. mallei* proteins as their murine orthologs. Finally, we found that 16 human proteins (9%) and 21 murine proteins (11%) interacted with multiple pathogen proteins, implying that these proteins may be involved in more than one stage of the infectious process and, potentially, have a multivariate pathogenic functionality.

*Selection of Proteins for Animal-model Experiments*—To identify the likely virulence factors for validation in animal experiments, we focused on the 19 *B. mallei* proteins that interacted with both hosts. We considered this to be a reasonable down selection criterion because murine animal models of *B. mallei* infection share many pathophysiological characteristics with human infections (62, 63). Thirteen of



TABLE III

List of proteins selected for experimental yeast two-hybrid evaluation. \*Proteins that have been experimentally linked to a particular secretion system. Genes associated with secretion system are classified according to their secretion system type and chromosomal locations

Group	Locus tag	Protein name	Secretion system	Successfully cloned	Number of PPIs detected		
					Human	Murine	Conserved
1	BMAA0410*	VgrG	6/c3 <sup>a</sup>	Yes			
1	BMAA0445*	VgrG	6/c2 <sup>a</sup>	Yes	2	1	
1	BMAA0446*	VgrG	6/c2 <sup>a</sup>	Yes	7	7	
1	BMAA0737*	VgrG	6/c1 <sup>a</sup>	No			
1	BMAA1901*	VgrG	6/c4 <sup>a</sup>	Yes			
1	BMAA0749*	BimA	5 <sup>b</sup>	Yes	80	199	5
1	BMA0278*	PilA	-	Yes	42	64	2
1	BMAA0079*	PlcN	2 <sup>c</sup>	Yes			
1	BMAA1521*	BopA	3 <sup>d</sup>	Yes	1	1	
1	BMAA1523*	BopE	3 <sup>d</sup>	Yes		27	
1	BMAA1525*	BapB	3 <sup>d</sup>	Yes	9	13	
1	BMAA1528*	BipD	3 <sup>d</sup>	Yes	41	49	3
1	BMAA1530*	BipC	3 <sup>d</sup>	Yes	2		
1	BMAA1531*	BipB	3 <sup>d</sup>	Yes	24	25	1
1	BMA0855*	VgrG Rhs domain	6 <sup>e</sup>	No			
1	BMAA1269	VgrG	6 <sup>e</sup>	Yes	1	1	
1	BMAA1538*	BsaU	3 <sup>d</sup>	Yes	14	42	1
1	BMAA0729*	TssM	6/c1 <sup>a</sup>	Yes		2	
2	BMA2097	Hypothetical protein	3 <sup>f</sup>	Yes			
2	BMAA1865	Hypothetical protein	3 <sup>f</sup>	Yes	61	89	6
2	BMAA1899	Hypothetical protein	3 <sup>f</sup>	No			
2	BMA0666	CysD-1	3 <sup>f</sup>	Yes		3	
2	BMAA1487	Methyl-accepting chemotaxis protein	3 <sup>f</sup>	No			
2	BMA3281	FlIF	3 <sup>f</sup>	Yes	42	43	1
2	BMA1750	HrpA	3 <sup>f</sup>	Yes			
2	BMAA0728	TssN	3 <sup>g</sup>	Yes	72	87	3
2	BMA2787	IpgF	3 <sup>f</sup>	Yes			
2	BMA0267	Pseudogene	-	Yes	10	4	1
2	BMAA0038	Hypothetical protein	-	Yes			
2	BMAA1662	Response regulator	3 <sup>f</sup>	Yes		19	
2	BMA0429	Cmk	3 <sup>f</sup>	Yes	63	58	4
2	BMA2469	Tkt	3 <sup>f</sup>	Yes	24	6	
2	BMAA0238	Hypothetical protein	2 <sup>h</sup>	Yes	25	28	2
2	BMAA1173	Hypothetical protein	-	Yes			
2	BMA0886	AcpA	2 <sup>h</sup>	No			
3	BMAA1618	Hypothetical protein	3 <sup>f</sup>	Yes			
3	BMAA1619	Hypothetical protein	3 <sup>f</sup>	Yes	14	9	1
3	BMAA0679	Chemotaxis protein CheC	-	Yes	7		
3	BMAA0553	Ser/Thr protein phosphatase	2 <sup>h</sup>	Yes	59	60	3
3	BMAA1559	Thermolysin metalloproteinase	2 <sup>h</sup>	Yes			
3	BMAA0582	Hypothetical protein	3 <sup>f</sup>	Yes			
4	BMAA1566	Serine-type carboxypeptidase	3 <sup>f</sup>	Yes			
4	BMAA1641	Hypothetical protein	3 <sup>f</sup>	Yes			
4	BMAA0605	Hypothetical protein	3 <sup>f</sup>	Yes			
4	BMAA1111	Hypothetical protein	3 <sup>f</sup>	No			
4	BMAA1648	Hypothetical protein	3 <sup>f</sup>	Yes		9	
4	BMAA1997	Phosphatidylserine decarboxylase	3 <sup>f</sup>	Yes			
4	BMAA2052	Polysaccharide deacetylase	3 <sup>f</sup>	Yes			
4	BMAA0418	NAD-dependent deacetylase	3 <sup>f</sup>	Yes			

<sup>a</sup> Gene located within the loci of a type 6 secretion system (T6SS) clusters (*cluster 1* [c1]: BMAA0744-0729, *cluster 2* [c2]: BMAA0438-0455, *cluster 3* [c3]: BMAA0396-0412, and *cluster 4* [c4]: BMAA1897-1915) (13).

<sup>b</sup> Type 5 secretion system effector (auto-transporter), located adjacent to T6SS c1.

<sup>c</sup> Predicted phospholipase, a putative type 2 secretion system (T2SS) effector (60).

<sup>d</sup> Gene located within the locus of the type 3 secretion system (T3SS) (animal pathogen): BMAA1520-1552 (14).

<sup>e</sup> T6SS-associated effector not located in any known or annotated T6SS cluster.

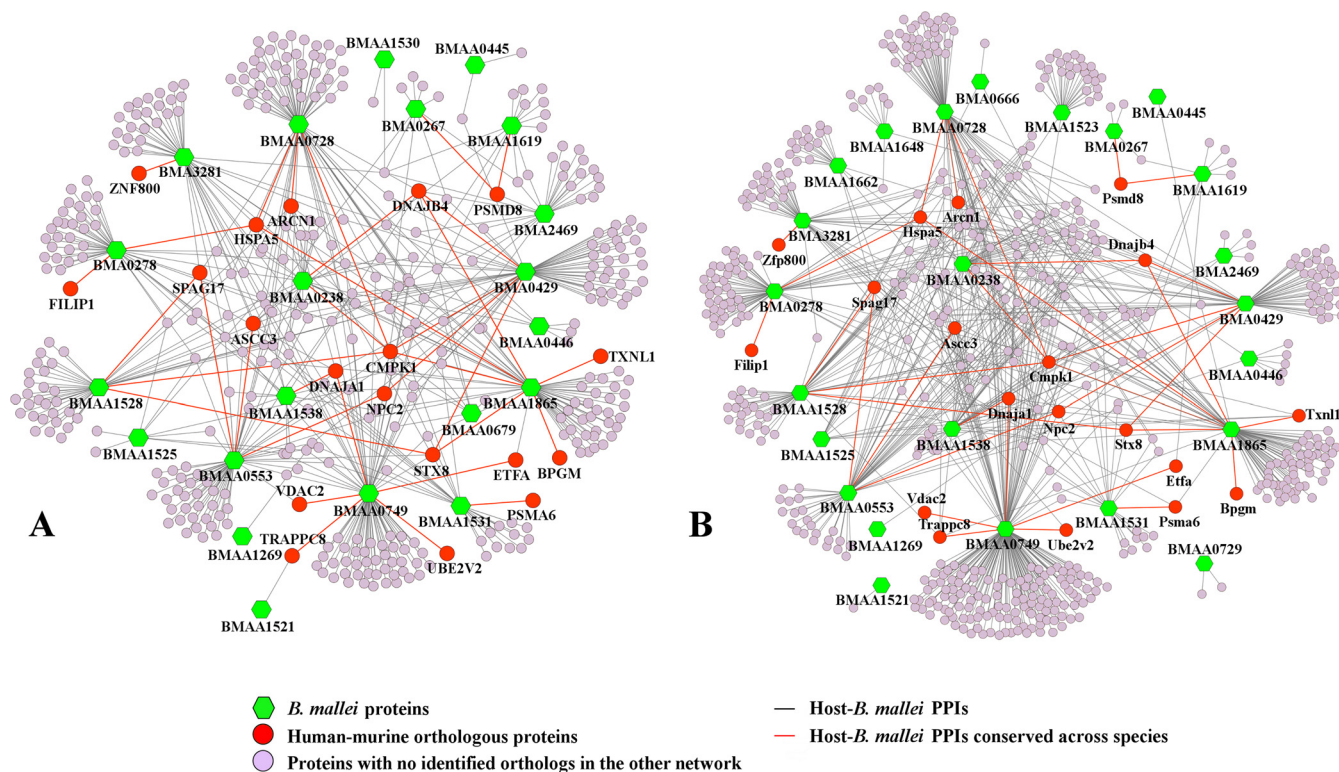
<sup>f</sup> T3SS-predicted effector not located in any known or annotated T3SS.

