A Widespread Bacterial Secretion System with Diverse Substrates

Alex S. Grossman,[a] Terra J. Mauer,[b*] Katrina T. Forest,[b] Heidi Goodrich-Blair[a,b]

[a]University of Tennessee—Knoxville, Department of Microbiology, Knoxville, Tennessee, USA
[b]University of Wisconsin—Madison, Department of Bacteriology, Madison, Wisconsin, USA

Alex S. Grossman and Terra J. Mauer contributed equally to this work. Author order was determined on the basis of Grossman's more recent contributions.

ABSTRACT In host-associated bacteria, surface and secreted proteins mediate acquisition of nutrients, interactions with host cells, and specificity of tissue localization. In Gram-negative bacteria, the mechanism by which many proteins cross and/or become tethered to the outer membrane remains unclear. The domain of unknown function 560 (DUF560) occurs in outer membrane proteins throughout *Proteobacteria* and has been implicated in host-bacterium interactions and lipoprotein surface exposure. We used sequence similarity networking to reveal three subfamilies of DUF560 homologs. One subfamily includes those DUF560 proteins experimentally characterized thus far: NilB, a host range determinant of the nematode-mutualist *Xenorhabdus nematophila*, and the surface lipoprotein assembly modulators Slam1 and Slam2, which facilitate lipoprotein surface exposure in *Neisseria meningitidis* (Y. Hooda, C. C. Lai, A. Judd, C. M. Buckwalter, et al., Nat Microbiol 1:16009, 2016, https://doi.org/10.1038/nmicrobiol.2016.9; Y. Hooda, C. C. L. Lai, T. F. Moraes, Front Cell Infect Microbiol 7:207, 2017, https://doi.org/10.3389/fcimb.2017.00207). We show that DUF560 proteins from a second subfamily facilitate secretion of soluble, nonlipidated proteins across the outer membrane. Using *in silico* analysis, we demonstrate that DUF560 gene complement correlates with bacterial environment at a macro level and host association at a species level. The DUF560 protein superfamily represents a newly characterized Gram-negative secretion system capable of lipoprotein surface exposure and soluble protein secretion with conserved roles in facilitating symbiosis. In light of these data, we propose that it be titled the type 11 secretion system (TXISS).

IMPORTANCE The microbial constituency of a host-associated microbiome emerges from a complex physical and chemical interplay of microbial colonization factors, host surface conditions, and host immunological responses. To fill unique niches within a host, bacteria encode surface and secreted proteins that enable interactions with and responses to the host and co-occurring microbes. Bioinformatic predictions of putative bacterial colonization factor localization and function facilitate hypotheses about the potential of bacteria to engage in pathogenic, mutualistic, or commensal activities. This study uses publicly available genome sequence data alongside experimental results from *Xenorhabdus nematophila* to demonstrate a role for DUF560 family proteins in secretion of bacterial effectors of host interactions. Our research delineates a broadly distributed family of proteins and enables more accurate predictions of the localization of colonization factors throughout *Proteobacteria*.

All plants and animals exist in association with bacterial symbionts that contribute to nutrition, protection, development, and reproduction. These symbionts express surface and secreted proteins that facilitate host interactions through a variety of

functions, including acquisition of nutrients (1, 2), interaction with host cells (3), and specificity in host range and tissue localization (4). Possibly due to the complexity of bacterial membranes and the breadth of biophysical characteristics of secreted proteins, there is no singular export pathway.

Bacteria have two broadly distributed export systems, including the inner membrane spanning Sec and twin arginine translocation (Tat) systems which are shared between Gram-positive and Gram-negative bacteria. Diderms require additional secretion systems for transport into and across the outer membrane. The type I, II, V, IX, and X secretion systems work to transport substrates across that outer membrane, while types III, IV, and VI go one step further, moving effector proteins across the outer membrane and directly into another organism (5–7). Despite the many secretion systems described thus far, there are still a number of secreted proteins which are known to contribute to symbiosis but for which no transport system is known. Here we describe a machinery present throughout the phylum *Proteobacteria* that is responsible for secreting soluble proteins and lipoproteins from the periplasm across the outer membrane.

Recently, the Slam (surface lipoprotein assembly modulator) mechanism of lipoprotein surface tethering was identified in *Neisseria meningitidis* (8, 9). Slam proteins containing the β-barrel DUF560 (domain of unknown function 560) (also termed SlipAM domain) (10) are required for surface presentation of certain lipoproteins, including those that capture metal-carrying compounds used by hosts to sequester nutrients (8). Two *N. meningitidis* Slam proteins have been characterized, with distinct lipoprotein substrates which contribute to symbiosis. Slam activity also has been demonstrated for DUF560 representatives from pathogens *Pasteurella multocida*, *Moraxella catarrhalis*, and *Haemophilus influenzae* (8, 9). However, most Slam homologs have no bioinformatically predicted substrate thus far, and one study has found that *N. meningitidis* Slam1 can surface expose unprocessed factor H binding protein (fHbp) single nucleotide polymorphism (SNP) variants (11). Thus, the full functional potential of DUF560 proteins is not yet known.

The DUF560 homolog NilB is a host association and species specificity factor in the nematode symbiont *Xenorhabdus nematophila*, a proteobacterium in the family *Morganellaceae* (12–14). A screen for *X. nematophila* mutants defective in colonizing *Steinernema carpocapsae* intestines revealed the nematode intestinal localization (*nil*) locus (14, 15). The *nil* locus contains the genes *nilB* and *nilC*, each of which is independently necessary for colonization of nematodes. Biochemical and bioinformatic analyses have established that NilC is an outer membrane-associated lipoprotein, and NilB is an outer membrane β-barrel in the DUF560 family with an ~140-amino-acid periplasmic N-terminal domain that contains tetratricopeptide repeats (15–18).

To begin to understand the functional range of DUF560 proteins, we assessed ecological distribution, genomic context, and relatedness. We experimentally examined the *X. nematophila* DUF560 homolog HrpB, which is predicted to transport an unlipidated protein. Finally, to better understand the potential role of DUF560 proteins in host-symbiont interactions, we analyzed distribution among symbiotic *Xenorhabdus*. Our data demonstrate that biological activities of the DUF560 family extend beyond lipoprotein surface presentation and constitute the type 11 bacterial secretion system (TXISS) which, like the type II secretion system, is capable of acting on membrane-anchored or soluble proteins (Fig. 1) (19).

(This article was submitted to an online preprint archive [20].)

RESULTS

TXISS cluster according to environment. Using homology to NilB or Slam proteins, previous work identified a wide distribution of DUF560 proteins within mucosa-associated bacteria (9, 14, 15). To quantifiably delineate subfamilies within the TXISS, we generated a sequence similarity network (SSN) using the Enzyme Function Initiative toolset (EFI) (21–23) and annotated it to highlight environmental source or taxonomic grouping of microbes containing DUF560 homologs (Fig. 2; see also Table S1 in the supplemental material). In this
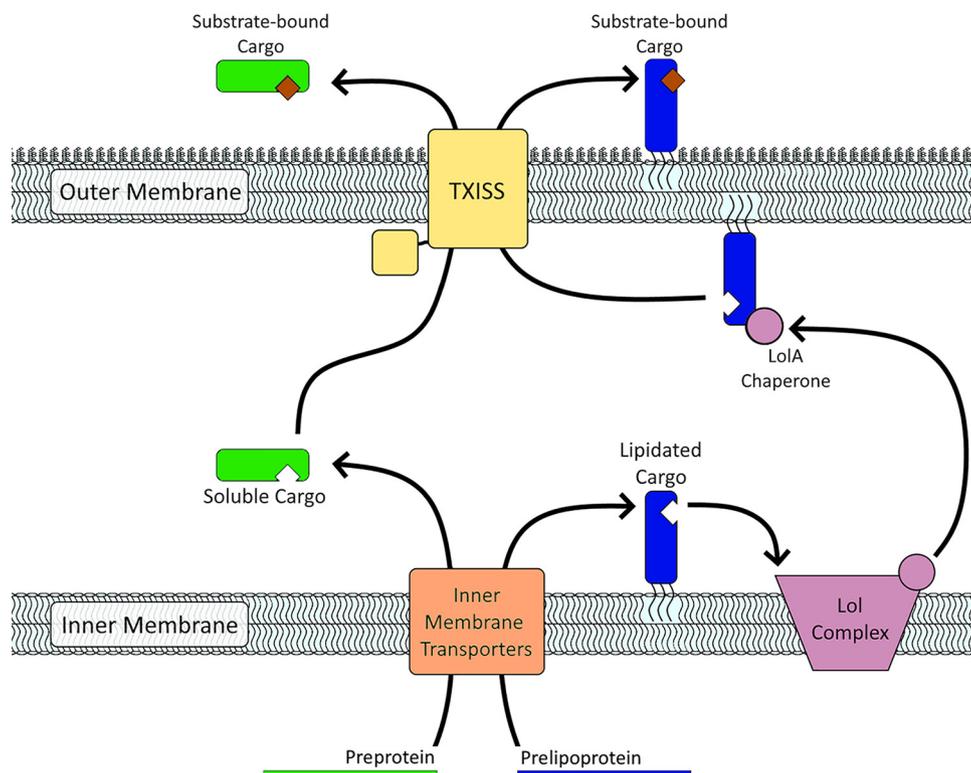
FIG 1 Conceptual model of the type XI secretion system. TXISS outer membrane proteins (yellow) of the DUF560 family are necessary and sufficient to secrete lipidated (blue) and soluble (green) cargo proteins across the outer membrane to the extracellular milieu. TXISS-dependent cargo proteins are exported into the periplasm via inner membrane transporters (orange). Once in the periplasm, lipidated TXISS cargo proteins, such as TbpB, are expected to be chaperoned across the hydrophilic periplasm by the localization of lipoproteins (Lol) complex. It is currently unknown whether periplasmic soluble TXISS cargo proteins have dedicated molecular chaperones to reach the outer membrane. Once secreted by TXISS, cargo proteins can bind to their specific substrates.

