Supplemental information

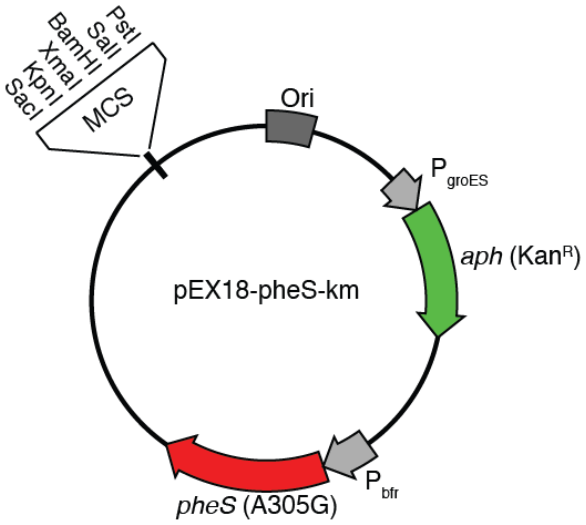

Figure S1. An Improved Allelic Exchange Vector for *F. novicida* (Related to Figure 1).
The parental plasmid (Charity et al., 2007) was modified to encode a cassette containing the *F. novicida* bacterioferitin (FTN_1410) promoter (P_{bfr}) driving expression of a PheS point mutant (A305G) with relaxed substrate specificity (Kast, 1994). Unique restriction sites in the multiple cloning site (MCS) used for inserting deletion alleles are indicated.

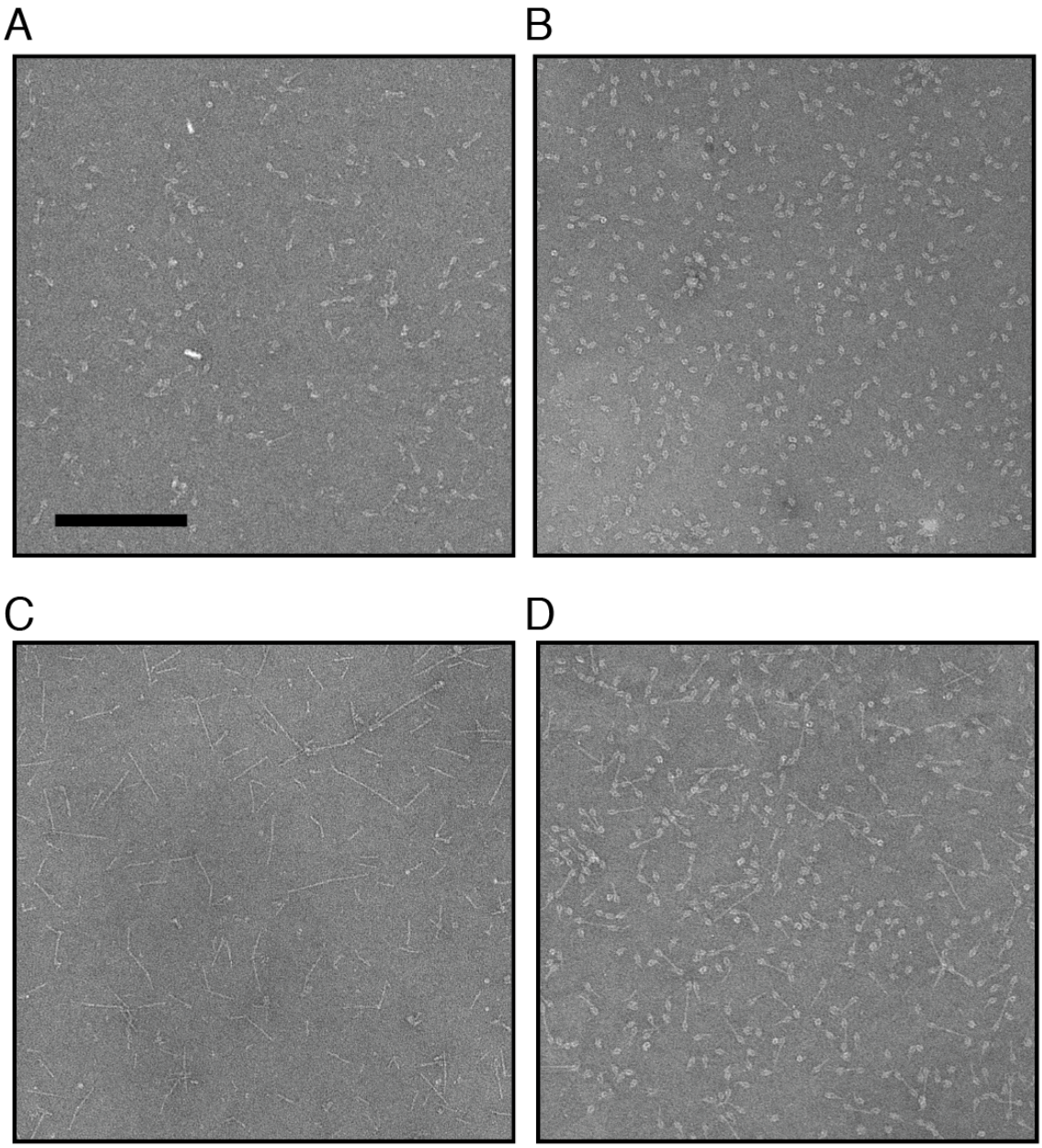

Figure S2. Representative Negative Stain Electron Micrographs (Related to Figure 4).
(A-D) Micrographs of VgrG immunoprecipitated from *F. novicida* (A), heterologously expressed PdpA (B), VgrG (C), or PdpA co-expressed with VgrG (D).