<sup>g</sup> Predicted as a T3SS effector, but located adjacent to T6SS c1.

<sup>h</sup> Predicted to be secreted by the general Sec pathway and therefore possibly related to the type 2 secretion system.

<sup>i</sup> Protein located within the locus of the putative plant-pathogenicity-associated T3SS: BMAA1617-1640 (61).





**FIG. 2. Y2H host-pathogen protein-protein interactions.** Our yeast two-hybrid (Y2H) assays against whole human and murine proteome libraries identified 600 protein-protein interactions (PPIs) between 21 *B. mallei* proteins and 409 human proteins (A) and 846 PPIs between 24 *B. mallei* proteins and 574 murine proteins (B). Green nodes represent *B. mallei* proteins, whereas purple and red nodes represent host proteins. There are 33 conserved interactions between these two data sets (red edges), i.e. in 33 PPIs, human proteins (red nodes in A) interacted with the same *B. mallei* proteins as their murine orthologs (red nodes in B).

these 19 proteins participated in interactions that were conserved between the two different hosts, and eight of them participated in more than one conserved PPI. Three of these proteins, PilA, BipD, and BimA, have been previously characterized as virulence factors in animal models (64–68). The other five proteins represent targets that have not been previously studied in animal models of glanders or melioidosis, and we considered them as promising proteins for animal model experiments. Because one of these five proteins consists of only 62 amino acids, it was too small for the creation of insertion mutants and was removed from further consideration. Another protein, cytidylate kinase BMA0429, was identified as an essential *B. mallei* protein [it is an ortholog of STY0980, an essential gene of *Salmonella enterica* Typhimurium (69)], and, as such, was not likely to be successfully mutated and was not considered further. Thus, we identified BMAA1865, BMAA0728, and BMAA0553 as three relatively uncharacterized and novel *B. mallei* candidate virulence factors.

We initially examined PPIs among these three pathogen proteins and host proteins to identify sets of overrepresented Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways among the host proteins. We found that each of the three proteins interacted with a statistically significant ( $p \leq 0.02$ ) number of host proteins that are

exclusively part of pathways related to ubiquitin processing, receptor signaling, and protein processing in the endoplasmic reticulum, i.e. they interacted with host proteins that are primarily part of these pathways and do not appear in a large number of other pathways (see Experimental Procedures). In summary, we found that each of these three proteins interacted with host proteins involved in biological processes and/or molecular pathways often implicated in pathogenicity and were, therefore, of interest for experimental validation in a murine model of glanders.

**Mutants Showed Attenuated Virulence in a BALB/c Aerosol Challenge Model**—Based on the above evidence, we created separate insertion mutants in the *B. mallei* ATCC 23344 strain for each of the three genes (BMAA0728, BMAA1865, and BMAA0553). These mutants did not seem to have any ascertainable *in vitro* growth defect and appeared fully competent. We subsequently used a mouse aerosol challenge model to assess potential attenuation afforded by these *B. mallei* mutants. The fully virulent wild-type *B. mallei* ATCC 23344 strain was used as a positive control.

Forty BALB/c mice (10 mice for each of the three mutant strains and 10 mice for the wild-type strain) were exposed to inhalation doses of  $\geq 10^4$  colony-forming units (cfu; corresponding to  $\geq 10$  LD<sub>50</sub> for the wild-type strain) and monitored

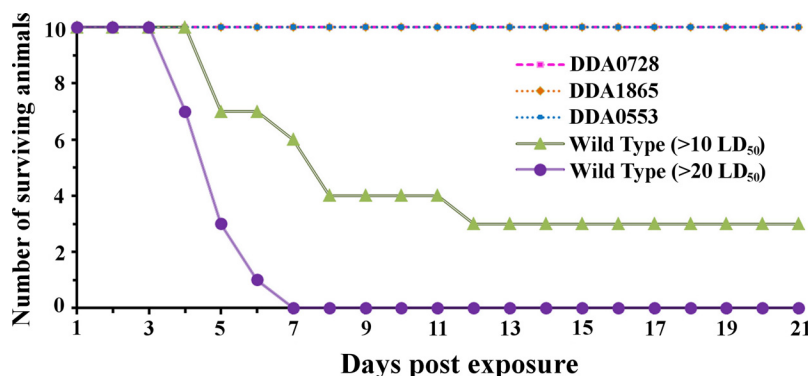


FIG. 3. **Mouse aerosol challenge model results.** In the first set of experiments, 40 BALB/c mice (10 mice for each of the three *B. mallei* mutants + 10 mice exposed to the wild-type strain) received inhalation doses of  $\geq 2 \times 10^4$  colony-forming units (cfu), corresponding to  $\geq 10$  LD<sub>50</sub> for the wild-type strain. In the second set of experiments, additional 40 BALB/c mice received inhalation doses of  $\geq 2 \times 10^5$  cfu, corresponding to  $\geq 20$  LD<sub>50</sub> for the wild-type strain. In both experiments, animals were monitored for 21 days. At the end of the first set of experiments, all 30 mice exposed to mutant strains survived 21 days post-exposure (pink, orange, and blue lines), whereas seven mice exposed to the wild-type strain did not survive (green line). At the end of the second set of experiments, all 30 mice exposed to mutant strains also survived 21 days post-exposure (pink, orange, and blue lines), whereas all 10 mice infected with *B. mallei* ATCC 23344 died (purple line). There was a statistically significant difference in the survival rate of mice exposed to mutant strains and mice exposed to the wild-type strain in both sets of experiments ( $p = 4 \times 10^{-3}$  and  $p = 9 \times 10^{-5}$  for the first and second set of experiments, respectively). These results support the notion that each of the three mutants attenuates virulence when infected in mice via the inhalational route of infection.

for 21 days. Fig. 3 shows the survival curves of mice exposed to mutant strains and mice exposed to the wild-type strain. For this dose, each of the three mutants appeared to attenuate infection, as all 30 mice exposed to the mutants survived 21 days post-exposure, whereas seven mice exposed to the wild-type strain did not survive. A survival-curve analysis indicated that this difference was statistically significant ( $p = 4 \times 10^{-3}$ ). However, because not all animals exposed to the wild-type challenge died at the considered dose ( $\geq 10^4$  cfu), we repeated the experiment using a higher aerosol challenge dose.

Accordingly, 40 additional BALB/c mice ( $3 \times 10 + 10$ ) were exposed to inhalation doses of  $\geq 2 \times 10^5$  cfu (corresponding to  $\geq 20$  LD<sub>50</sub> for the wild-type strain) and monitored for 21 days. Similarly, we found statistically significant differences ( $p = 9 \times 10^{-5}$ ) in the survival rate of mice exposed to mutants and mice exposed to the wild-type strain. All three of the *B. mallei* mutants appeared to attenuate infection, *i.e.* all mice infected with the mutants survived 21 days post-exposure, whereas all of the mice infected with *B. mallei* ATCC 23344 strain died within 7 days post exposure (Fig. 3). These results support the hypothesis that each of the three mutants attains a virulence defect when infected in mice via the inhalational route of infection.