analysis, sequences with high identity (≥40%) were gathered into nodes and connected with edges based on sequence similarity. Using all homologs in Interpro 73 and UniProt 2019-02, we identified 10 major clusters of TXISS proteins. Cluster 1 was chosen for in-depth analysis as it contained the majority of nodes in the network (62.4%) and could be clearly divided into three subclusters (1A, 1B, and 1C) using force-directed node placement (Fig. 3). The remaining nine clusters displayed a preponderance of water- and soil-associated organisms and contained no characterized proteins (see Table S1 for details). Consistent with our previous observations (15), the majority of cluster 1 nodes (75%) comprise sequences from animal-associated isolates, while another 20% contain sequences from marine, freshwater, soil, or built-environment isolates.

The division of cluster 1 nodes strongly reflected environmental origin (Fig. 3A). Subcluster 1A almost exclusively comprises animal-associated bacteria and contains all previously characterized TXISS, which separate according to the predicted substrate when analyzed with higher stringency (see Fig. S1 in the supplemental material). Subclusters 1B and 1C have no previously characterized representatives. Subcluster 1B contains a mixture of sequences from host-associated and free-living bacteria, while subcluster 1C contains sequences largely from environmentally isolated *Sphingomonadaceae*. Subclusters 1A and 1B correlate poorly with taxonomy (Fig. 3B). For example, cluster 1 contains 81 nodes with *Moraxellaceae* sequences. Of these, 79% are animal associated and are enriched in subcluster 1A, while 12.3% are environmental isolates and are enriched in subcluster 1B. These data demonstrate a correlation with lifestyle (e.g., free-living versus host associated) as opposed to taxonomy and suggest that subclusters indicate divergent functions.

TXISS cluster according to substrate. The Slam acronym was defined on the basis that DUF560 homologs from *N. meningitidis* facilitate surface exposure of lipoproteins,
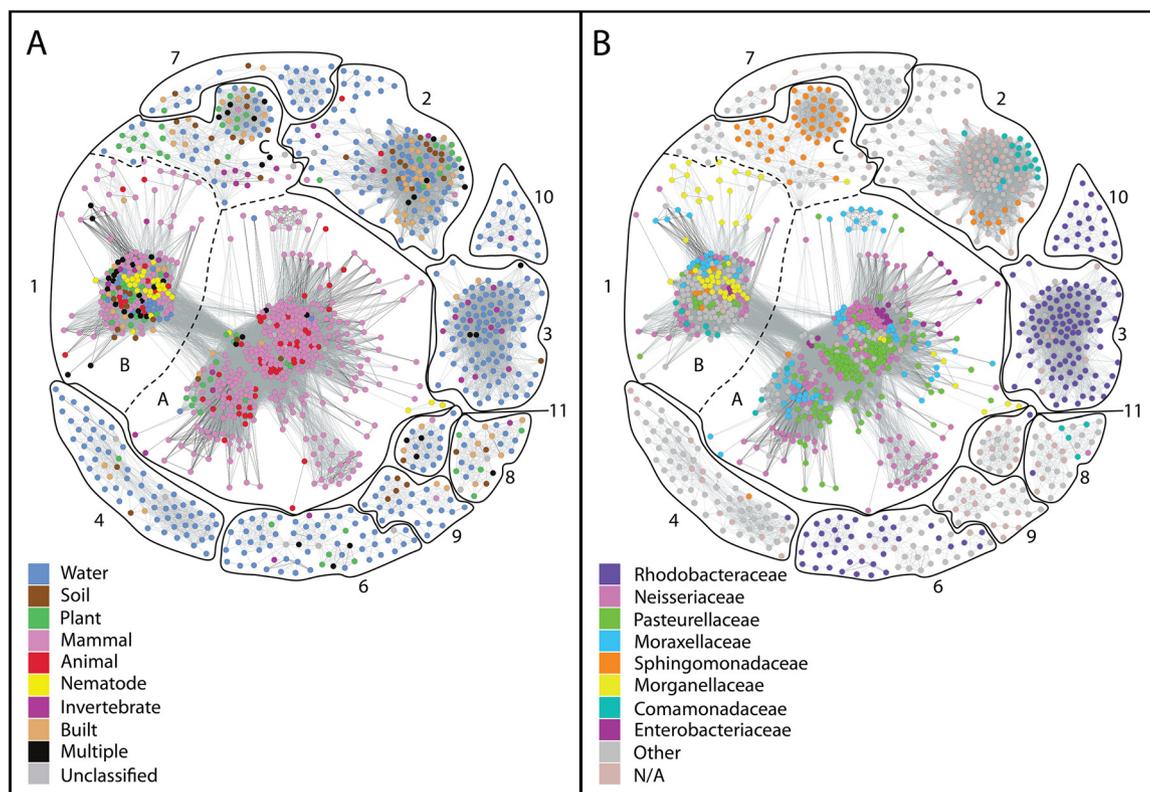
FIG 2 Sequence similarity network (SSN) of all DUF560 proteins. Network generated by EFI-EST as accessed on 24 April 2019 (1–3). Edges were drawn using an alignment score of 38, and any sequences which shared ≥40% identity were placed in a single node to allow the separation of clusters. Each node represents a group of highly similar sequences, with edge darkness demonstrating similarity, and the distance between nodes being determined via the Fruchterman-Reingold algorithm to optimize edge lengths (59). Each node was color coded to show the isolates' environmental origin(s) (A) and taxonomic class (B). N/A, not applicable.

such as transferrin binding protein B (TbpB), lactoferrin binding protein B (LbpB), hemoglobin/haptoglobin binding protein A (HpuA), and fHbp, which are frequently encoded nearby (8, 9). Yet the lipid tail is not essential for Slam-dependent surface exposure of a target (8, 11, 24), prompting us to consider whether cooccurrence with lipoproteins is a constant. We used the EFI Genome-Neighborhood Tool (21, 23, 25) to assay the genomic context of cluster 1 members. This analysis corroborated previous work demonstrating genomic association of DUF560 proteins with TonB, TonB-dependent receptors, and proteins that have a Pfam TbpB_B_D domain, which will be referred to hereafter as TbpBBD (9, 26).