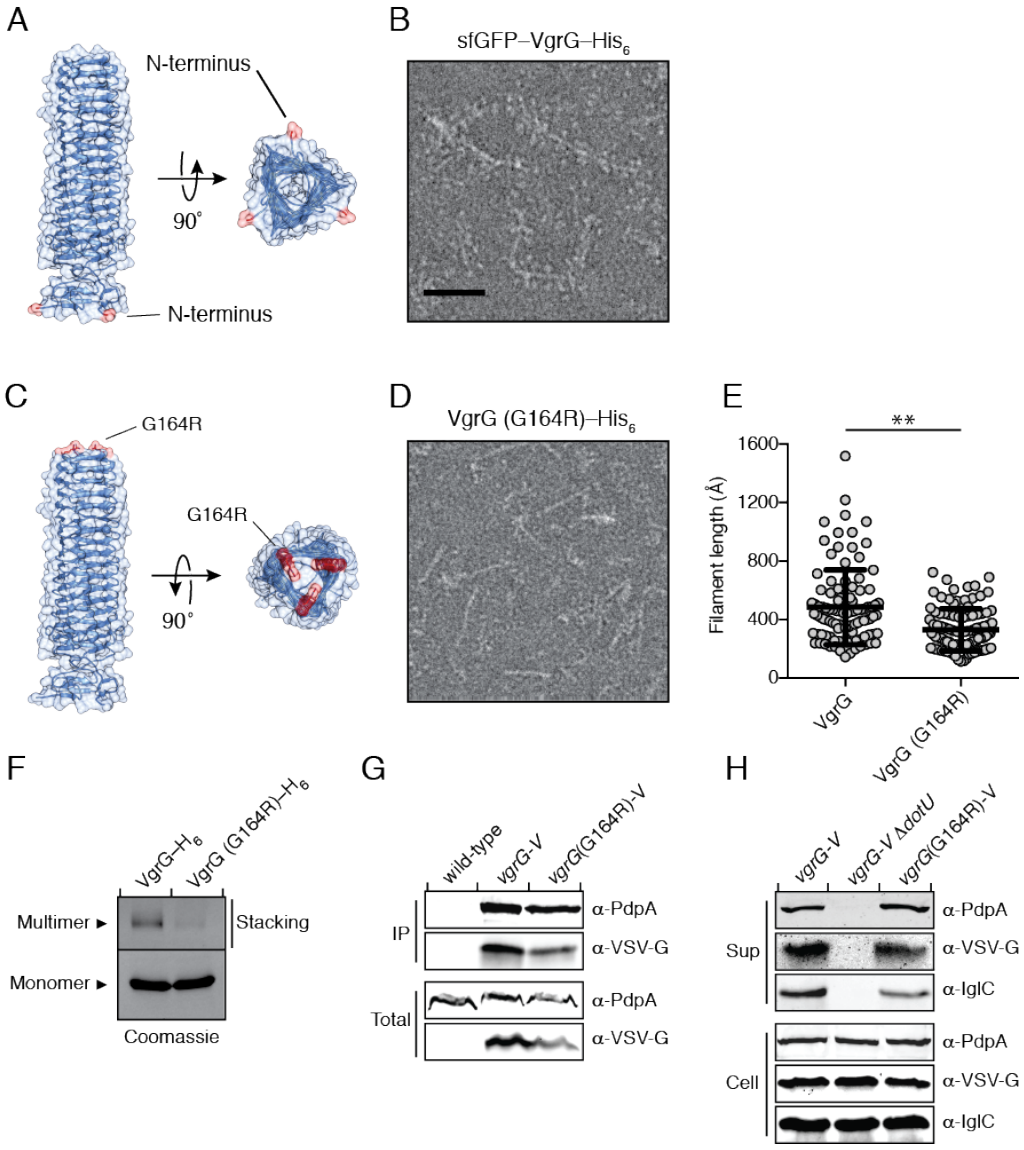

Figure S3. VgrG Polymerization is Inhibited by Mutation of the C-terminus (Related to Figure 4).
(A) Structural model of a VgrG trimer with the N-terminus highlighted in red.
(B) Negative stain electron micrograph of heterologously expressed sfGFP–VgrG. sfGFP is apparent as negatively stained punctae along the length of VgrG filaments. Scale bar represents 50 nm.
(C) Structural model of a VgrG trimer with the G164R mutation highlighted in red.
(D) Negative stain electron micrograph of VgrG (G164R) expressed and purified from *E. coli.*
(E) Filament lengths of wild-type VgrG (G164R) expressed and purified from *E. coli* (n=100/group). Horizontal lines represent mean and standard deviation for each group. Asterisk denotes a statistical difference (**$p \leq 10^{-3}$).
(F) Coomassie-stained SDS-PAGE analysis of heterologously expressed wild-type VgrG and VgrG (G164R). The presence of the SDS-resistant VgrG mulitmer is noted in the stacking portion of the gel.
(G) Western blot analysis of supernatant (Sup) and cell fractions from the indicated *F. novicida* strains.
(H) Western blot analysis of samples derived from the indicated *F. novicida* strains before (total) or following (IP) anti-VSV-G immunoprecipitation.