#### DISCUSSION

*In vitro* high-throughput technologies have become versatile tools for discovery and hypothesis generation involving proteins and protein interactions that underlie infectious pathogenesis (70–72). These technologies have been successfully applied to both viral (73–78) and bacterial (21–23, 79–82) pathogens to shed light on how these organisms interact with and subvert host defense processes. For exam-

ple, Li *et al.* used a small set of conserved herpesvirus-encoded protein kinases to identify key targeted host proteins using human protein microarray experiments (77). The analysis of these interactions identified the DNA damage response pathway and other key host regulators as central targets in sustaining the viral life cycle.

*In Vitro Approaches to Study B. mallei Pathogenesis*—A key component of *B. mallei* pathogenesis is its ability to enter, survive, and replicate within the host cells. Secretion systems play an important role in pathogenicity of the Gram-negative *B. mallei* by allowing for controlled transport and insertion of pathogen proteins into the host cell cytosol. The identities of secretion systems' proteins are not all cataloged, and the interactions between specific *B. mallei* proteins and host proteins and other host factors are not well characterized. Given that most *Burkholderia* spp. are opportunistic infections, one cannot expect all secretion system proteins to have evolved host-specific protein interactions. Instead, the bulk of the interactions must encompass generic host target interactions to broadly cope with multiple and variable hosts, *e.g.* plants, nematodes, and other soil microorganisms, as well as opportunistic infections in mammalian hosts. Even for the equine-adapted and, thus, more genetically constrained, *B. mallei* pathogen, we cannot expect that either human or murine interactions will have high specificity. As Y2H techniques are sensitive to weak interactions, this method is ideally suited to probe interactions nonspecific to a distinct host (83). Although many *Burkholderia* proteins could potentially interact with multiple host proteins, nonsecreted proteins are less likely to encounter host proteins and, hence, focusing on secretion systems' proteins will increase the likelihood of finding biologically important interactions relevant to establishing the

infection. Our approach was to use *in silico* methods to specifically search for virulence factors that are likely to be a part of a secretion system and gauge their potential relevance by identifying important host-pathogen protein interactions based on *in vitro* data. The *in vitro* Y2H results thus represent an important source for generating hypothesis on *B. mallei* virulence and pathogenicity.

**Y2H Detection of Host Interactions**—Table III shows the number of interactions between the tested *B. mallei* proteins and human and murine proteins as well as the corresponding number of conserved interactions in terms of orthologous pairs of interactions. The small overlap between murine and human prey proteins for which PPIs were identified is potentially an artifact from the Y2H experiments. A comprehensive detection of all possible host-prey and pathogen-bait protein interactions requires exhaustive repetition (84) and, hence, our detected PPIs represent a subset of all interactions that can occur. These results suggest that more comprehensive screens may be required to exhaustively delineate the total set of possible interactions for the two hosts.

The largest numbers of interactions were detected for proteins in *groups 1* and *2*, *i.e.* for proteins with multiple evidences, whereas the number of detected interactions decreased with a decreased number of evidences and an increased fraction of hypothetical proteins within a group. Thus, the proteins in *group 4*, derived from *ab initio* prediction methods and annotated as enzymes or hypothetical proteins, did not notably interact with host proteins in our Y2H screens. However, a lack of host interactions does not necessarily indicate a lack of a potential biological role, as proteins may function as modulators of metabolites or other host factors without involving detectable protein interactions.

**Known Virulence Factors and Their Virulence Phenotypes**—Among the known virulence factors shown in Table III, those linked to the T3SS in *group 1* interacted with at least one host protein. Some of these proteins, *e.g.* BipD, BipB, BopA, and BopE, have already been shown to be required for different aspects of pathogenicity in different animal models. For example, the BipD mutant of *B. pseudomallei* has been shown to attenuate infection after intraperitoneal challenge of BALB/c mice (67), whereas the BipB mutant of *B. pseudomallei* was not able to produce multinucleated giant cells during intranasal challenge of the same host and, consequently, improved host survival time (19). These two proteins, along with BipC, are positioned at the tip of the T3SS needle and form translocation pores (56). Furthermore, studies have shown that the effector (secreted protein) BopA is required for intracellular survival of both *B. pseudomallei* and *B. mallei* in phagocytic cells (56) and the *B. pseudomallei* effector BopE is required for optimal invasion of nonphagocytic cells. However, because BopE deletion mutants showed hardly any attenuation during intraperitoneal or intranasal infection of *B. pseudomallei* in BALB/c mice (56, 67), these results indicate

that multiple effectors, possibly in concert with BopE, contribute to a robust infection.

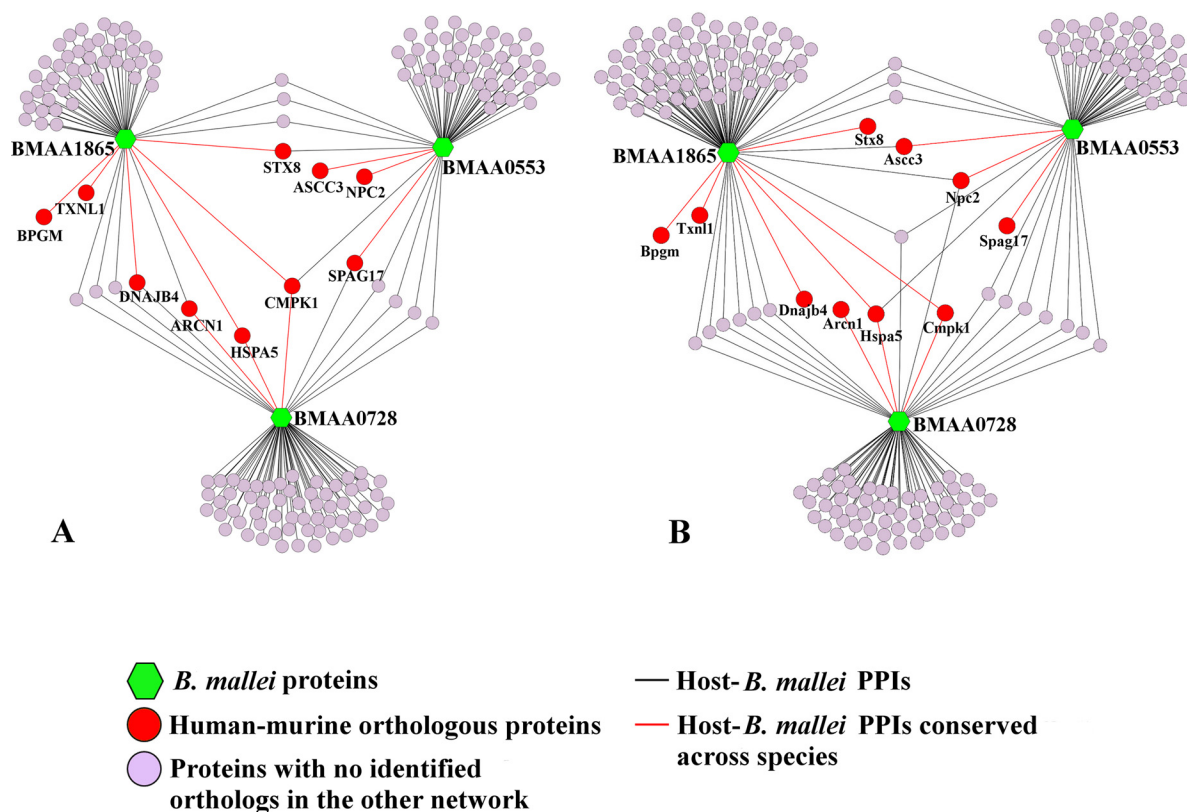
*Group 1* also contained a number of proteins from T6SS, the majority of which were annotated as proteins containing a VgrG domain. A number of large VgrGs (complex trimeric proteins containing transmembrane parts and a C-terminal domain) has been identified in Gram-negative bacteria and shown to be secreted and interact with host cells (85). The largest VgrG protein in *B. mallei*, BMAA0737, is an ortholog of the *B. pseudomallei* VgrG protein, BPSS1503, a protein identified as essential for virulence (11). Although we were unable to successfully clone BMAA0737, we identified a handful of PPIs for three of the five VgrG protein constructs we could assay. Two of them, BMAA0445 and BMAA0446, belonging to the second T6SS cluster, were recently identified in *B. pseudomallei* as important for interactions with other bacteria (86), whereas the third one, BMAA1269, was located outside of any known T6SS clusters (13). The small number of identified PPIs for VgrG proteins is likely a consequence of difficulties in expressing large proteins with transmembrane domains (87). Preliminary data also suggest that recombinant expression of VgrG proteins in *Escherichia coli* is challenging (K. Kwon, unpublished data).