Given the prevalence of TbpBBD domain-containing genes in the neighborhoods of DUF560 genes, and their known occurrence in lipoproteins exposed by Slams, we examined their open reading frames (ORFs) for potential patterns. Using a combination of genome-neighborhood-network data (23), Rapid ORF Description & Evaluation Online (RODEO) data (27), and manual annotation, we analyzed 851 unique TbpBBD domain proteins coinherited with cluster 1 DUF560s (Table S2). The majority (75.1%) of TbpBBD-bearing proteins associated with subcluster 1A have a signal peptidase 2 (SPII) signal peptide, indicating lipidation, and two TbpBBD domains, similar to TbpB in *N. meningitidis* (9). These are referred to here as TbpBBD$_{lip}$ (lipidated TbpBBD). In contrast, the TbpBBD-bearing proteins associated with subcluster 1B are almost exclusively predicted to be soluble proteins (97.8%) with signal peptidase 1 (SPI) signal peptides and a single TbpBBD domain, similar to hemophilin in *Haemophilus haemolyticus*. Hereafter we refer to this class of proteins as TbpBBD$_{sol}$ (for soluble TbpBBD) (Fig. 3C and Fig. 4). TbpBBD$_{lip}$ and TbpBBD$_{sol}$ proteins are predicted to be translocated across the inner membrane through the Sec secretion pathway. However, at least some TbpBBD$_{sol}$ proteins have similarities to *Escherichia coli* OmpA and DsbA, which can be secreted through either
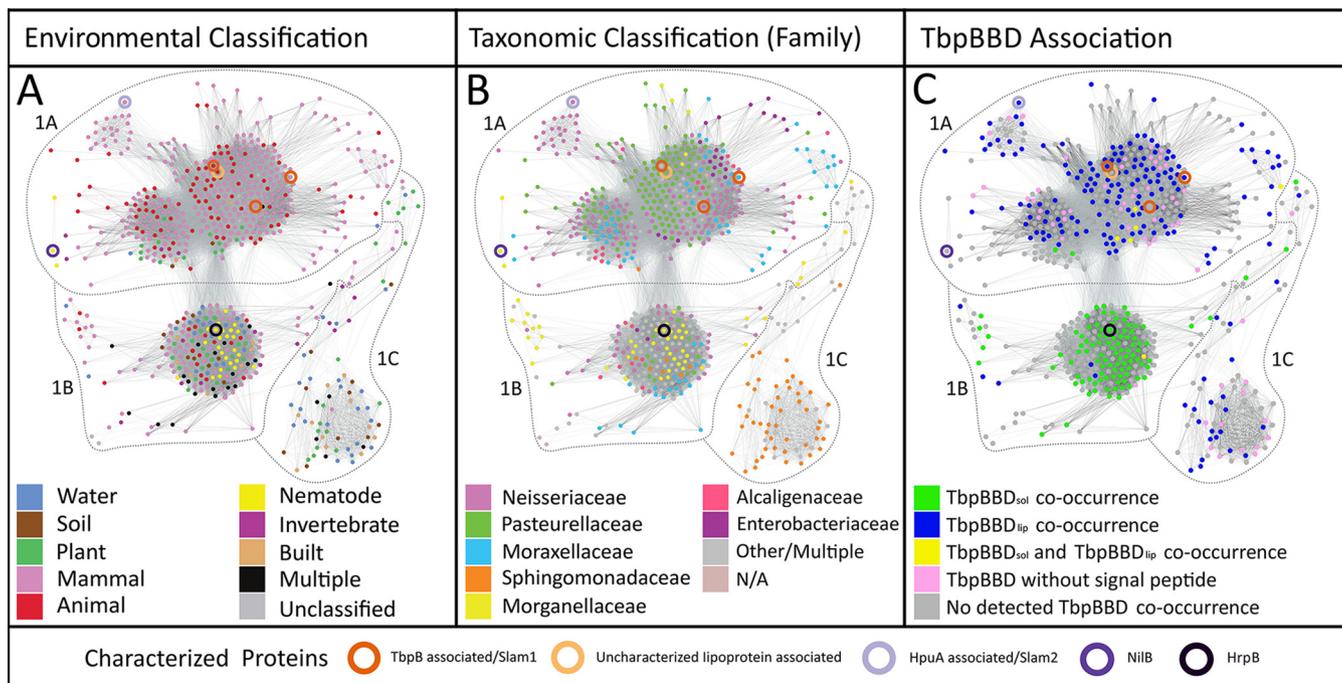
FIG 3 Cluster 1 of the TXISS SSN forms subclusters according to environment of isolation and signal sequence of predicted cargo. All nodes from cluster 1 of the TXISS SSN were positioned using the Fruchterman-Reingold algorithm. (A to C) The resulting graph was annotated according to environmental origin(s) (A), taxonomic class of the isolates (B), or whether the node homolog(s) cooccur with a TbpBBD domain-encoding gene (C). Nodes containing experimentally characterized proteins are highlighted using colored circles as indicated at the bottom of the figure.

the Sec or Tat secretion pathways (Fig. 4), leaving open the possibility that some TXISS-1B TbpBBD$_{sol}$ proteins are conditionally Tat secreted (28).

Biochemical and structural evidence support the conclusion that the TbpBBD$_{sol}$ protein hemophilin is a soluble secreted protein that binds free heme and facilitates heme uptake (29). Three-dimensional homology modeling (Phyre[2]) was used to visualize potential structural similarities between hemophilin and several TbpBBD$_{sol}$ proteins: a previously described homolog *X. nematophila* XNC1_0075 and two of its most closely related homologs, *Providencia rettgeri* PROVRETT_08181 and *Proteus mirabilis* WP_134940027.1 (15, 30–32) (Fig. S2). No high confidence models were found for the first ~50 residues, which correspond to the position of the variable N-terminal handle domains of TbpBBD$_{lip}$ proteins. However, the structures of the central regions and the C-terminal TbpBBD β-barrel domains were predicted with high confidence (>99%) based on hemophilin. In light of sequence and structural level similarities, we predict two functional TXISS subfamilies; TXISS-1A members flip TbpBBD$_{lip}$ substrates across the outer membrane, and TXISS-1B members secrete TbpBBD$_{sol}$ substrates into the extracellular milieu.

TXISS-1B activity reconstructed *in vivo*. To experimentally evaluate our prediction that TXISS-1B can secrete TbpBBD$_{sol}$ substrates, we investigated the heme receptor protein (Hrp) locus of *X. nematophila*. This locus, conserved across the *Xenorhabdus* genus, encodes TonB, a TonB-dependent heme receptor HrpA (XNC1_0073), the TXISS-1B homolog HrpB (XNC1_0074), and its predicted TbpBBD$_{sol}$ substrate HrpC (XNC1_0075) (15), the hemophilin homolog described above (29). Specifically, we sought to test whether HrpB mediates secretion of HrpC. Immunotagged HrpC was expressed with or without immunotagged HrpB in *E. coli*. Whole-cell and culture supernatant fractions were separated and analyzed by immunoblotting with anti-FLAG antibody (33). Whole-cell lysates demonstrated equivalent expression of HrpC in all treatments (Fig. 5A). In the presence of FLAG-tagged HrpB (FLAG-HrpB), the levels of HrpC-FLAG detected in culture supernatants increased 9.9-fold at 1 h postinduction and 17.0-fold at 2.5 h postinduction (Fig. 5A and B), demonstrating that secretion of HrpC is mediated by HrpB.

A trivial explanation for these data could be that HrpB expression leads to cell lysis. The supernatant protein profile revealed by Coomassie blue staining of sodium dodecyl
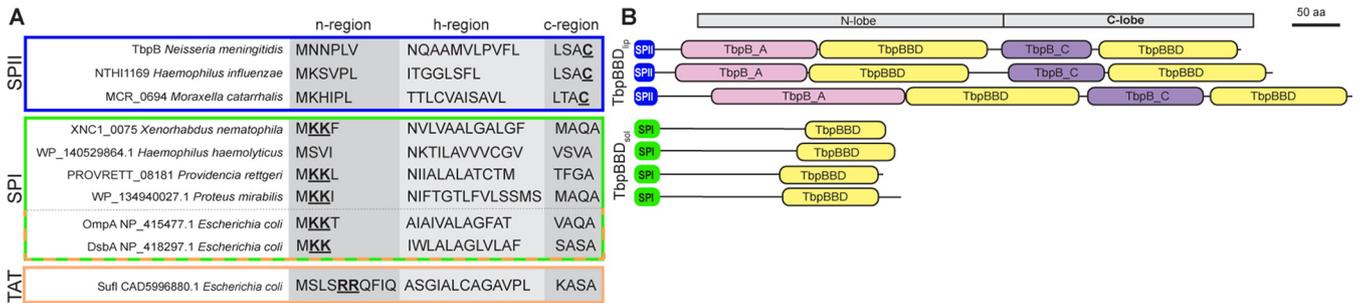
FIG 4 Examples of TbpBBD$_{lip}$ and TbpBBD$_{sol}$ signal peptides and domain architectures. TXISS-1A-associated proteins from *N. meningitidis*, *H. influenzae*, and *M. catarrhalis* have signal peptidase 2 (SPII; blue) signal sequences, while TXISS-1B associated proteins from *X. nematophila*, *H. hemolyticus*, *P. rettgeri*, and *P. mirabilis* have signal peptidase 1 (SPI; green) signal sequences. (A) SPII and SPI signal sequences, comprising n-, h-, and c-regions of representative TXISS-1A (blue) and TXISS-1B (green, above dashed line) cargo proteins and two *E. coli* proteins (OmpA and DsbA) that can be secreted through either the Sec or Tat machinery (below dashed line, green-orange border), are compared to a canonical Tat secretion signal (orange) found in the *E. coli* protein SufI. Underlined amino acids highlight conserved features of signal sequences, including the acylated cysteine of lipoproteins, the twin arginine motif of Tat signal peptides, and n-region twin lysines that are permissive for Tat secretion (28). (B) The schematic diagram shows features found in select TbpBBD domain-containing proteins predicted to be exported by TXISS mechanisms. TXISS-1A-associated cargo proteins shown have SPII (blue), N-lobe and C-lobe handle domains TbpB_A (pink) and TbpB_C (purple), respectively, and two TbpBBD domains (yellow). TXISS-1B-associated cargo proteins shown have SPI (green), lack annotated handle domains, and have a single TbpBBD domain (yellow).

sulfate-polyacrylamide gels (SDS-PAGs) did not indicate cell lysis or nonspecific protein secretion (Fig. S3). To further rule out this possibility, we created an HrpC−glutathione *S*-transferase (GST) fusion protein and coexpressed it with FLAG-HrpB. Previous observations had demonstrated that GST-fused TbpB is not surface exposed by Slam1 (8). Consistent with this finding, the levels of HrpC-GST in a culture supernatant were unaffected by coexpression with FLAG-HrpB, supporting the conclusion that expression of the outer membrane protein HrpB does not cause bacterial lysis (Fig. 5A and B) and moreover demonstrating that HrpB cannot secrete its substrate when HrpC is fused to a 26-kDa protein.