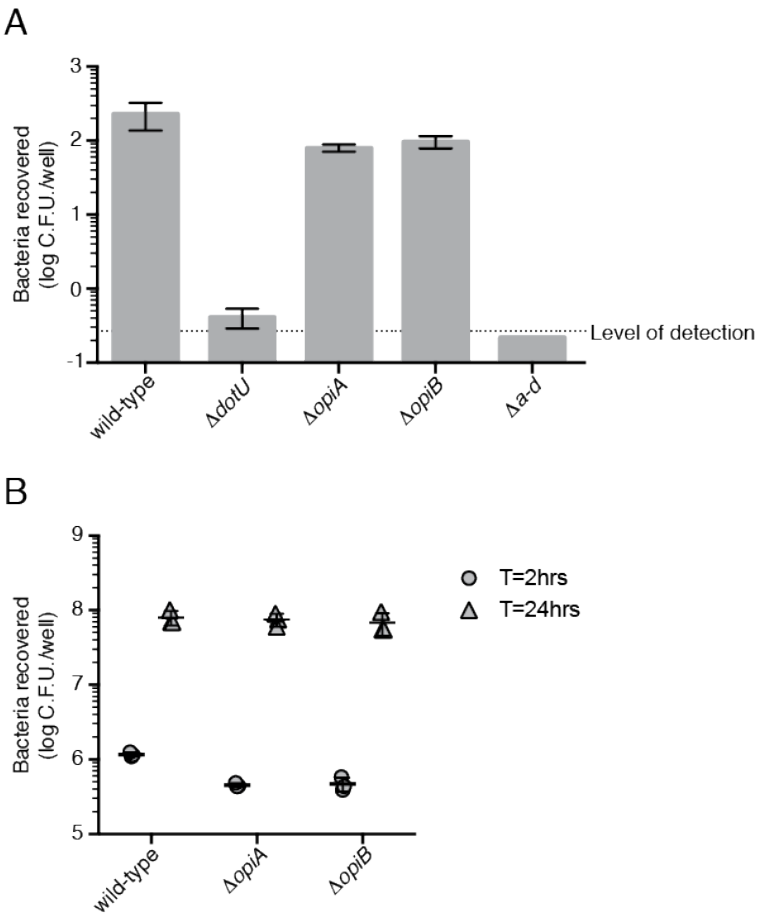

Figure S4. *opiA* and *opiB* do not contribute to intracellular growth (Related to Figure 5).
(A) Growth of the indicated *F. novicida* strains in A549 human alveolar epithelial-like cells 24 hours post infection. Data presented is normalized to the number of intracellular bacteria immediately post infection. (B) Growth of the indicated *F. tularensis* subsp. holarctica LVS strains in PMA-differentiated THP-1 cells. The 2 hour timepoint represents the number of gentamicin protected bacteria post infection.