In addition to T3SS and T6SS proteins, *group 1* also contained BimA, a well-characterized auto-transporter of *B. pseudomallei* and *B. mallei* that is exported through a type 5 secretion system (12). The role of BimA in actin-based motility is accepted, but the mechanisms are not fully understood, and it is likely that they differ between *B. pseudomallei* and *B. mallei* (12). In *B. mallei*, BimA was not required for pathogenicity in a hamster model of glanders (13), whereas in *B. pseudomallei* it was found to contribute to pathogenicity in murine models (12).

Thus, although many of the known virulence factors in *group 1* did include a large number of interactions, their specific biological function and resultant infectious phenotype varied dependent on both the animal model and the pathogen species used. Given that the three *B. mallei* proteins, BMAA0728, BMAA1865, and BMAA0553, all similarly attenuated virulence in our animal model and, hence, generated a common nonvirulent infectious phenotype, we now turn to our Y2H interaction data to infer and discuss the putative role of these novel virulence factors in invasion and survival within the host cell.

**Novel Virulence Factors and Their Host Protein Interactions**—Focusing on the three novel *B. mallei* virulence factors and their interacting human and murine proteins (Table III), these proteins participated in 192 and 236 PPIs with human and murine proteins, respectively ([supplementary Data set S3](#) and [supplementary Data set S4](#)), containing 170 unique human proteins and 191 unique murine proteins. Fig. 4 graphically shows these protein interactions for both hosts. There are 12 conserved (or overlapping) PPIs between these two data sets, containing 10 unique host proteins. To maximize





**FIG. 4. Host-pathogen interactions for three *B. mallei* virulence factors.** Three *B. mallei* virulence factors (green nodes) interact with 192 human (A) and 236 murine (B) proteins (purple and red nodes). There are 12 conserved protein-protein interactions (PPIs) between these two PPI data sets (red edges), *i.e.* in 12 PPIs, human proteins (red nodes in A) interacted with the same *B. mallei* proteins as their murine orthologs (red nodes in B). The host proteins (human/murine) in the conserved interactions include: activating signal cointegrator 1 complex subunit 3 (ASCC3/Ascc3), epididymal secretory protein E1 (NPC2/Npc2), sperm-associated antigen 17 (SPAG17/Spag17), ARCN1 protein-coatomer subunit- $\delta$  (ARCN1/Arcn1), UMP-cytidine monophosphate kinase (CMPK1/Cmpk1), heat shock 70-kDa protein 5 (HSPA5/Hspa5), bisphosphoglycerate mutase (BPGM/Bpgm), DnaJ homolog subfamily B, member 4 (DNAJB4/Dnajb4), syntaxin-8 (STX8/Stx8), and thioredoxin-like protein 1 (TXNL1/Txn1).

the coverage of interactions, we combined all human- and murine-*B. mallei* PPIs detected in our screens in subsequent analysis. This resulted in a data set consisting of 382 interactions among three *B. mallei* proteins and 336 human proteins (170 human proteins + 166 human-murine orthologs).

**Pathogen Targeting Ubiquitin-related Processes**—We found that all three proteins interacted with a statistically significant ( $p \leq 2 \times 10^{-3}$ ) number of host proteins involved in ubiquitin-related processes (supplemental Table S5). This finding is noteworthy, as intracellular pathogens are known to interfere with host ubiquitination processes, thus disturbing immune signaling cascades and attenuating the immune response (88–91). Additionally, pathogens interfere with the host ubiquitin proteasome system to prevent their degradation as well as to ensure their destruction when no longer required for establishing the infection (88–91).

One of the major known *B. mallei* virulence factors in the modulation of host ubiquitination system is TssM (BMAA0729), a protein that encodes a broadly acting deubiquitinase (17, 92). Although the exact function of TssM is unknown, it has been suggested that it acts to remove ubiq-

uitin from pathogen proteins, preventing their recognition and, consequently, the activation of early inflammatory responses (64). Using the PFAM-A database (43), we mapped the ubiquitin C-terminal hydrolase (UCH) domain to TssM. Interestingly, we found that the identified BMAA0728 protein (TssN) interacted with human and murine proteins that contain the UCH domain. This implies that it is possible that BMAA0728 also interacts with the UCH domain of its chromosome-adjacent protein TssM, a protein with which it is also coexpressed (D. DeShazer, unpublished data). Even though we tried to identify PPIs for TssM (Table III), we were unable to obtain a significant number of interactions for this protein, despite the large number of interactions detected for TssN. We further found that BMAA0728 interacts with the polyubiquitin-B protein (UBB) and with the cullin-1a protein (CUL1), a core component of multiple cullin-RING-based SCF E3 ubiquitin-protein ligase complexes. From the available human PPI data (42), we found that UBB and CUL1 physically interact with TNF receptor-associated factor 6 (TRAF-6) and I-kappa-B inhibitor alpha (IKB- $\alpha$ ), two host proteins that have been shown to be targets of TssM (17, 92). These observations lead



to the hypothesis that BMAA0728 either interferes with host ubiquitination processes directly or in collaboration with TssM. However, based on currently available domain annotations, no known domain could be assigned to BMAA0728.

**Pathogen Targeting Host Entry Mechanisms**—We identified Y2H interactions between BMAA0728 and GABA(A) receptor-associated protein-like 1 (GABARAPL1), a host protein possibly involved in the closure of phagosomal membranes (93, 94). This observation implies a possible role of BMAA0728 in *B. mallei* (auto)phagosomal escape. Besides GABARAPL1, we found many other murine and human host proteins that are involved in processes that can be related to cellular internalization or degradation escape and evasion and that interact with all three proteins (BMAA0728, BMAA1865, and BMAA0553; [supplemental Table S5](#)). However, based on the available data, we cannot determine a precise role for these three proteins in (auto)phagosomal escape/evasion processes.

To gain entry into host cells, multiple pathogen proteins target the host cytoskeleton through direct interactions with actin cytoskeleton regulators (95, 96). One such protein, Rho family GTPase (guanosine triphosphatase) CDC42, is a known pathogen target that induces actin rearrangement in the host cell (64, 96–98). Previous studies have suggested that *B. pseudomallei* BopE (*B. mallei* orthologous protein BopE) acts as a guanine nucleotide exchange factor for the Rho family of GTPases and that, through an interaction with a Rho GTPase CDC42, *B. pseudomallei* facilitates membrane ruffling and uptake into the host cell (64, 99). Studies have also implied that multiple pathogen proteins may act together with BopE to facilitate bacterial invasion (64, 99). Our study identified that BMAA1865 also interacts with RhoGTPases, such as CDC42 and RhoH (Ras homolog gene family, member H), as well as with a number of other proteins involved in processes related to actin cytoskeleton rearrangement ([supplemental Table S5](#)), and we hypothesize that BMAA1865 represents another *B. mallei* protein involved in host membrane ruffling and bacterial invasion.

**Pathogen Targeting Phosphorylation Signaling**—Although the *B. mallei* serine/threonine phosphatase BMAA0553 did not interact with any host protein already known to be involved in bacterial invasion and survival, we found it interacted with a large number of host proteins involved in various signaling processes and, hence, it is likely that it interferes with host intracellular signaling pathways ([supplemental Table S5](#)). We found that BMAA0553 interacts with Rho GDP (Rho family of guanosine diphosphate) dissociation inhibitor- $\beta$  (ARHGDI $\beta$ ), which regulates the GDP/GTP exchange reaction of Rho proteins, as well as with CDC42. Furthermore, as serine/threonine phosphorylation has been linked to actin cytoskeleton rearrangement processes (97, 100), it is possible that BMAA0553 also has a role in cytoskeletal rearrangement.