In the absence of Slam1, a fraction of the cargo protein fHbp is surface exposed when expressed in *E. coli* BL21(DE3) (34, 35). Similarly, we found that some HrpC-FLAG and HrpC-GST reached the supernatant in the absence of FLAG-HrpB. Furthermore, FLAG-HrpB was unexpectedly detected in the supernatant fraction (Fig. 5B). To distinguish soluble supernatant proteins from those that may be associated with insoluble membrane components (e.g., outer membrane vesicles [OMVs]), we depleted insoluble components of the supernatant via ultracentrifugation and tested the clarified soluble fraction for Hrp proteins. Effectively all HrpB in the supernatant was removed with ultracentrifugation, suggesting it is localized in OMVs (Fig. 5C) (36). Also, in the absence of HrpB coexpression, the levels of HrpC detected in the supernatant were reduced upon removal of insoluble material, while supernatant levels of HrpC detected after coexpression with HrpB were not affected by removal of insoluble material (Fig. 5D). We conclude that HrpB is membrane bound and that it secretes soluble HrpC protein into the extracellular milieu.

Host environment drives TXISS class. Having established that DUF560 homologs represent a bona fide secretion system, we next used the *Xenorhabdus* system to expand on our observation that the presence and type of TXISS correspond to environmental niche. *Xenorhabdus* bacteria are species-specific obligate mutualists of *Steinernema* nematodes, and NilB is a known host range determinant. Therefore, we considered host species as an environmental niche and bioinformatically examined whether the complement of DUF560 genes in a microbe corresponds with host phylogeny (37–41). All 46 *Xenorhabdus* genomes on the Magnifying Genome (MaGe) platform (42) encode between one and three TXISS, with five distinct homologs represented across the genus (one TXISS-1A and four TXISS-1B) (43). The one TXISS homolog that is conserved among all *Xenorhabdus* species is TXISS-1B HrpB. The genomic contexts of six representative homologs are shown in Fig. S4. Each unique combination of TXISS homologs was assigned one of six classes (A to F) (Fig. S4 and Table S3). To visualize correlations between TXISS class and host identity, we constructed maximum likelihood and Bayesian phylogenetic trees for *Xenorhabdus* and *Steinernema*. *Xenorhabdus* trees were generated with whole-genome data, while *Steinernema* trees used

A Fold Change of Supernatant HrpC Induced by HrpB-FLAG
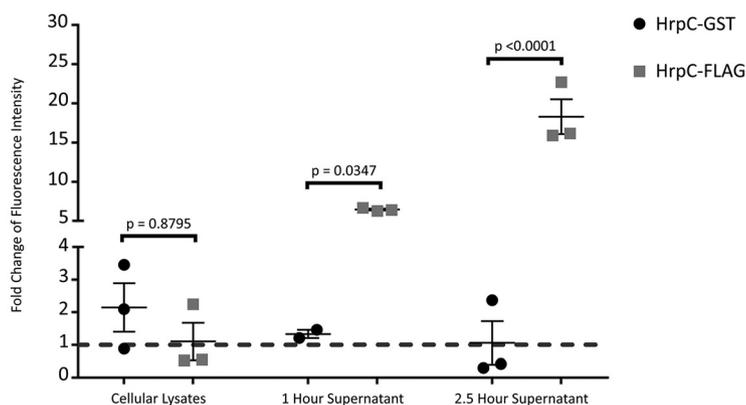
B

C HrpB-FLAG Within Supernatant Fractions
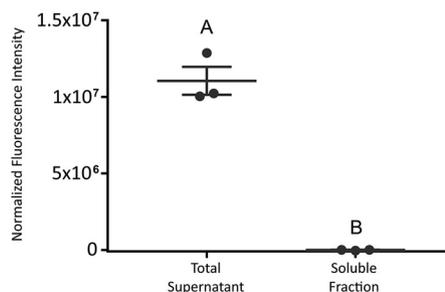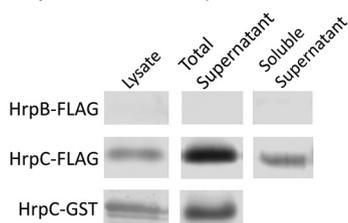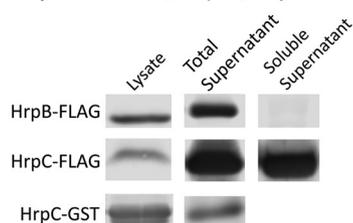
D HrpC-FLAG Within Supernatant Fractions
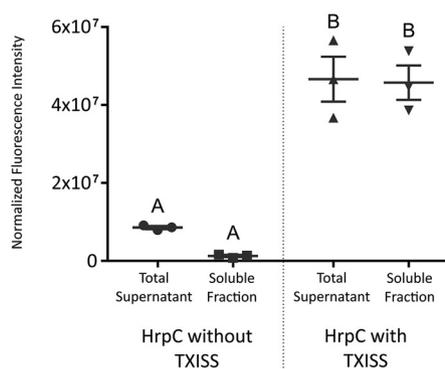
FIG 5 HrpB increases secretion of HrpC. (A) Demonstration of HrpB-dependent secretion. HrpC was detected and quantified as fluorescence intensity observed in immunoblots with anti-tag (FLAG or GST) antibodies (see Fig. S3 for complete blots). Fold change values shown are the fluorescence intensity of HrpC coexpressed with HrpB, divided by the intensity of HrpC expressed alone as observed from immunoblots. Each dot represents fold change derived from a distinct biological replicate pair and standard error is shown. At both assayed time points, the total supernatant concentration of HrpC-FLAG was increased by coexpression with FLAG-HrpB compared to the HrpC-GST treatment which was seemingly unaffected by the presence of FLAG-HrpB (unpaired *t* test). (B) Representative immunoblots comparing FLAG-HrpB, HrpC-FLAG, and HrpC-GST in cellular and supernatant fractions. (C) Fluorescence intensity of HrpB in supernatant fractions normalized to the Coomassie blue loading control. Each dot represents a distinct technical replicate. HrpB is not present in the soluble fraction, suggesting that it is likely in OMVs (unpaired *t* test). (D) Fluorescence intensity of HrpC-FLAG in supernatant fractions normalized to the Coomassie blue loading control. Each dot represents a distinct technical replicate pair. The TXISS-secreted HrpC is mostly located in the soluble fraction (Tukey's HSD test).

five available loci (Fig. S5 and Table S4). Aligning the maximum likelihood phylogenies reveals that the TXISS complement of a *Xenorhabdus* species is more predictive of nematode host than phylogenetic position (Fig. 6).

This alignment of the phylogenetic placement of a given *Steinernema* host with the TXISS complement of the symbiont provides insights into the nematode internal environment experienced by the symbiont. For example, *Xenorhabdus* with two HrpB paralogs at the *hrp* locus are symbionts of nematodes within the phylogenetic clade IV, suggesting that these nematodes present a distinctive environment in which a second *hrp* locus is beneficial. *Xenorhabdus* with an HrpB paralog encoded adjacent to genes predicted to encode a cobalt ABC transporter are solely symbionts of clade V nematodes. *Xenorhabdus innexi* and *Xenorhabdus stockiae* have seemingly diverged from this lineage through acquisition of a *nilB* homolog and switching to hosts within clade II. Similarly, *X. nematophila* independently gained *nilB* and switched into a clade II host (Fig. 6B). These acquisitions, alongside the varied genomic contexts of *nilB/nilC* pairs, are consistent with previous suggestions that the *nil* locus was horizontally acquired among *Xenorhabdus* (18).