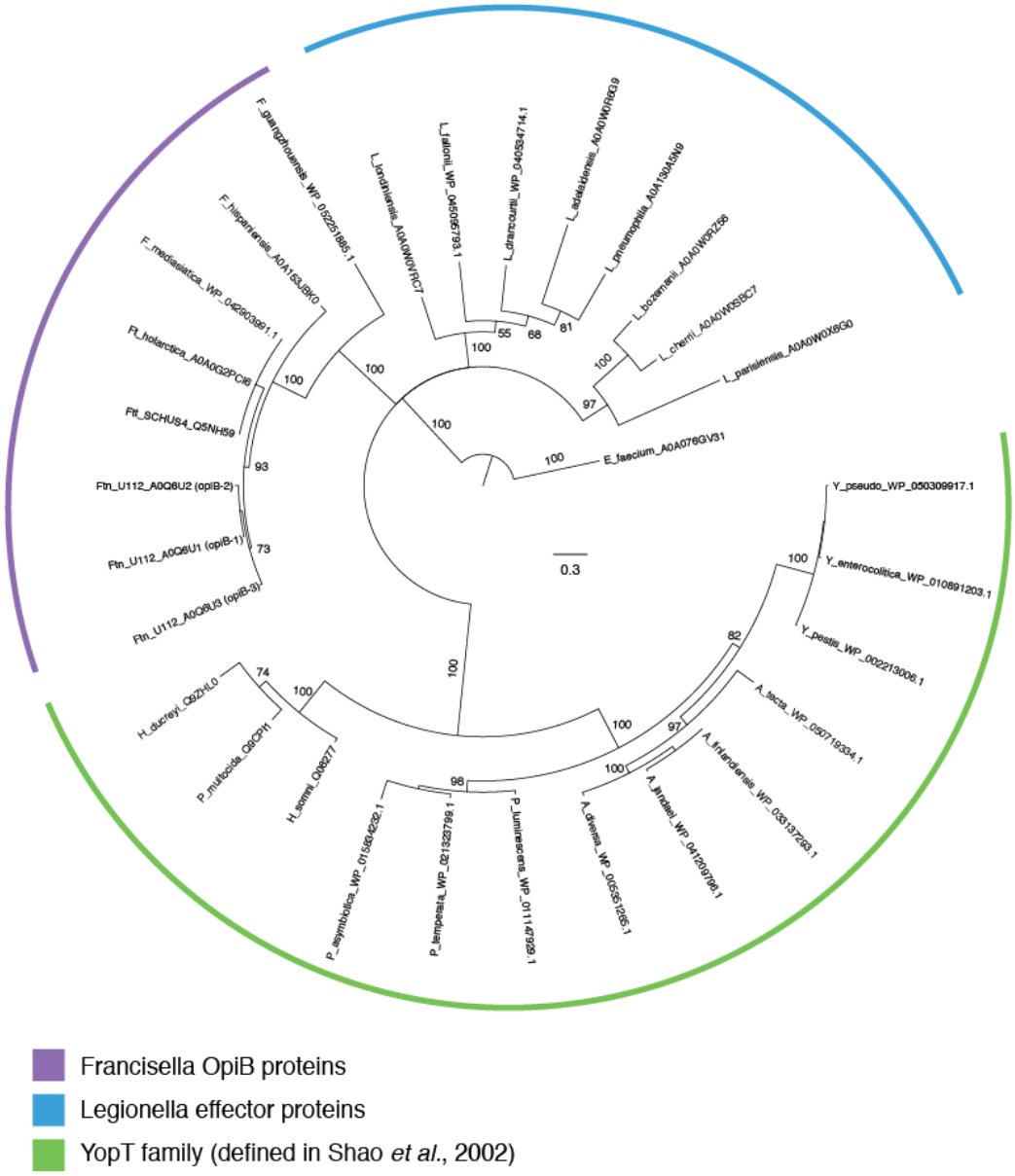

Figure S5. Phylogeny of select predicted and validated cysteine protease effectors belonging to the OpiB family (Related to Figure 6).
Sequences from representatives of type III-VI secretion system cysteine protease effectors related to the N-terminal 240 amino acids of *F. novicida* OpiB-1 were recovered from Jackhmmer, trimmed, and aligned with MAFFT (see methods, (Katoh and Standley, 2013)). The tree was constructed with MrBayes using a mixture of amino acid models with fixed rate matrices (Huelsenbeck and Ronquist, 2001). OpiB family members from Francisella, Legionella, and those belonging to the YopT family are denoted (Shao et al., 2002). The scale bar corresponds to mean amino acid substitutions per site.

Table S1. Mass spectrometry based secretome analysis of *F. novicida* T6SS^{ii} substrates (Related to Figure 1).

Locus ID	Protein name	Wild-type spectral counts[1]	Δ*dotU* spectral counts	Fold change	Protein assignment[2]
		Absent in mutant			
FTN_1309*	PdpA	47.3	0		Structural
FTN_1312*	VgrG	48.8	0		Structural
FTN_1319*	PdpC	83.6	0		Effector
FTN_1325*	PdpD	99.1	0		Effector
FTN_1071	OpiB.3	13.3	0		Effector
		Not absent in mutant			
FTN_0131	OpiA	55.8	1	55.8	Effector
FTN_1069	OpiB.1	19.1	0.5	38.2	Effector
FTN_1322*	IglC	298.4	43	6.9	Structural
FTN_1097		10.4	3	3.5	-
FTN_0111	RibH	9.4	3	3.1	-
FTN_1631	TpiA	20.8	7	3.0	-
FTN_0443	MaeA	8.5	3	2.8	-
FTN_1640	GltA	33.7	12	2.8	-
FTN_0893		15.4	7	2.2	-
FTN_0264	RpoA1	8.6	4	2.2	-
FTN_1157	EF-Tu	11.6	6	1.9	-
FTN_0116	IpdC	13.2	8	1.7	-
FTN_1044		8.2	5	1.6	-
FTN_1332	GapA	30.7	19	1.6	-
FTN_1189	RpmI	20.9	13.5	1.5	-
FTN_0279		10.7	7	1.5	-
FTN_1465		11.3	7.5	1.5	-
FTN_1444		23.2	16	1.5	-
FTN_0951	RpsF	20.7	14.5	1.4	-
FTN_1341	FabF	12.8	9	1.4	-
FTN_0266	HtpG	22.3	16.5	1.4	-
FTM_1109		10.1	7.5	1.3	-
FTN_1682	FrgA	21.3	16	1.3	-
FTN_0064	PpdK	11.3	8.5	1.3	-
FTN_1074		8.5	6.5	1.3	-
FTN_1637	SdhA	9.8	7.5	1.3	-
FTN_0980	Mdh	20.2	15.5	1.3	-
FTN_0621	Eno	13.5	10.5	1.3	-
FTN_1635	SucA	14	11	1.3	-
FTN_0255	RplR	8.2	6.5	1.3	-

FTN_1532	GdhA	27.5	22.5	1.2	-
FTN_1338	FabD	19.4	16	1.2	-
FTN_1329	FbaA	23.3	19.5	1.2	-
FTN_1485	CbpA	23.2	19.5	1.2	-
FTN_0165	LpxC	19.5	16.5	1.2	-
FTN_1743	ClpB	24.8	21	1.2	-
FTN_1333	TktA	17.2	15	1.1	-
FTN_1434	Icd	35.5	31	1.1	-
FTN_0841		21.3	19	1.1	-
FTN_0441	IleS	10	9	1.1	-
FTN_1492	LpdA	13.8	12.5	1.1	-
FTN_0505	GcvT	9.7	9	1.1	-
FTN_1087	CynT	9.1	8.5	1.1	-
FTN_1339	FabG	20	19	1.1	-
FTN_1331	Pgk	9.8	9.5	1.0	-
FTN_0088		10.3	10	1.0	-
FTN_0601	PdxS	8.2	8	1.0	-
FTN_1569	RplL	20.9	21	1.0	-
FTN_0485	GlmS	10.4	10.5	1.0	-
FTN_1634	SucB	21	21.5	1.0	-
FTN_0660	PepA	13.8	14.5	1.0	-
FTN_0669	DeoD	11.3	12	0.9	-
FTN_1494	AceE	52.1	56.5	0.9	-
FTN_0298	GplX	15.5	17	0.9	-
FTN_0337	FumA	9.1	10	0.9	-
FTN_1642	SodB	16	18	0.9	-
FTN_0085	UspA	19.3	22	0.9	-
FTN_1674	NuoG	14	16	0.9	-
FTN_0594	SucC	20	23	0.9	-
FTN_1744	ChiB	20	23	0.9	-
FTN_0593	SucD	42	49.5	0.8	-
FTN_0175		10.9	13	0.8	-
FTN_1538	GroL	54.7	65.5	0.8	-
FTN_1040	IlvC	27.4	33	0.8	-
FTN_1661	NusA	13.5	16.5	0.8	-
FTN_1569	RplL	12.9	16	0.8	-
FTN_0973		37.8	48	0.8	-
FTN_0408		14.5	18.5	0.8	-
FTN_0958		10.4	13.5	0.8	-
FTN_1537		10.4	13.5	0.8	-
FTN_1210		8.8	11.5	0.8	-
FTN_0237	FusA	38	51	0.7	-
FTN_1410	Bfr	29	39.5	0.7	-