**Understanding Host-Pathogen Interactions**—Given the inherent nonspecific protein interactions and the lack of exhaustive experimental testing of protein interactions, our ap-

proach was to delineate the possible effects of select *B. mallei* protein by investigating which host processes they would have a high likelihood to interfere with. Indeed, the new virulence factors we identified appear to have a multifunctional biological role, perhaps because of nonspecific interactions.

The nonspecific nature of the interactions is both a source of weakness and strength in the ability of the pathogen to establish an infection. Avoiding highly specific and strong interactions avoids the risk of being limited to infecting only those hosts that have the requisite specific proteins. However, it requires that the pathogen maintain many, possibly alternative, indiscriminate interactions to achieve the same goal. The apparent multivariate capability of the virulence factors is also commensurate with the notion of using multiple, possibly overlapping, interactions to influence the host environment and to maximize the likelihood of establishing successful infections in multiple hosts and host environments.

### CONCLUSIONS

Our study is the first study to specifically and systematically generate host-protein interactions associated with known and putative virulence factors of *B. mallei*. We used bioinformatics approaches to identify 49 proteins involved in *B. mallei* pathogenicity. Using Y2H assays to screen these proteins against normalized whole human and murine proteome libraries, we identified a large number of host-pathogen PPIs. Analyzing these interactions allowed us to characterize host proteins targeted by *B. mallei* proteins and to identify three novel virulence factor candidates. Using a mouse aerosol challenge model, we demonstrated that these three proteins attenuated *B. mallei* ATCC 23344 virulence, implicating them and their host targets in *B. mallei* pathogenicity.

**Acknowledgments**—We thank Drs. P. Uetz and W. Nierman for helpful discussions.

\* This work was supported by the Defense Threat Reduction Agency (project CBS.MEDBIO.02.10.BH.021).

§ This article contains [supplemental Tables S1 to S5 and Data sets S1 to S4](#).

|| To whom correspondence should be addressed: Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD 21702. Tel.: (301) 619-8130; Fax: +1 301 619 1983; E-mail: jaques.reifman.civ@mail.mil.

Author contributions: N.Z., R.P., D.D.S., J.R., and A.W. designed research; V.M., N.Z., R.P., S.V.R., K.K., K.T., and D.D.S. performed research; V.M., N.Z., R.P., C.Y., X.Y., and D.D.S. analyzed data; and V.M., N.Z., R.P., D.D.S., J.R. and A.W. wrote the paper.

Conflict of interest: The authors declare no conflict of interest.

Disclaimer: The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the U.S. Army or the U.S. Department of Defense. This paper has been approved for public release with unlimited distribution.