The TXISS NilB and the lipoprotein NilC are necessary for *X. nematophila* to colonize the clade II nematode *S. carpocapsae* (14, 18). However, *X. nematophila* also colonizes two nematodes, *Steinernema anatoliense* and *Steinernema websteri*, which are phylogenetically separate from clade II (Fig. 6). Our hypothesis that TXISS are involved in host-environment adaptations leads to the prediction that *X. nematophila* will require the *nil* locus to colonize these nematodes. To test this hypothesis, bacterium-free *S. anatoliense*, *S. websteri*,
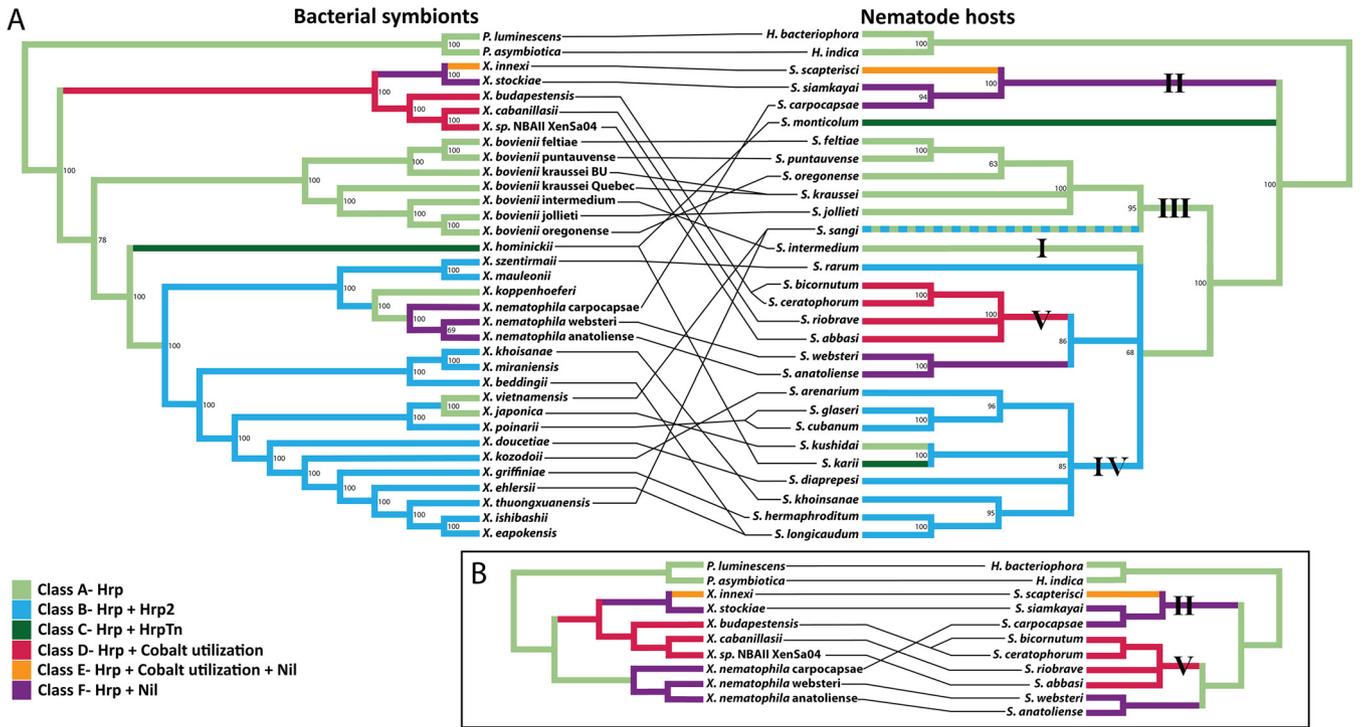
FIG 6 Cladograms of *Xenorhabdus* and *Steinernema* color coded according to *Xenorhabdus* DUF560 class. (A and B) Cophylogeny of nematode species and their colonizing bacteria across the *Steinernema* genus (A) or with a focus on specific clades (B). Numbers on branches indicate bootstrap support values. Bootstrap values below 60% were contracted. Lines connecting the phylogenies indicate mutualist pairs. Roman numerals highlight the five *Steinernema* clades described in reference 41. Colored overlays indicate the DUF560 class of a given bacterium or a given nematode's symbiont as follows: class A (light green), class B (light blue), class C (dark green), class D (red), class E (orange), and class F (purple).

and *S. carpocapsae* eggs (44) were exposed to an *X. nematophila* carpocapsae ATCC 19061 Δ*nil* mutant and a *nil*-complemented strain (15). Consistent with our prediction, the Δ*nil* mutant was deficient in infective juvenile colonization in all three nematode species (Fig. 7), demonstrating that *nil* genes are necessary for infective juvenile colonization of *S. anatoliense* and *S. websteri* and supporting our hypothesis that TXISS promote adaptation to host environments.

DISCUSSION

Bacteria have evolved specialized secretion systems for delivery of effectors that facilitate the host-associated lifestyle. Knowledge of cargo protein identities and sorting processes facilitates predictions from genomic information of bacterial secretome composition, regulation, and localization. Despite the diverse secretion systems now recognized, more remain to be identified based on the fact that some proteins predicted to be secreted lack known secretion pathways (45). The type X secretory pathway was described in 2020 shedding light on long-standing mysteries surrounding the dependence on a muramidase for secretion of typhoid toxin across the cell wall (7). Enabled by availability of genomic data from myriad environments and ever-improving bioinformatic visualization tools, we have presented data that the TXISS is a broadly distributed molecular vehicle for moving proteins across the Gram-negative outer membrane. Our network analysis has revealed functionally relevant TXISS clusters, with cluster 1 members having a conserved role in host-microbe interactions.

The DUF560 (domain of unknown function) family presence in animal-associated bacteria was first recognized when it was noted that the host colonization factor NilB has homologs in several human pathogens (14). This observation was strengthened by subsequent demonstration that Slam1 and Slam2 from these bacteria facilitate surface presentation of host metal acquisition proteins (8, 9, 14). Using the *Steinernema*-*Xenorhabdus* symbiosis, here we provide evidence that the composition of TXISS in a
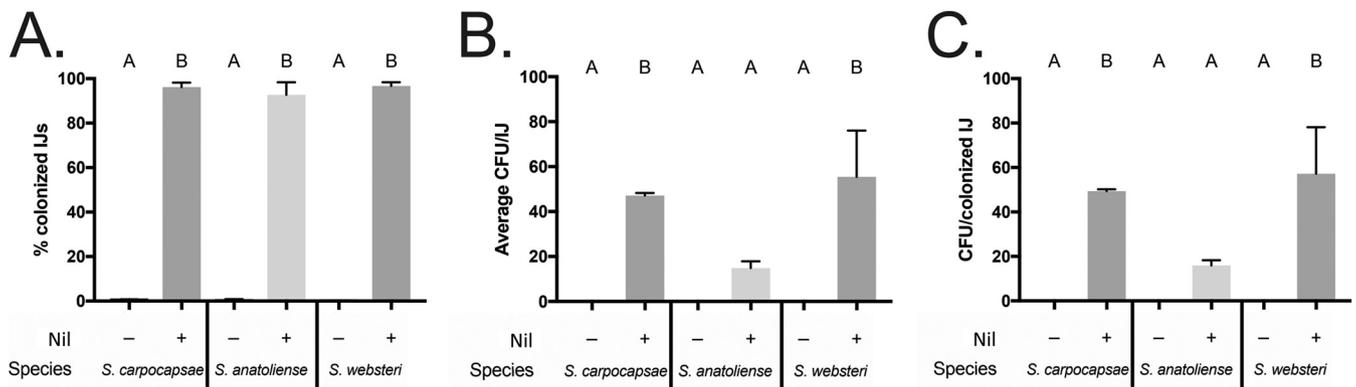
FIG 7 The *nil* locus is necessary for colonization of *S. anatoliense*, *S. carpocapsae*, and *S. websteri* infective juveniles. Bacterium-free *S. carpocapsae*, *S. anatoliense*, and *S. websteri* were exposed to GFP-expressing *X. nematophila* lacking or bearing the *nil* locus. (A and B) The resulting progeny infective juveniles (IJs) were monitored for colonization either by microscopy (A) or by plating lysates for average CFU/IJ. (B) (C) The average CFU per colonized IJ combines both of these values to show bacterial load per organism. Treatments were analyzed via one-way ANOVA and Tukey's *post hoc* test.

bacterial symbiont genome correlates with host organism. During the evolutionary history of *Xenorhabdus*, the gain or loss of TXISS loci correlated with host switching events (Fig. 6). These data indicate that TXISS activity contributes to bacterial adaptation to new host environments, which is particularly relevant given the varied distribution of TXISS homologs among human pathogens. For instance, our work enables categorization of TXISS among animal-associated *Neisseria* species, including the human pathogens *Neisseria meningitidis* and *Neisseria gonorrhoeae* (46). *Neisseria* strains can encode up to 6 TXISS paralogs. *N. meningitidis* MC58 has two functionally characterized TXISS, Slam1 and Slam2. Our network analysis indicates *N. meningitidis* encodes a third Slam, NMB1466/NP_274965, that also falls within cluster 1A. *N. gonorrhoeae* TXISS are represented in more nodes than *N. meningitidis*, notably occupying 19.3% of nodes in subcluster 1B (Fig. 3). The cluster framework we present here indicates that *Neisseria* use TXISS to secrete both lipoproteins and soluble cargo and that variation in TXISS composition among *Neisseria* strains may be predictive of host or tissue association phenotypes.