FTN_1054	HupB	8.8	12	0.7	-
FTN_1323*	IglB	30.7	42	0.7	Structural
FTN_0564	AccC	12.6	17.5	0.7	-
FTN_1284	DnaK	27	37.5	0.7	-
FTN_1539	GroS	9.8	14	0.7	-
FTN_1568	RpoB	10.3	15	0.7	-
FTN_1058	Tig	16.5	24.5	0.7	-
FTN_0239	RplC	8.4	13	0.6	-
FTN_1186	PepO	52.7	83.5	0.6	-
FTN_1031	FtnA	8.2	13	0.6	-
FTN_1576	Tuf	67.2	108.5	0.6	-
FTN_0228	Tsf	20	32.5	0.6	-
FTN_1228	FabI	8.5	14.5	0.6	-
FTN_1260		15.3	28	0.5	-
FTN_0689	PpiC	10.4	21.5	0.5	-
FTN_1261		20.8	46.5	0.4	-
FTN_0164	FtsZ	13.2	30.5	0.4	-

*Gene located in Francisella pathogenicity island
[1]Average of normalized spectral counts from two biological replicates
[2]Protein assignments refer only to FPI-related activities and are based on this study with the exception of IglC and IglB

Table S2. Strains and plasmids used in this study (Related to Experimental Procedures)

Strains	Description
F. novicida U112	Wild-type
F. novicida Δ*dotU*	DotU residues 5 to 204 substituted with AAA
F. novicida Δ*vgrG*	VgrG residues 6 to 160 substituted with AAA
F. novicida Δ*pdpA*	PdpA residues 6 to 816 substituted with GSAAAPG
F. novicida Δ*pdpC*	PdpC residues 6 to 1321 substituted with GSAAAPG
F. novicida Δ*pdpD*	PdpD residues 8 to 1239 deleted
F. novicida Δ*iglG*	IglG residues 6 to 169 substituted with AAA
F. novicida Δ*opiA*	OpiA residues 20 to 427 deleted
F. novicida Δ*opiB*	Deletion of residue 62 of OpiB-3 to residue 771 of OpiB-1
F. novicida Δ*opiA*Δ*opiB*Δ*pdpC*Δ*pdpD*	Combined features from the appropriate strains above
F. novicida vgrG-VSV-G	Stop of VgrG substituted with codon encoding YTDIEMNRLGK-Stop; 17 bases at end of vgrG sequences repeated after stop codon to preserve RBS of downstream gene
F. novicida VSV-G-*pdpA*	MYTDIEMNRLGK replaced N-terminal M of PdpA
F. novicida VSV-G-*pdpC*	MYTDIEMNRLGK replaced N-terminal M of PdpC
F. novicida pdpD-VSV-G	Stop of PdpD substituted with codons encoding DIEMNRLGK-Stop
F. novicida iglG-VSV-G	Stop of IglG substituted with codons encoding YTDIEMNRLGK-Stop
F. novicida opiA-VSV-G	Stop of OpiA substituted with codons encoding YTDIEMNRLGK-Stop
F. novicida opiB-VSV-G	Stop of OpiB-1 substituted with codons encoding YTDIEMNRLGK-Stop
F. novicida Δ*dotU* VSV-G-*pdpC*	Combined features from the appropriate strains above
F. novicida Δ*vgrG* VSV-G-*pdpC*	Combined features from the appropriate strains above
F. novicida Δ*pdpA* VSV-G-*pdpC*	Combined features from the appropriate strains above
F. novicida Δ*pdpD* VSV-G-*pdpC*	Combined features from the appropriate strains above
F. novicida Δ*dotU pdpD*-VSV-G	Combined features from the appropriate strains above
F. novicida Δ*vgrG pdpD*-VSV-G	Combined features from the appropriate strains above
F. novicida Δ*pdpA pdpD*-VSV-G	Combined features from the appropriate strains above
F. novicida Δ*pdpC pdpD*-VSV-G	Combined features from the appropriate strains above
F. novicida Δ*pdpA vgrG*-VSV-G	Combined features from the appropriate strains above
F. novicida Δ*dotU vgrG*-VSV-G	Combined features from the appropriate strains above
F. novicida Δ*vgrG* VSV-G-pdpA	Combined features from the appropriate strains above
F. novicida Δ*dotU* VSV-G-*pdpA*	Combined features from the appropriate strains above
F. novicida Δ*iglG vgrG*-VSV-G	Combined features from the appropriate strains above
F. tularensis subspecies holarctica LVS	Wild-type
F. tularensis subspecies holarctica LVS Δ*opiA*	291 bases upstream of the OpiA ORF (WP_003014106.1) and residues 1 to 314 substituted with AAA
F. tularensis subspecies holarctica LVS Δ*opiB*	OpiB (WP_011457431.1) residues 4 to 315 substituted with AAA