## REFERENCES

- Estes, D. M., Dow, S. W., Schweizer, H. P., and Torres, A. G. (2010) Present and future therapeutic strategies for melioidosis and glanders. *Expert Rev. Anti-Infective Ther.* **8**, 325–338
- Sarkar-Tyson, M., and Titball, R. W. (2010) Progress toward development of vaccines against melioidosis: A review. *Clin. Therap.* **32**, 1437–1445
- Godoy, D., Randle, G., Simpson, A. J., Aanensen, D. M., Pitt, T. L., Kinoshita, R., and Spratt, B. G. (2003) Multilocus sequence typing and evolutionary relationships among the causative agents of melioidosis and glanders, *Burkholderia pseudomallei* and *Burkholderia mallei*. *J. Clin. Microbiol.* **41**, 2068–2079
- Galyov, E. E., Brett, P. J., and DeShazer, D. (2010) Molecular insights into *Burkholderia pseudomallei* and *Burkholderia mallei* pathogenesis. *Ann. Rev. Microbiol.* **64**, 495–517
- Ulrich, R. L., Deshazer, D., Brueggemann, E. E., Hines, H. B., Oyston, P. C., and Jeddelloh, J. A. (2004) Role of quorum sensing in the pathogenicity of *Burkholderia pseudomallei*. *J. Med. Microbiol.* **53**, 1053–1064
- Ulrich, R. L., Deshazer, D., Hines, H. B., and Jeddelloh, J. A. (2004) Quorum sensing: A transcriptional regulatory system involved in the pathogenicity of *Burkholderia mallei*. *Infection Immunity* **72**, 6589–6596
- Fernandes, P. J., Guo, Q., Waag, D. M., and Donnenberg, M. S. (2007) The type IV pilin of *Burkholderia mallei* is highly immunogenic but fails to protect against lethal aerosol challenge in a murine model. *Infection Immunity* **75**, 3027–3032
- DeShazer, D., Waag, D. M., Fritz, D. L., and Woods, D. E. (2001) Identification of a *Burkholderia mallei* polysaccharide gene cluster by subtractive hybridization and demonstration that the encoded capsule is an essential virulence determinant. *Microb. Pathogen.* **30**, 253–269
- Burtneck, M. N., DeShazer, D., Nair, V., Gherardini, F. C., and Brett, P. J. (2010) *Burkholderia mallei* cluster 1 type VI secretion mutants exhibit growth and actin polymerization defects in RAW 264.7 murine macrophages. *Infection Immunity* **78**, 88–99
- Stevens, J. M., Ulrich, R. L., Taylor, L. A., Wood, M. W., Deshazer, D., Stevens, M. P., and Galyov, E. E. (2005) Actin-binding proteins from *Burkholderia mallei* and *Burkholderia thailandensis* can functionally compensate for the actin-based motility defect of a *Burkholderia pseudomallei* *bimA* mutant. *J. Bacteriol.* **187**, 7857–7862
- Burtneck, M. N., Brett, P. J., Harding, S. V., Ngugi, S. A., Ribot, W. J., Chantratita, N., Scorpio, A., Milne, T. S., Dean, R. E., Fritz, D. L., Peacock, S. J., Prior, J. L., Atkins, T. P., and Deshazer, D. (2011) The cluster 1 type VI secretion system is a major virulence determinant in *Burkholderia pseudomallei*. *Infection Immunity* **79**, 1512–1525
- Lazar Adler, N. R., Stevens, J. M., Stevens, M. P., and Galyov, E. E. (2011) Autotransporters and their role in the virulence of *Burkholderia pseudomallei* and *Burkholderia mallei*. *Frontiers Microbiol.* **2**, 151
- Schell, M. A., Ulrich, R. L., Ribot, W. J., Brueggemann, E. E., Hines, H. B., Chen, D., Lipscomb, L., Kim, H. S., Mrazek, J., Nierman, W. C., and Deshazer, D. (2007) Type VI secretion is a major virulence determinant in *Burkholderia mallei*. *Mol. Microbiol.* **64**, 1466–1485
- Ulrich, R. L., and DeShazer, D. (2004) Type III secretion: A virulence factor delivery system essential for the pathogenicity of *Burkholderia mallei*. *Infection Immunity* **72**, 1150–1154
- Chen, Z. J. (2005) Ubiquitin signalling in the NF-kappaB pathway. *Nat. Cell Biol.* **7**, 758–765
- Sitthidet, C., Stevens, J. M., Chantratita, N., Currie, B. J., Peacock, S. J., Korbsrisate, S., and Stevens, M. P. (2008) Prevalence and sequence diversity of a factor required for actin-based motility in natural populations of *Burkholderia* species. *J. Clin. Microbiol.* **46**, 2418–2422
- Shanks, J., Burtneck, M. N., Brett, P. J., Waag, D. M., Spurgers, K. B., Ribot, W. J., Schell, M. A., Panchal, R. G., Gherardini, F. C., Wilkinson, K. D., and Deshazer, D. (2009) *Burkholderia mallei* *tssM* encodes a putative deubiquitinase that is secreted and expressed inside infected RAW 264.7 murine macrophages. *Infection Immunity* **77**, 1636–1648
- Whitlock, G. C., Valbuena, G. A., Popov, V. L., Judy, B. M., Estes, D. M., and Torres, A. G. (2009) *Burkholderia mallei* cellular interactions in a respiratory cell model. *J. Med. Microbiol.* **58**, 554–562
- Suparak, S., Kespichayawattana, W., Haque, A., Easton, A., Damnin, S., Lertmengkolchai, G., Bancroft, G. J., and Korbsrisate, S. (2005) Multinucleated giant cell formation and apoptosis in infected host cells is mediated by *Burkholderia pseudomallei* type III secretion protein BipB. *J. Bacteriol.* **187**, 6556–6560
- Oh, Y. M., Kim, M., Lee-Cruz, L., Lai-Hoe, A., Go, R., AINUDDIN, N., Rahim, R. A., Shukor, N., and Adams, J. M. (2012) Distinctive bacterial communities in the rhizosphere of four tropical tree species. *Microbial Ecol.* **64**, 1018–1027
- Suwannasaen, D., Mahawantung, J., Chaowagul, W., Limmathurotsakul, D., Felgner, P. L., Davies, H., Bancroft, G. J., Titball, R. W., and Lertmengkolchai, G. (2011) Human immune responses to *Burkholderia pseudomallei* characterized by protein microarray analysis. *J. Infect. Dis.* **203**, 1002–1011
- Tuanyok, A., Tom, M., Dunbar, J., and Woods, D. E. (2006) Genome-wide expression analysis of *Burkholderia pseudomallei* infection in a hamster model of acute melioidosis. *Infection Immunity* **74**, 5465–5476
- Ulett, G. C., Labrooy, J. T., Currie, B. J., Barnes, J. L., and Ketheesan, N. (2005) A model of immunity to *Burkholderia pseudomallei*: Unique responses following immunization and acute lethal infection. *Microbes Infect.* **7**, 1263–1275
- Yu, C., Desai, V., Cheng, L., and Reifman, J. (2012) QuartetS-DB: a large-scale orthology database for prokaryotes and eukaryotes inferred by evolutionary evidence. *BMC Bioinformatics* **13**, 143
- Yu, C., Zavaljevski, N., Desai, V., and Reifman, J. (2011) QuartetS: A fast and accurate algorithm for large-scale orthology detection. *Nucleic Acids Res.* **39**, e88
- Benson, D. A., Karsch-Mizrachi, I., Clark, K., Lipman, D. J., Ostell, J., and Sayers, E. W. (2012) *GenBank*. *Nucleic Acids Res.* **40**, D48–D53
- Brinkac, L. M., Davidsen, T., Beck, E., Ganapathy, A., Caler, E., Dodson, R. J., Durkin, A. S., Harkins, D. M., Lorenzi, H., Madupu, R., Sebastian, Y., Shrivastava, S., Thiagarajan, M., Orvis, J., Sundaram, J. P., Crabtree, J., Galens, K., Zhao, Y., Inman, J. M., Montgomery, R., Schobel, S., Galinsky, K., Tanenbaum, D. M., Resnick, A., Zafar, N., White, O., and Sutton, G. (2010) Pathema: A clade-specific bioinformatics resource center for pathogen research. *Nucleic Acids Res.* **38**, D408–D414
- Yu, C., Zavaljevski, N., Desai, V., Johnson, S., Stevens, F. J., and Reifman, J. (2008) The development of PIPA: An integrated and automated pipeline for genome-wide protein function annotation. *BMC Bioinformatics* **9**, 52
- Arnold, R., Brandmaier, S., Kleine, F., Tischler, P., Heinz, E., Behrens, S., Niinikoski, A., Mewes, H. W., Horn, M., and Rattei, T. (2009) Sequence-based prediction of type III secreted proteins. *PLOS Pathogens* **5**, e1000376
- Lower, M., and Schneider, G. (2009) Prediction of type III secretion signals in genomes of gram-negative bacteria. *PLOS ONE* **4**, e5917
- McDermott, J. E., Corrigan, A., Peterson, E., Oehmen, C., Niemann, G., Cambronne, E. D., Sharp, D., Adkins, J. N., Samudrala, R., and Heffron, F. (2011) Computational prediction of type III and IV secreted effectors in gram-negative bacteria. *Inf. Immunity* **79**, 23–32
- Samudrala, R., Heffron, F., and McDermott, J. E. (2009) Accurate prediction of secreted substrates and identification of a conserved putative secretion signal for type III secretion systems. *PLOS Pathogens* **5**, e1000375
- Stellberger, T., Hauser, R., Baiker, A., Pothineni, V. R., Haas, J., and Uetz, P. (2010) Improving the yeast two-hybrid system with permuted fusions proteins: The *Varicella zoster* virus interactome. *Proteome Sci.* **8**, 8
- Walhout, A. J., Temple, G. F., Brasch, M. A., Hartley, J. L., LORSON, M. A., van den Heuvel, S., and Vidal, M. (2000) GATEWAY recombinational cloning: Application to the cloning of large numbers of open reading frames or ORFs. *Methods Enzymol.* **328**, 575–592
- James, P., Halladay, J., and Craig, E. A. (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**, 1425–1436
- Rajagopala, S. V., and Uetz, P. (2011) Analysis of protein-protein interactions using high-throughput yeast two-hybrid screens. *Methods Mol. Biol.* **781**, 1–29
- Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005) BioMart and Bioconductor: A powerful link between biological databases and microarray data analysis. *Bioinformatics* **21**, 3439–3440
- Durinck, S., Spellman, P. T., Birney, E., and Huber, W. (2009) Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protocols* **4**, 1184–1191
- Lopes, C. T., Franz, M., Kazi, F., Donaldson, S. L., Morris, Q., and Bader,