While DUF560 proteins were originally thought to represent a mechanism for lipoprotein surface exposure (8), recent studies have expanded that functional range to include peripheral membrane proteins (11) and now soluble secreted proteins, establishing the function of DUF560 outer membrane proteins in secretion of varied substrates. We predict that further study will uncover even more chemically diverse cargo for the distinct classes of TXISS, revealed through our network analysis. Here we focused on a single cluster of a DUF560 sequence similarity network (SSN). The remaining nine clusters likely represent diverse subfamilies responsible for transporting as-yet-unknown cargo. The lipoprotein cargo proteins for which structural conformations are known, share a C-terminal eight-stranded β-barrel (9). Our discovery that HrpC is a cargo protein for TXISS-1B HrpB strengthens the concept that this barrel is an important characteristic of TXISS cargo; HrpC is a homolog of the *H. haemolyticus* hemophilin, the structure of which likewise adopts a C-terminal eight-stranded β-barrel (29) (see Fig. S2 in the supplemental material). These data support the concept that TXISS cargo have bifunctional structures in which the N terminus is the host effector domain, while the C terminus targets secretion. This framework will facilitate identification of as-yet-unknown TXISS cargo among the genes that co-occur with TXISS outer membrane proteins. It is our hope that the DUF560-TbpBBD co-occurrence table used for our network annotation will become a resource for other microbiologists studying secreted proteins (see Table S2 in the supplemental material). This database identifies 851 potential TXISS/cargo pairs spread across 463 bacterial isolates, only 4 of which have been experimentally investigated thus far. Furthermore, our analyses suggest that SSN clusters have predictive power for other characteristics of TXISS cargo, including whether they are surface attached or secreted. The network-enabled classification

presented here will facilitate the investigation of both TXISS outer membrane proteins (OMPs) and their cargo in diverse bacteria.

Slam1 and Slam2 TXISS-dependent cargo with known molecular function include those that serve as coreceptors for TonB-dependent metal uptake systems (2). Similarly, HrpC, which we show here is a TXISS-dependent cargo protein, and its homologs also likely function in metal acquisition by acting as hemophores akin to hemophilin and HasA (29, 47, 48). This idea is supported by the fact that an *N. gonorrhoeae hrpC* homolog (NGO0554) is repressed by iron, upregulated in response to oxidative stress, and contributes to resistance to peroxide challenge (49–51). Further, the *hrp* locus, which is conserved across proteobacteria, is predicted to encode TonB and a TonB-dependent receptor (Fig. S4). Our working model is that the metal bound by secreted HrpC is passed to its respective TonB-dependent receptor and imported into the cell through TonB energization (2). Given the conservation of the TXISS *hrp* locus among all *Xenorhabdus* and throughout human microbiome constituents, it will be important in the future to examine the regulation of HrpC TXISS-dependent secretion and the roles of *hrp* machinery in binding and acquiring host metals in a mucosal environment.

In addition to their long-recognized roles in import, TonB and TonB-dependent receptors may also function in protein export. The *Myxococcus xanthus* TonB-dependent transporter Oar can export the protease PopC using the proton motive force to energize membrane translocation (52). Despite this expansion of the known functional range of TonB-dependent transporters, in the case of the TXISS clusters, TonB is likely responsible for energizing uptake systems, but not TXISS-mediated export. This suggestion is supported by the fact that both the TXISS-1A function of Slam-1 (8) and the TXISS-1B function of HrpB shown here could be reconstructed in *E. coli* without coexpression of native TonB. However, it will be interesting for future studies to investigate possible functional interactions between TonB-dependent transporter-mediated secretion and TXISS-mediated secretion.

Our initial network generated from the entire PF04575 (DUF560) data set included cluster 5, comprising 14 nodes of *Klebsiella* homologs and 2 nodes with *Klebsiella* and *Escherichia* homologs annotated as PgaA/HmsH. This cluster was removed based on its limited number of nodes, predicted topology differences relative to the rest of the network (16-stranded versus 14-stranded barrel), and the fact that the majority of known PgaA homologs were not represented within the cluster. PgaA is a component of one of several machineries for secretion of exopolysaccharides (EPSs) that comprise biofilm matrices of Gram-negative bacteria (53–55). Despite the topological and substrate (polysaccharide versus protein) differences, the PF04575 assignment of cluster 5 PgaA homologs hints that there could be evolutionary or structural parallels between exopolysaccharide synthase-dependent secretion systems and TXISS. For instance, both TXISS and EPS secretion machineries either have, or associate with proteins that have, tetratricopeptide repeat (TPR) domains (53). In EPS secretion systems, the TPR are necessary for secretion or for cargo modification (54, 56). TXISS TPR domains may similarly modulate activities of other proteins that influence secretion of TXISS substrates. Support for a role of TPR domains in TXISS activity comes from the fact that *X. nematophila* expressing versions of the NilB TPR domain with small deletions display colonization defects that are ameliorated by deletion of the entire N-terminal periplasmic domain (15). Our establishment of TXISS as a bona fide secretion system opens new avenues for exploration of its integration with other export and secretion machineries and its coordinated contributions to host-associated phenotypes, including metal homeostasis, aggregation, and biofilm formation.

MATERIALS AND METHODS

All primers and strains are listed in Table S5 in the supplemental material.

DUF560 SSN analysis. The EFI Enzyme Similarity Tool (EFI-EST) was used to collect all predicted DUF560 domain-containing sequences from Interpro 73 and UniProt 2019-02 (accessed 24 April 2019) and compare using BLAST (21, 22, 25). Representative networks collapsed nodes which shared ≥40%

identity. On an EFI-EST network, edges are drawn according to a database-independent alignment score. Greater scores correspond to fewer edges. For separation of DUF560 clusters, an alignment score of 38 was chosen (Fig. 2). For subclusters 1A and 1B, 89 and 100 were chosen, respectively (see Fig. S1 in the supplemental material). Networks were visualized and interpreted using Cytoscape v3.7.1 (57) and Gephi v0.9.2 (58). Nodes were arranged with the Fruchterman-Reingold force-directed layout *algorithm* (59).

The contents of each cluster were compared to Pfam DUF560 (PF04575) to ensure that clusters contained legitimate DUF560 proteins (26). Any clusters for which fewer than 18% of sequences were present in Pfam, or with fewer than 20 sequences, were excluded. This filtering removed cluster 5, composed mostly of *Klebsiella pneumoniae* PgaA. The resulting SSN contains 10 clusters, 1,222 nodes, and 52,190 edges with 1,589 TaxIDs represented (Fig. 2). Using NCBI Taxonomy Browser, each node was manually curated for the isolate's environmental origin(s) among the following categories: water, soil, plant, mammal, animal, invertebrate, nematode, built (environments such as sewage, bioreactors, etc.), multiple environments, and unclassified (Table S1.). If no citation was available, the source was curated from BioSample or BioProject records or if neither was available, from other resources (NCBI linkout, Google search). Animal-associated microbes were designated as mammal, insect, nematode, or generic animal associated. Our fine-scale analysis focused on cluster 1 and its subclusters (Fig. 3 and Fig. S1).

Three different techniques were used to determine whether DUF560 proteins within our network were genomically associated with TbpBBD domain-containing proteins. First, using EFI-GNN, genome-neighborhood-networks were generated for subcluster 1A (alignment score, 89; 40 total flanking ORFs, 20 upstream and 20 downstream of the DUF560), subcluster 1B (alignment score, 38; 20 total flanking ORFs, 10 upstream and 10 downstream of the DUF560), and subcluster 1C (alignment score, 38; 20 total flanking ORFs, 10 upstream and 10 downstream of the DUF560), resulting in 352 DUF560-TbpBBD pairs (22). Next, the RODEO web tool was used to analyze cooccurrence using profile hidden Markov models to assign domains to local ORFs resulting in 712 DUF560-TbpBBD pairs (27). Seven additional DUF560-TbpBBD pairs in *Xenorhabdus* were manually annotated. All three data sets were combined for a total of 851 nonredundant protein pairs (Table S2). SignalP-5 (60, 61) was used to predict the signal peptides of all TbpBBD domain-containing proteins and annotate each node (Fig. 3C). Any node associated with both signal peptide-bearing and signal peptide-free proteins was annotated according to those with a signal.