Plasmids	Relevant features
pEX18-pheS-km	*F. novicida pheS* A305G expressed by *bfr* promoter, *aph* expressed by *F. novicida groES* promoter
pEX	Kanamycin resistance marker and sacB counterselection marker.
pET29b-*vgrG*-His6	*vgrG* cloned into NdeI and XhoI sites of pET29b
pET28b-His6-*pdpA*	*pdpA* cloned into BamHI and SalI sites of pET28b
pET21a-pdpA	*pdpA* cloned into BamHI and SalI sites of pET21a
pET29b-*vgrG*(G164R)-His6	*vgrG* (G164R) cloned into NdeI and XhoI sites of pET29b
pET29b-sfGFP-vgrG-His6	sfGFP, linker (AGCAGCGGC) and *vgrG* cloned into the NdeI and XhoI sites of pET29b
pF5	*Francisella* constituative expression vector
pF5:opiB-2	opiB-2 ORF cloned into the XmaI and XhoI sites of pF5
pF5:opiB-3	opiB-3 ORF cloned into the XmaI and XhoI sites of pF5

Supplemental experimental procedures

Strains and Culture Conditions. *F. novicida* strain U112 was obtained from Colin Manoil (University of Washington, Seattle, WA) and grown in tryptic soy broth or agar supplemented with 0.1% (w/v) cysteine (TSBC or TSAC). *F. tularensis* subspecies holarctica strain LVS was grown on Chocolate II agar supplemented with 1% hemoglobin and 1% IsoVitaleX (Becton, Dickinson and Company). For macrophage infections, *F. novicida* and all derived strains were grown in *F. novicida* defined media (FnDM) (Gallagher et al., 2007). *E. coli* DH5α and BL21 Rosetta 2 DE3 (EMD Millipore) were employed for cloning and expression, respectively. All bacterial strains were grown while shaking at 37°C.

Human THP-1 monocyte-like cells and A549 alveolar epithelial cells were maintained at 37°C in the presence of 5% CO_2 and cultured in RPMI 1640 or DMEM media supplemented with 10% fetal bovine serum, 4.5 g/L glucose, 2 mM glutamine, 110 mg/L sodium pyruvate, 100 U/mL penicillin, 100 µg/mL streptomycin, and 0.05 mM beta-mercaptoethanol. All mammalian cells were maintained at 37°C in the presence of 5% CO_2.

Allelic Exchange. To generate in-frame markerless chromosomal deletions and insertions in *F. novicida*, an exchange vector was constructed by using the Gibson reaction to clone the *F. novicida bfr* (FTN_1410) promoter upstream of the *F. novicida pheS* gene (FTN_0883) into pEX18km (Charity et al., 2007; Gibson et al., 2009; Zaide et al., 2011). Subsequently, site directed mutagenesis was performed to mutate the alanine at position 305 in *pheS* to glycine. Deletion and insertion constructs were generated as previously described (Mougous et al., 2006). Briefly, deletion constructs were generated by amplifying 500-1500 bases upstream and downstream of the open reading frame and then splicing by overlap extension. VSV-G insertion constructs were generated by inserting a linker encoding VSV-G just downstream of the start codon or upstream of the translation stop site for N- and C-terminal fusions, respectively. Specific sequence information is provided in Table S1. Plasmids to generate *F. tularensis* subsp. holarctica LVS deletion strains by allelic exchange were designed and integrated into LVS as previously described (Charity et al., 2007).

Allelic exchange vectors were transformed into *F. novicida* as described previously and plated on TSAC with 15 µg/mL kanamycin (Gallagher et al., 2008). Resulting merodiploids were grown overnight in non-selective TSBC, diluted 1:100 into FnDM containing 0.1% p-chlorophenylalanine and allowed to grow to stationary phase. These cultures were streaked onto non-selective TSAC and resulting kanamycin sensitive colonies were screened for deletions or insertions by colony PCR.

Mass Spectrometry of Extracellular Proteome. Samples were vacuum dried and resolved in 10 µL of acetonitrile/H_2O/formic acid (5/94.9/0.1, v/v/v) for LC-MS/MS analysis. The samples were analyzed in duplicates in a nanoLC-MS system comprising a nanoflow LC (NanoAcquity; Waters Corporation, Milford, MA) and a hybrid quadrupole-orbitrap mass spectrometer (Fusion™ Tribrid™ Thermo Scientific Corp.). Samples were loaded on a 100 µm i.d. x 20 mm 200Å, 5µm C18AQ in-house packed pre-column at flow rate of 4 µL/min for 5 min, using a loading buffer of acetonitrile/H_2O/formic acid (5/94.9/0.1, v/v/v). Peptides were separated on a 75 µm i.d. x 180 mm 100 Å, 5 µm C18AQ analytical column in a 95-min gradient at a flow rate of 250 nL/min, using mobile phase A of 0.1% formic acid in water and mobile phase B of 0.1% FA in acetonitrile. The gradient elution started at 5% mobile phase B, increased to 35% at 60 min, 80% at 65 min and held at 80% for 5 min before a 25 min re-equilibration at 5%. The eluting peptides were interrogated with an Orbitrap Fusion mass spectrometer running a data-dependent LC-MS/MS method with the Top Speed decisions selection. FTMS1 spectra were collected using the following parameters: scan range 350–1800 m/z, resolving power 120k, AGC target 4E5, and maximum injection time of 50 ms. ITMS2 spectra were collected using the following parameters: rapid scan rate, CID NCE 35, 1.6 m/z isolation window, AGC target 1E4, and maximum injection time of 50 ms. MS2 precursors were selected for a 3 sec cycle. Precursors with an assigned monoisotopic m/z and a charge state of 2–7 were interrogated. Precursors were filtered using a 60 sec dynamic exclusion window.