- G. D. (2010) Cytoscape Web: An interactive web-based network browser. *Bioinformatics* **26**, 2347–2348
40. Sayers, E. W., Barrett, T., Benson, D. A., Bolton, E., Bryant, S. H., Canese, K., Chetvermin, V., Church, D. M., DiCuccio, M., Federhen, S., Feolo, M., Fingerman, I. M., Geer, L. Y., Helmberg, W., Kapustin, Y., Krasnov, S., Landsman, D., Lipman, D. J., Lu, Z., Madden, T. L., Madej, T., Maglott, D. R., Marchler-Bauer, A., Miller, V., Karsch-Mizrachi, I., Ostell, J., Panchenko, A., Phan, L., Pruitt, K. D., Schuler, G. D., Sequeira, E., Sherry, S. T., Shumway, M., Sirotkin, K., Slotta, D., Souvorov, A., Starchenko, G., Tatusova, T. A., Wagner, L., Wang, Y., Wilbur, W. J., Yaschenko, E., and Ye, J. (2012) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **40**, D13–D25
  41. Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Church, D. M., DiCuccio, M., Edgar, R., Federhen, S., Helmberg, W., Kenton, D. L., Khovayko, O., Lipman, D. J., Madden, T. L., Maglott, D. R., Ostell, J., Pontius, J. U., Pruitt, K. D., Schuler, G. D., Schriml, L. M., Sequeira, E., Sherry, S. T., Sirotkin, K., Starchenko, G., Suzek, T. O., Tatusov, R., Tatusova, T. A., Wagner, L., and Yaschenko, E. (2005) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **33**, D39–D45
  42. Yu, X., Wallqvist, A., and Reifman, J. (2012) Inferring high-confidence human protein-protein interactions. *BMC Bioinformatics* **13**, 79
  43. Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E. L., Eddy, S. R., Bateman, A., and Finn, R. D. (2012) The Pfam protein families database. *Nucleic Acids Res.* **40**, D290–D301
  44. Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Statist. Soc.* **57**, 289–300
  45. Dutta, B., Wallqvist, A., and Reifman, J. (2012) PathNet: A tool for pathway analysis using topological information. *Source Code Biology Med.* **7**, 10
  46. Lopez, C. M., Rholl, D. A., Trunck, L. A., and Schweizer, H. P. (2009) Versatile dual-technology system for markerless allele replacement in *Burkholderia pseudomallei*. *Appl. Environmental Microbiol.* **75**, 6496–6503
  47. Simon, R., Priefer, U., and Puhler, A. (1983) A broad host range mobilization system for in vivo genetic engineering: Transposon mutagenesis in gram negative bacteria. *Nat. Biotechnol.* **1**, 784–791
  48. Niernan, W. C., DeShazer, D., Kim, H. S., Tettelin, H., Nelson, K. E., Feldblyum, T., Ulrich, R. L., Ronning, C. M., Brinkac, L. M., Daugherty, S. C., Davidsen, T. D., Deboy, R. T., Dimitrov, G., Dodson, R. J., Durkin, A. S., Gwinn, M. L., Haft, D. H., Khouri, H., Kolonay, J. F., Madupu, R., Mohammud, Y., Nelson, W. C., Radune, D., Romero, C. M., Sarria, S., Selengut, J., Shamblin, C., Sullivan, S. A., White, O., Yu, Y., Zafar, N., Zhou, L., and Fraser, C. M. (2004) Structural flexibility in the *Burkholderia mallei* genome. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14246–14251
  49. Wilson, K. (2001) Preparation of genomic DNA from bacteria. Current Protocols in Molecular Biology Chapter 2, Unit 2.4
  50. Roy, C. J., Hale, M., Hartings, J. M., Pitt, L., and Duniho, S. (2003) Impact of inhalation exposure modality and particle size on the respiratory deposition of ricin in BALB/c mice. *Inhalation Toxicol.* **15**, 619–638
  51. Bewick, V., Cheek, L., and Ball, J. (2004) Statistics review 12: Survival analysis. *Crit. Care* **8**, 389–394
  52. Losada, L., Ronning, C. M., DeShazer, D., Woods, D., Fedorova, N., Kim, H. S., Shabalina, S. A., Pearson, T. R., Brinkac, L., Tan, P., Nandi, T., Crabtree, J., Badger, J., Beckstrom-Sternberg, S., Saqib, M., Schutzer, S. E., Keim, P., and Niernan, W. C. (2010) Continuing evolution of *Burkholderia mallei* through genome reduction and large-scale rearrangements. *Genome Biol. Evol.* **2**, 102–116
  53. Gillespie, J. J., Wattam, A. R., Cammer, S. A., Gabbard, J. L., Shukla, M. P., Dalay, O., Driscoll, T., Hix, D., Mane, S. P., Mao, C., Nordberg, E. K., Scott, M., Schulman, J. R., Snyder, E. E., Sullivan, D. E., Wang, C., Warren, A., Williams, K. P., Xue, T., Yoo, H. S., Zhang, C., Zhang, Y., Will, R., Kenyon, R. W., and Sobral, B. W. (2011) PATRIC: The comprehensive bacterial bioinformatics resource with a focus on human pathogenic species. *Infection Immunity* **79**, 4286–4298
  54. Pruitt, K. D., Tatusova, T., Brown, G. R., and Maglott, D. R. (2012) NCBI Reference Sequences (RefSeq): Current status, new features and genome annotation policy. *Nucleic Acids Res.* **40**, D130–D135
  55. Tseng, T. T., Tyler, B. M., and Setubal, J. C. (2009) Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC Microbiol.* **9**, S2
  56. Sun, G. W., and Gan, Y. H. (2010) Unraveling type III secretion systems in the highly versatile *Burkholderia pseudomallei*. *Trends Microbiol.* **18**, 561–568
  57. Drevinek, P., Holden, M. T., Ge, Z., Jones, A. M., Ketchell, I., Gill, R. T., and Mahenthalingam, E. (2008) Gene expression changes linked to antimicrobial resistance, oxidative stress, iron depletion and retained motility are observed when *Burkholderia cenocepacia* grows in cystic fibrosis sputum. *BMC Infectious Dis.* **8**, 121
  58. Haraga, A., West, T. E., Brittnacher, M. J., Skerrett, S. J., and Miller, S. I. (2008) *Burkholderia thailandensis* as a model system for the study of the virulence-associated type III secretion system of *Burkholderia pseudomallei*. *Infection Immunity* **76**, 5402–5411
  59. Shalom, G., Shaw, J. G., and Thomas, M. S. (2007) *In vivo* expression technology identifies a type VI secretion system locus in *Burkholderia pseudomallei* that is induced upon invasion of macrophages. *Microbiology* **153**, 2689–2699
  60. DeShazer, D., Brett, P. J., Burtnick, M. N., and Woods, D. E. (1999) Molecular characterization of genetic loci required for secretion of exoproducts in *Burkholderia pseudomallei*. *J. Bacteriol.* **181**, 4661–4664
  61. Whitlock, G. C., Estes, D. M., and Torres, A. G. (2007) Glanders: Off to the races with *Burkholderia mallei*. *FEMS Microbiol. Lett.* **277**, 115–122
  62. Jeddeloh, J. A., Fritz, D. L., Waag, D. M., Hartings, J. M., and Andrews, G. P. (2003) Biodefense-driven murine model of pneumonic melioidosis. *Infection Immunity* **71**, 584–587
  63. Leakey, A. K., Ulett, G. C., and Hirst, R. G. (1998) BALB/c and C57Bl/6 mice infected with virulent *Burkholderia pseudomallei* provide contrasting animal models for the acute and chronic forms of human melioidosis. *Microbial Pathogenesis* **24**, 269–275
  64. Allwood, E. M., Devenish, R. J., Prescott, M., Adler, B., and Boyce, J. D. (2011) Strategies for intracellular survival of *Burkholderia pseudomallei*. *Frontiers Microbiol.* **2**, 170
  65. Essex-Lopresti, A. E., Boddey, J. A., Thomas, R., Smith, M. P., Hartley, M. G., Atkins, T., Brown, N. F., Tsang, C. H., Peak, I. R., Hill, J., Beacham, I. R., and Titball, R. W. (2005) A type IV pilin, PilA, contributes to adherence of *Burkholderia pseudomallei* and virulence *in vivo*. *Infection Immunity* **73**, 1260–1264
  66. Gong, L., Cullinane, M., Treerat, P., Ramm, G., Prescott, M., Adler, B., Boyce, J. D., and Devenish, R. J. (2011) The *Burkholderia pseudomallei* type III secretion system and BopA are required for evasion of LC3-associated phagocytosis. *PLoS ONE* **6**, e17852
  67. Stevens, M. P., and Galyov, E. E. (2004) Exploitation of host cells by *Burkholderia pseudomallei*. *Int. J. Med. Microbiol.* **293**, 549–555
  68. Stevens, M. P., Stevens, J. M., Jeng, R. L., Taylor, L. A., Wood, M. W., Hawes, P., Monaghan, P., Welch, M. D., and Galyov, E. E. (2005) Identification of a bacterial factor required for actin-based motility of *Burkholderia pseudomallei*. *Mol. Microbiol.* **56**, 40–53
  69. Langridge, G. C., Phan, M. D., Turner, D. J., Perkins, T. T., Parts, L., Haase, J., Charles, I., Maskell, D. J., Peters, S. E., Dougan, G., Wain, J., Parkhill, J., and Turner, A. K. (2009) Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants. *Genome Res.* **19**, 2308–2316
  70. Schmidt, F., and Volker, U. (2011) Proteome analysis of host-pathogen interactions: Investigation of pathogen responses to the host cell environment. *Proteomics* **11**, 3203–3211
  71. Schoolnik, G. K. (2002) Microarray analysis of bacterial pathogenicity. *Adv. Microbial Physiol.* **46**, 1–45
  72. Sobek, J., Bartscherer, K., Jacob, A., Hoheisel, J. D., and Angenendt, P. (2006) Microarray technology as a universal tool for high-throughput analysis of biological systems. *Combinatorial Chem. High Throughput Screening* **9**, 365–380
  73. Brass, A. L., Huang, I. C., Benita, Y., John, S. P., Krishnan, M. N., Feeley, E. M., Ryan, B. J., Weyer, J. L., van der Weyden, L., Fikrig, E., Adams, D. J., Xavier, R. J., Farzan, M., and Elledge, S. J. (2009) The IFITM proteins mediate cellular resistance to influenza A H1N1 virus, West Nile virus, and dengue virus. *Cell* **139**, 1243–1254
  74. Calderwood, M. A., Venkatesan, K., Xing, L., Chase, M. R., Vazquez, A., Holthaus, A. M., Ewence, A. E., Li, N., Hirozane-Kishikawa, T., Hill, D. E., Vidal, M., Kieff, E., and Johannsen, E. (2007) Epstein-Barr virus and virus human protein interaction maps. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7606–7611