DUF560 genome neighborhood analysis. Subclusters 1A to C were separated and analyzed in EFI-EST with an alignment score of 38 as described above. Each network was then analyzed through EFI-GNN and visualized in Cytoscape v3.7.1 (57) (alignment score, 38; 10 ORFs up- and downstream).

Phyre^2 analysis. After removal of signal peptide sequences, TbpBBD$_{sol}$ proteins *X. nematophila* HrpC, *P. rettgeri* PROVETT_08181 and PROVETT_05852, and *Proteus mirabilis* WP_134940027.1 were submitted to Phyre^2 Protein Homology/analogy Recognition Engine v.2.0 to predict potential three-dimensional (3-D) structures (32). The top predicted match for each, hemophilin (PDB accession no. 6OM5), was used to generate structural models (Fig. S2).

Expression of HrpB and HrpC. The gene encoding HrpB25_26insDYKDDDDK (FLAG-HrpB) was synthesized and cloned into the second multiple cloning site (MCS) of pETDuet-1 by Genscript. The genomic region containing HrpC was amplified from the *X. nematophila* ATCC 19061 (HGB800) genome using primers 1 and 2, digested with restriction enzymes SacI and SalI, and ligated into the pUC19 MCS. Site-directed mutagenesis was used to add a C-terminal FLAG tag onto HrpC using primers 3 and 4. The gene encoding HrpC1_2insV246_247insDYKDDDDK (henceforth HrpC-FLAG) was amplified from pUC19 using primers 5 and 6. The product was ligated into MCS1 of both pETDuet-1 and pETDuet-1/FLAG-HrpB, resulting in expression plasmids pETDuet-1/HrpC-FLAG and pETDuet-1/HrpC-FLAG/FLAG-HrpB. HrpC-FLAG was used to create HrpC1_2insV246_247ins3x(GGGGS)-GST domain (HrpC-GST) using Hi-Fi assembly and primer pair 7 and 8 and primer pair 9 and 10 to integrate a 15-amino-acid linker and GST into the previous FLAG locus. This process yielded pETDuet-1/HrpC-GST and pETDuet-1/HrpC-GST/FLAG-HrpB. All clones were confirmed by Sanger sequencing using primers 11 and 14 at the Univeriity of Tennessee (UT) Genomics Core.

Expression plasmids were transformed into *E. coli* B21(DE3) and *E. coli* C43(DE3) via electroporation. Results were similar for the two strains. Strains were grown in defined medium with 150 μg/ml ampicillin (62). Bacteria were subcultured into 100 ml of broth at an initial optical density at 600 nm (OD$_{600}$) of 0.028, grown for 18 h at 37°C to reach late logarithmic growth, and induced with 500 μM isopropyl-β-D-thiogalactopyranoside (IPTG). After 1 h, 700 μl of each culture was filtered for subsequent use. At 2.5 h, whole cells were collected by centrifugation and lysed using a bead beater. Remaining supernatants were filtered. Protein concentration of whole-cell lysates was measured via Bradford assay. For supernatant samples, 700 μl of supernatant was precipitated via 10% trichloroacetic acid (TCA) (63).

Samples were analyzed by 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting. For lysates, wells were loaded with 9.5 μg of protein. For supernatants, wells were loaded with half the TCA precipitate collected media. Western blots were probed with either rat anti-FLAG or rabbit anti-GST primary antibody anti-IgG secondary antibody that fluoresces at 680 nm. Intensities were recorded for FLAG or GST reactive bands which were the correct size for mature HrpB and HrpC. For every supernatant sample, a distinct band from the Coomassie blue-stained gel was used as a loading control to normalize intensities of supernatant samples prior to analysis (Fig. S3). Unpaired *t* tests were used to compare HrpC-FLAG to HrpC-GST secretion and supernatant fraction HrpB. One replicate of the HrpC-GST 1-h supernatant was excluded from analysis

due to cellular lysis. A Tukey's honestly significant difference (HSD) test was used for comparing supernatant fraction HrpC.

For additional analysis of OMVs, bacteria producing each expression plasmid were subcultured in triplicate into 40 ml of broth at an initial OD_{600} of 0.04. After 5 h at 37°C to reach mid-logarithmic growth, cultures were induced with 500 μM IPTG. After an additional 2.5 h, each culture was clarified via centrifugation, then the supernatant was filtered and centrifuged at 150,000 relative centrifugal force (RCF) for 3 h. The entire OMV pellet was solubilized in SDS sample buffer. Six hundred microliters of the total supernatant or the clarified (postultracentrifugation) supernatant were precipitated via 10% TCA and resuspended in SDS sample buffer (63). Samples were analyzed by SDS-PAGE and Western blotting as described above.

Phylogenetic tree generation. Phylogenetic analysis was performed as described previously (64). Briefly, select *Xenorhabdus* and *Photorhabdus* species were analyzed using MicroScope MaGe's Gene Phyloprofile tool (37, 42) to identify homologous protein sets which were conserved across all assayed genomes. Putative paralogs were excluded from downstream analysis to ensure homolog relatedness, resulting in 665 homologous sets (Table S5). Homolog sets were retrieved via locus tag indexing using BioPython (65), individually aligned using Muscle v3.8.31 (66), concatenated using Sequence Matrix v1.8 (67), and trimmed of nucleotide gaps using TrimAL v1.3 (68). JmodelTest v2.1.10 (69) was used to choose the GTR+γ substitution model for maximum likelihood and Bayesian analysis.

For nematode phylogenetic analysis, select *Steinernema* and *Heterorhabditis* species were analyzed. The internal transcribed spacer, 18S rRNA, 28S rRNA, cytochrome oxidase I, and 12S rRNA loci were collected from GenBank and used as homologous sets (Table S4). Nematode species with fewer than three of five loci sequenced were excluded. Homologous sets were processed using the same methods as the *Xenorhabdus* sequences, although for these sequences Modettes v2.1.10 (70) chose the GTR+γ+I substitution model.

Maximum likelihood analyses were performed via RAxML v8.2.10 (70) using rapid bootstrapping and 1,000 replicates and were visualized via Dendroscope v3.6.2 (71). Nodes with less than 60% bootstrap support were collapsed. Bayesian analyses were performed via MrBayes v3.2.6 with BEAGLE (72–74) on the CIPRES Science Gateway platform (75). A total of 500,000 or 4,000,000 Markov chain Monte Carlo (MCMC) replicates were performed for the bacterial or nematode tree, respectively. Twenty-five percent were discarded as burn-in, and posterior probabilities were sampled every 500 replicates. Two runs were performed with three heated chains and one cold chain. The final standard deviation of split frequencies was less than 10^{-6} for the bacterial tree, and it was 0.002557 for the nematode tree. Bayesian trees were visualized with FigTree v1.4.4 (76). Bayesian and maximum likelihood methods generated phylogenies with consistent topologies (Fig. 6 and Fig. S5).

Wild type versus Δ*nil* colonization of nematodes. Bacterium-free eggs of *S. anatoliense*, *S. carpocapsae*, and *S. websteri* (44) exposed to a green fluorescent protein (GFP)-expressing Nil mutant (HGB1495) or its complemented strain (HGB1496) were grown on lipid agar plates for 2 days at 25°C (15). Plates were placed into White traps 1 week after plating to collect infective juvenile (IJ) nematodes. Nematode colonization was visualized using fluorescence microscopy on a Keyence BZX-700 to observe bacteria in the receptacle (in biological triplicate and technical duplicate). To determine the number of CFU per IJ, nematodes were surface sterilized and ground for 2 min using a Fisher brand motorized tissue grinder (catalog no. 12-1413-61) to homogenize the nematodes and release colonizing bacteria. Serial dilutions in phosphate-buffered saline (PBS) were plated on LB agar and then incubated at 30°C for 1 day before enumerating CFU (Fig. 7). To calculate the CFU per colonized IJ, the percent colonized nematodes was divided by the CFU/IJ for each biological replicate. The data were analyzed using a one-way analysis of variance (ANOVA) with Tukey's multiple-comparison test to compare the mean of each treatment.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, PDF file, 0.4 MB.
FIG S2, PDF file, 0.2 MB.
FIG S3, PDF file, 2.7 MB.
FIG S4, PDF file, 0.2 MB.
FIG S5, PDF file, 0.2 MB.
TABLE S1, XLSX file, 0.4 MB.
TABLE S2, XLSX file, 0.2 MB.
TABLE S3, XLSX file, 0.4 MB.
TABLE S4, XLSX file, 0.2 MB.
TABLE S5, XLSX file, 0.01 MB.