MS data were analyzed by MaxQuant (version 1.5.0.25) using standard settings and a UniprotKB database of *F. novicida*. Peptide-spectrum matches (PSM) and protein identifications were filtered at a false

discovery rate of 0.01 (Cox and Mann, 2008). Label-free quantification and MS/MS spectral counts were extracted and used for statistical analysis of differential expression. Three technical replicates were performed on each of two biological replicates and proteins identified in both replicates were included in further analysis. The total spectral counts for each protein were normalized to the aggregated spectral counts of ribosomal proteins for each biological replicate and proteins with average WT spectral count of less than eight were removed from the data set.

Immunoprecipitation Assays. Immunoprecipitation (IP) experiments were performed as reported previously (Silverman et al., 2013). Briefly, overnight cultures were diluted 1:100 in 1 L TSBC and allowed to grow to late log phase. Cells were harvested by centrifugation at 6,000 x G for 15 min. Pellets were resuspended in 10 mL IP buffer (20 mM Tris HCl pH 7.5, 150 mM NaCl, 1 mM AEBSF, 10 μM leupeptin, 1 μM pepstatin) supplemented with 1 mg/mL lysozyme, sonicated and centrifuged at 4°C for 30 min at 40,000 x G to pellet debris. The supernatants were incubated with 30 μL anti-VSV-G agarose beads (Sigma Aldrich) for 1 hour at 4°C and washed three times with 10 mL of IP buffer prior to analysis by SDS-PAGE and either western blot or silver stain (Invitrogen). Silver stained bands of interest were then excised and prepped for mass spectrometry analysis as per the manufacturer's instructions.

Western Blotting. Western blot analysis was performed using rabbit anti-VSV-G (Sigma Aldrich), mouse anti-IglC (BEI resources), and mouse anti-PdpA (BEI resources) antibodies diluted 1:3000. For anti-PdpD western blots, purified rabbit antisera were generated against a PdpD peptide (RGYSIDDLNRKYRP; GenScript) and used at a dilution of 1:1000. To detect primary antibodies, anti-mouse and anti-rabbit antibodies conjugated to either horse radish peroxidase (diluted 1:5000, Sigma Aldrich) or fluorescent molecules (1:20,000, LI-COR) were employed. Membranes were visualized using FluorChem Q (Protein Simple) or LI-COR.

Protein Expression and Purification. For expression and purification of proteins, DNA sequences were cloned from *F. novicida* U112 genomic DNA into pET vectors (sites and vectors are listed in Table S1). Plasmids were transformed into chemically competent *E. coli* BL21 Rosetta 2DE3 and grown to stationary phase overnight. These cultures were then used to inoculate 1L of 2X YT broth and the cultures were grown to OD600 0.6 and then induced with 1mM IPTG overnight at 18°C. Cells were harvested by centrifugation, resuspended in 20 mM Tris-HCl, 500 mM NaCl, 15mM imidazole, 1 mM AEBSF, 10 μM leupeptin, and 1 μM pepstatin. Cells were disrupted by sonication and cellular debris was removed by centrifugation at 40,000 x G for 30 min. His-tagged proteins were purified from cleared lysates using a 1mL HisTrap FF NI-TA cartridge on an AKTA FPLC purification system (GE Healthcare). Bound proteins were eluted using a linear imidazole gradient to a final concentration of 500 mM. The purity of each protein sample was assessed by SDS-PAGE and Coomassie brilliant blue staining.

Negative Staining Electron Microscopy Image Analysis. Particles were picked automatically in a reference-free manner using DogPicker (Voss et al., 2009). Contrast transfer function (CTF) estimation was performed using CTFFIND3 (Mindell and Grigorieff, 2003). 2D classifications were performed using RELION 1.4 and included correction for the effect of the microscope CTF (Scheres, 2012). *E. coli* PdpA complex particles were extracted with a box size of 224 Å and the displayed class contains 6,989 particles from an original 75,388. *E. coli* VgrG-PdpA complex particles were extracted with a box size of 320 Å and the displayed class contains 2,102 particles from an original 65,976. *F. novicida* VgrG particles were extracted with a box size of 400 Å and the displayed class contains 4,323 particles from an original 135,128. Measurements of VgrG lengths were made using Image J.

Infection Assays. 1.3 x 10^6 THP-1 cells were seeded without antibiotic in 24 well plates, differentiated with 100-200nM phorbol 12-myristate 13-acetate (PMA) for 18-24 hours and rinsed with fresh media. The cells were washed and exposed to mid-log phase *Francisella* at a multiplicity of infection of 0.5-5. The infection was synchronized by centrifugation at 850 x G at room temperature for 15 min followed by a 45 min incubation at 37°C. After washing to remove extracellular bacteria, the cells were either processed for immunofluorescence or treated with gentamycin, rinsed again and lysed by the addition of Triton X100 or Saponin to 0.1% or 1%, respectively, for CFU enumeration.

Immunofluorescence. After staining with mouse anti-human LAMP1 and goat anti-*Francisella* primary antibodies (described in main body), nuclei were stained with Hoechst 33342 (Thermo Fisher) for 5 min in water and samples were mounted using Mowiol 4-88 mounting medium. Epifluorescence microscopy was performed using either a Leica DM4000 B LED for quantitative analysis or a Leica TCS SP8 Confocal Microscope for confocal image acquisition. Representative confocal micrographs were acquired as 1024 x 1024 pixels images using a 63x/1.4 NA HC PL APO objective and assembled using Adobe Photoshop CS6.