75. Karlas, A., Machuy, N., Shin, Y., Pleissner, K. P., Artarini, A., Heuer, D., Becker, D., Khalil, H., Ogilvie, L. A., Hess, S., Maurer, A. P., Muller, E., Wolff, T., Rudel, T., and Meyer, T. F. (2010) Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. *Nature* **463**, 818–822
76. Konig, R., Stertz, S., Zhou, Y., Inoue, A., Hoffmann, H. H., Bhattacharyya, S., Alamares, J. G., Tscherne, D. M., Ortigoza, M. B., Liang, Y., Gao, Q., Andrews, S. E., Bandyopadhyay, S., De Jesus, P., Tu, B. P., Pache, L., Shih, C., Orth, A., Bonamy, G., Miraglia, L., Ideker, T., Garcia-Sastre, A., Young, J. A., Palese, P., Shaw, M. L., and Chanda, S. K. (2010) Human host factors required for influenza virus replication. *Nature* **463**, 813–817
77. Li, R., Zhu, J., Xie, Z., Liao, G., Liu, J., Chen, M. R., Hu, S., Woodard, C., Lin, J., Taverna, S. D., Desai, P., Ambinder, R. F., Hayward, G. S., Qian, J., Zhu, H., and Hayward, S. D. (2011) Conserved herpesvirus kinases target the DNA damage response pathway and TIP60 histone acetyltransferase to promote virus replication. *Cell Host Microbe* **10**, 390–400
78. Shapira, S. D., Gat-Viks, I., Shum, B. O., Dricot, A., de Grace, M. M., Wu, L., Gupta, P. B., Hao, T., Silver, S. J., Root, D. E., Hill, D. E., Regev, A., and Hacohen, N. (2009) A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell* **139**, 1255–1267
79. Chen, Z., Li, B., Zhang, J., Qin, L., Zhou, D., Han, Y., Du, Z., Guo, Z., Song, Y., and Yang, R. (2006) Quorum sensing affects virulence-associated proteins F1, LcrV, KatY and pH6 etc. of *Yersinia pestis* as revealed by protein microarray-based antibody profiling. *Microbes Infection* **8**, 2501–2508
80. Keasey, S. L., Schmid, K. E., Lee, M. S., Meegan, J., Tomas, P., Minto, M., Tikhonov, A. P., Schweitzer, B., and Ulrich, R. G. (2009) Extensive antibody cross-reactivity among infectious gram-negative bacteria revealed by proteome microarray analysis. *Mol. Cell. Proteomics* **8**, 924–935
81. Liang, L., Juarez, S., Nga, T. V., Dunstan, S., Nakajima-Sasaki, R., Davies, D. H., McSorley, S., Baker, S., and Felgner, P. L. (2013) Immune profiling with a *Salmonella Typhi* antigen microarray identifies new diagnostic biomarkers of human typhoid. *Sci. Reports* **3**, 1043
82. Yang, H., Ke, Y., Wang, J., Tan, Y., Myeni, S. K., Li, D., Shi, Q., Yan, Y., Chen, H., Guo, Z., Yuan, Y., Yang, X., Yang, R., and Du, Z. (2011) Insight into bacterial virulence mechanisms against host immune response via the *Yersinia pestis*-human protein-protein interaction network. *Infection Immunity* **79**, 4413–4424
83. Yu, X., Ivanic, J., Memišević, V., Wallqvist, A., and Reifman, J. (2011) Categorizing biases in high-confidence high-throughput protein-protein interaction data sets. *Mol. Cell. Proteomics* **10**, M111 012500
84. Yu, H., Braun, P., Yildirim, M. A., Lemmens, I., Venkatesan, K., Sahalie, J., Hirozane-Kishikawa, T., Gebreab, F., Li, N., Simonis, N., Hao, T., Rual, J. F., Dricot, A., Vazquez, A., Murray, R. R., Simon, C., Tardivo, L., Tam, S., Svrikapa, N., Fan, C., de Smet, A. S., Motyl, A., Hudson, M. E., Park, J., Xin, X., Cusick, M. E., Moore, T., Boone, C., Snyder, M., Roth, F. P., Barabasi, A. L., Tavernier, J., Hill, D. E., and Vidal, M. (2008) High-quality binary protein interaction map of the yeast interactome network. *Science* **322**, 104–110
85. Pukatzki, S., McAuley, S. B., and Miyata, S. T. (2009) The type VI secretion system: Translocation of effectors and effector-domains. *Curr. Opin. Microbiol.* **12**, 11–17
86. Schwarz, S., West, T. E., Boyer, F., Chiang, W. C., Carl, M. A., Hood, R. D., Rohmer, L., Tolker-Nielsen, T., Skerrett, S. J., and Mougous, J. D. (2010) *Burkholderia* type VI secretion systems have distinct roles in eukaryotic and bacterial cell interactions. *PLOS Pathogens* **6**, e1001068
87. Fetchko, M., Auerbach, D., and Stagljar, I. (2003) Yeast genetic methods for the detection of membrane protein interactions: Potential use in drug discovery. *BioDrugs: Clin. Immunotherapeut., Biopharmaceut. Gene Therapy* **17**, 413–424
88. Anderson, D. M., and Frank, D. W. (2012) Five mechanisms of manipulation by bacterial effectors: A ubiquitous theme. *PLOS Pathogens* **8**, e1002823
89. Angot, A., Vergunst, A., Genin, S., and Peeters, N. (2007) Exploitation of eukaryotic ubiquitin signaling pathways by effectors translocated by bacterial type III and type IV secretion systems. *PLOS Pathogens* **3**, e3
90. Munro, P., Flatau, G., and Lemichez, E. (2007) Bacteria and the ubiquitin pathway. *Curr. Opin. Microbiol.* **10**, 39–46
91. Perrett, C. A., Lin, D. Y., and Zhou, D. (2011) Interactions of bacterial proteins with host eukaryotic ubiquitin pathways. *Frontiers Microbiol.* **2**, 143
92. Tan, K. S., Chen, Y., Lim, Y. C., Tan, G. Y., Liu, Y., Lim, Y. T., Macary, P., and Gan, Y. H. (2010) Suppression of host innate immune response by *Burkholderia pseudomallei* through the virulence factor TssM. *J. Immunol.* **184**, 5160–5171
93. Mizushima, N., Yoshimori, T., and Ohsumi, Y. (2011) The role of Atg proteins in autophagosome formation. *Ann. Rev. Cell Development. Biol.* **27**, 107–132
94. Noda, T., and Yoshimori, T. (2009) Molecular basis of canonical and bactericidal autophagy. *Int. Immunol.* **21**, 1199–1204
95. Haglund, C. M., and Welch, M. D. (2011) Pathogens and polymers: Microbe-host interactions illuminate the cytoskeleton. *J. Cell Biol.* **195**, 7–17
96. Rottner, K., Stradal, T. E., and Wehland, J. (2005) Bacteria-host-cell interactions at the plasma membrane: Stories on actin cytoskeleton subversion. *Development. Cell* **9**, 3–17
97. Bhavsar, A. P., Guttman, J. A., and Finlay, B. B. (2007) Manipulation of host-cell pathways by bacterial pathogens. *Nature* **449**, 827–834
98. Ham, H., Sreelatha, A., and Orth, K. (2011) Manipulation of host membranes by bacterial effectors. *Nat. Rev. Microbiol.* **9**, 635–646
99. Stevens, M. P., Friebel, A., Taylor, L. A., Wood, M. W., Brown, P. J., Hardt, W. D., and Galyov, E. E. (2003) A *Burkholderia pseudomallei* type III secreted protein, BopE, facilitates bacterial invasion of epithelial cells and exhibits guanine nucleotide exchange factor activity. *J. Bacteriol.* **185**, 4992–4996
100. Zaidel-Bar, R., Itzkovitz, S., Ma'ayan, A., Iyengar, R., and Geiger, B. (2007) Functional atlas of the integrin adhesome. *Nat. Cell Biol.* **9**, 858–867