Bioinformatics. BLASTp was utilized to query the Integrated Microbial Genomes database to find OpiA, OpiB, PdpC, PdpD, PdpA, and VgrG homologs with ≥30% identity. An organism was classified as containing an FPI if *pdpA*, *vgrG*, and *iglC* homologs were identified within 20 kilobases. The phylogenetic tree presented is based on the analysis of Francisella genomes by (Sjodin et al., 2012). The cysteine protease effector family represented by OpiB was identified by subjecting the N-terminal 240 amino acids of the protein to analysis by Jackhmmer (https://www.ebi.ac.uk/Tools/hmmer/search/jackhmmer). This analysis identified a conservative group of >150 members after three iterations. An NCBI conserved domain search in conjunction with secondary structure assignment by Phyre2 and the output of the ankyrin repeat domain prediction system (ARD; http://ard.cs.ntou.edu.tw/) domain was used to define the architecture of the ankyrin repeat domain of OpiB. The search for homologs of OpiB was refined by including related proteins described by Shao *et al.* in analyses utilizing PHI-BLAST and PSI-BLAST (Shao et al., 2002). The N- and C-terminal ends were then manually trimmed based on the preliminary alignments. A full multi-alignment was generated using MAFFT version 7.273 (Katoh and Standley, 2013). We employed the G-INS-i + VMS algorithm with an a_{max} 0.8 to guard against over-alignment (Katoh and Standley, 2016). The output was used to construct the phylogenetic tree using Mr. Bayes 3.2.6, according to the Mr. Bayes manual sections 2.2 and 4.1. Two runs with a million generations were processed with three heated chains and one cold chain. Summary statistics for partitions with frequency ≥ 0.10 in at least one run suggested that the number of generations was sufficient, with the average standard deviation of split frequencies at 0.007713 (maximum 0.044905) and the average Potential Scale Reduction Factor for parameter values at 1.000 (maximum 1.004) (Gelman and Rubin, 1992). Based on the hypothesis that clans evolve independently, we used another CA clan cysteine protease from the Gram-positive organism *Enterococcus faecium* (family C47) to root the tree (Barrett and Rawlings, 2001). The tree was displayed and adjusted using FigTree 1.4.2 17 (http://tree.bio.ed.ac.uk/software/figtree/). The sequence logo was generated with Jalview 2.9.0b29 (Katoh and Standley, 2013; Waterhouse et al., 2009).

Supplemental References

Barrett, A.J., and Rawlings, N.D. (2001). Evolutionary lines of cysteine peptidases. Biol Chem *382*, 727-733.

Charity, J.C., Costante-Hamm, M.M., Balon, E.L., Boyd, D.H., Rubin, E.J., and Dove, S.L. (2007). Twin RNA polymerase-associated proteins control virulence gene expression in Francisella tularensis. PLoS Pathog *3*, e84.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature biotechnology *26*, 1367-1372.

Gallagher, L.A., McKevitt, M., Ramage, E.R., and Manoil, C. (2008). Genetic dissection of the Francisella novicida restriction barrier. J Bacteriol *190*, 7830-7837.

Gallagher, L.A., Ramage, E., Jacobs, M.A., Kaul, R., Brittnacher, M., and Manoil, C. (2007). A comprehensive transposon mutant library of Francisella novicida, a bioweapon surrogate. Proc Natl Acad Sci U S A *104*, 1009-1014.

Gelman, A., and Rubin, D.B. (1992). Inference from Iterative Simulation Using Multiple Sequences. Statistical Science *7*, 457-472.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature methods *6*, 343-345.

Huelsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics (Oxford, England) *17*, 754-755.

Kast, P. (1994). pKSS--a second-generation general purpose cloning vector for efficient positive selection of recombinant clones. Gene *138*, 109-114.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular biology and evolution *30*, 772-780.

Katoh, K., and Standley, D.M. (2016). A simple method to control over-alignment in the MAFFT multiple sequence alignment program. Bioinformatics (Oxford, England) *32*, 1933-1942.

Mindell, J.A., and Grigorieff, N. (2003). Accurate determination of local defocus and specimen tilt in electron microscopy. J Struct Biol *142*, 334-347.

Mougous, J.D., Cuff, M.E., Raunser, S., Shen, A., Zhou, M., Gifford, C.A., Goodman, A.L., Joachimiak, G., Ordonez, C.L., Lory, S., *et al.* (2006). A virulence locus of Pseudomonas aeruginosa encodes a protein secretion apparatus. Science *312*, 1526-1530.

Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol *180*, 519-530.

Shao, F., Merritt, P.M., Bao, Z., Innes, R.W., and Dixon, J.E. (2002). A Yersinia effector and a Pseudomonas avirulence protein define a family of cysteine proteases functioning in bacterial pathogenesis. Cell *109*, 575-588.

Silverman, J.M., Agnello, D.M., Zheng, H., Andrews, B.T., Li, M., Catalano, C.E., Gonen, T., and Mougous, J.D. (2013). Haemolysin Coregulated Protein Is an Exported Receptor and Chaperone of Type VI Secretion Substrates. Molecular cell *51*, 584-593.

Sjodin, A., Svensson, K., Ohrman, C., Ahlinder, J., Lindgren, P., Duodu, S., Johansson, A., Colquhoun, D.J., Larsson, P., and Forsman, M. (2012). Genome characterisation of the genus Francisella reveals insight into similar evolutionary paths in pathogens of mammals and fish. BMC genomics *13*, 268.

Voss, N.R., Yoshioka, C.K., Radermacher, M., Potter, C.S., and Carragher, B. (2009). DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. J Struct Biol *166*, 205-213.

Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics (Oxford, England) *25*, 1189-1191.

Zaide, G., Grosfeld, H., Ehrlich, S., Zvi, A., Cohen, O., and Shafferman, A. (2011). Identification and characterization of novel and potent transcription promoters of Francisella tularensis. Applied and environmental microbiology *77*, 1608-1618.