T.J.M., A.S.G., and H.G.-B. wrote the text article and composed the figures. Bioinformatic analysis was performed by A.S.G. and T.J.M. Phylogenetic analysis was performed by A.S.G. Cloning and experiments were performed by A.S.G. and T.J.M. K.T.F. provided intellectual contributions to the research design and edits to the text and figures.

We have no competing interests that might be perceived to influence the results and/or discussion reported in this paper.

REFERENCES

1. Chu BC, Garcia-Herrero A, Johanson TH, Krewulak KD, Lau CK, Peacock RS, Slavinskaya Z, Vogel HJ. 2010. Siderophore uptake in bacteria and the battle for iron with the host; a bird's eye view. Biometals 23:601–611. https://doi.org/10.1007/s10534-010-9361-x.

2. Noinaj N, Guillier M, Barnard TJ, Buchanan SK. 2010. TonB-dependent transporters: regulation, structure, and function. Annu Rev Microbiol 64:43–60. https://doi.org/10.1146/annurev.micro.112408.134247.

3. Schweppe DK, Harding C, Chavez JD, Wu X, Ramage E, Singh PK, Manoil C, Bruce JE. 2015. Host-microbe protein interactions during bacterial infection. Chem Biol 22:1521–1530. https://doi.org/10.1016/j.chembiol.2015.09.015.

4. Chavez-Dozal AA, Gorman C, Lostroh CP, Nishiguchi MK. 2014. Gene-swapping mediates host specificity among symbiotic bacteria in a beneficial symbiosis. PLoS One 9:e101691. https://doi.org/10.1371/journal.pone.0101691.

5. Green ER, Mecsas J. 2016. Bacterial secretion systems: an overview. Microbiol Spectr 4:10.1128/microbiolspec.VMBF-0012-2015. https://doi.org/10.1128/microbiolspec.VMBF-0012-2015.

6. Lasica AM, Ksiazek M, Madej M, Potempa J. 2017. The type IX secretion system (T9SS): highlights and recent insights into its structure and function. Front Cell Infect Microbiol 7:215. https://doi.org/10.3389/fcimb.2017.00215.

7. Palmer T, Finney AJ, Saha CK, Atkinson GC, Sargent F. 2021. A holin/peptidoglycan hydrolase-dependent protein secretion system. Mol Microbiol 115:345–355. https://doi.org/10.1111/mmi.14599.

8. Hooda Y, Lai CC, Judd A, Buckwalter CM, Shin HE, Gray-Owen SD, Moraes TF. 2016. Slam is an outer membrane protein that is required for the surface display of lipidated virulence factors in Neisseria. Nat Microbiol 1:16009. https://doi.org/10.1038/nmicrobiol.2016.9.

9. Hooda Y, Lai CCL, Moraes TF. 2017. Identification of a large family of slam-dependent surface lipoproteins in Gram-negative bacteria. Front Cell Infect Microbiol 7:207. https://doi.org/10.3389/fcimb.2017.00207.

10. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, Finn RD, Bateman A. 2021. Pfam: the protein families database in 2021. Nucleic Acids Res 49:D412–D419. https://doi.org/10.1093/nar/gkaa913.

11. da Silva RAG, Karlyshev AV, Oldfield NJ, Wooldridge KG, Bayliss CD, Ryan A, Griffin R. 2019. Variant signal peptides of vaccine antigen, FHbp, impair processing affecting surface localization and antibody-mediated killing in most meningococcal isolates. Front Microbiol 10:2847. https://doi.org/10.3389/fmicb.2019.02847.

12. Akhurst RJ. 1983. Neoplectana species: specificity of association with bacteria of the genus Xenorhabdus. Exp Parasitol 55:258–263. https://doi.org/10.1016/0014-4894(83)90020-6.

13. Chaston JM, Suen G, Tucker SL, Andersen AW, Bhasin A, Bode E, Bode HB, Brachmann AO, Cowles CE, Cowles KN, Darby C, de Léon L, Drace K, Du Z, Givaudan A, Herbert Tran EE, Jewell KA, Knack JJ, Krasomil-Osterfeld KC, Kukor R, Lanois A, Latreille P, Leimgruber NK, Lipke CM, Liu R, Lu X, Martens EC, Marri PR, Médigue C, Menard ML, Miller NM, Morales-Soto N, Norton S, Ogier J-C, Orchard SS, Park D, Park Y, Qurollo BA, Sugar DR, Richards GR, Rouy Z, Slominski B, Slominski K, Snyder H, Tjaden BC, van der Hoeven R, Welch RD, Wheeler C, Xiang B, Barbazuk B, et al. 2011. The entomopathogenic bacterial endosymbionts Xenorhabdus and Photorhabdus: convergent lifestyles from divergent genomes. PLoS One 6:e27909. https://doi.org/10.1371/journal.pone.0027909.

14. Heungens K, Cowles CE, Goodrich-Blair H. 2002. Identification of Xenorhabdus nematophila genes required for mutualistic colonization of Steinernema carpocapsae nematodes. Mol Microbiol 45:1337–1353. https://doi.org/10.1046/j.1365-2958.2002.03100.x.

15. Bhasin A, Chaston JM, Goodrich-Blair H. 2012. Mutational analyses reveal overall topology and functional regions of NilB, a bacterial outer membrane protein required for host association in a model of animal-microbe mutualism. J Bacteriol 194:1763–1776. https://doi.org/10.1128/JB.06711-11.

16. Blatch GL, Lassle M. 1999. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. Bioessays 21:932–939. https://doi.org/10.1002/(SICI)1521-1878(199911)21:11<932::AID-BIES5>3.0.CO;2-N.

17. Cowles CE, Goodrich-Blair H. 2004. Characterization of a lipoprotein, NilC, required by Xenorhabdus nematophila for mutualism with its nematode host. Mol Microbiol 54:464–477. https://doi.org/10.1111/j.1365-2958.2004.04271.x.

18. Cowles CE, Goodrich-Blair H. 2008. The Xenorhabdus nematophila nilABC genes confer the ability of Xenorhabdus spp. to colonize Steinernema carpocapsae nematodes. J Bacteriol 190:4121–4128. https://doi.org/10.1128/JB.00123-08.

19. d'Enfert C, Chapon C, Pugsley AP. 1987. Export and secretion of the lipoprotein pullulanase by Klebsiella pneumoniae. Mol Microbiol 1:107–116. https://doi.org/10.1111/j.1365-2958.1987.tb00534.x.

20. Grossman AS, Mauer TJ, Forest KT, Goodrich-Blair H. 19 February 2021. A widespread bacterial secretion system with chemically diverse protein substrates. bioRxiv https://doi.org/10.1101/2020.01.20.912956.

21. Gerlt JA, Bouvier JT, Davidson DB, Imker HJ, Sadkhin B, Slater DR, Whalen KL. 2015. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): a web tool for generating protein sequence similarity networks. Biochim Biophys Acta 1854:1019–1037. https://doi.org/10.1016/j.bbapap.2015.04.015.

22. Zallot R, Oberg NO, Gerlt JA. 2018. 'Democratized' genomic enzymology web tools for functional assignment. Curr Opin Chem Biol 47:77–85. https://doi.org/10.1016/j.cbpa.2018.09.009.

23. Zallot R, Oberg N, Gerlt JA. 2019. The EFI web resource for genomic enzymology tools: leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. Biochemistry 58:4169–4182. https://doi.org/10.1021/acs.biochem.9b00735.

24. Ostberg KL, DeRocco AJ, Mistry SD, Dickinson MK, Cornelissen CN. 2013. Conserved regions of gonococcal TbpB are critical for surface exposure and transferrin iron utilization. Infect Immun 81:3442–3450. https://doi.org/10.1128/IAI.00280-13.

25. Gerlt JA. 2017. Genomic enzymology: web tools for leveraging protein family sequence–function space and genome context to discover novel functions. Biochemistry 56:4293–4308. https://doi.org/10.1021/acs.biochem.7b00614.

26. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE, Finn RD. 2019. The Pfam protein families database in 2019. Nucleic Acids Res 47:D427–D432. https://doi.org/10.1093/nar/gky995.

27. Choe K, Mitchell Laboratory. 2018. RODEO (Rapid ORF Description & Evaluation Online) is an algorithm to help biosynthetic gene cluster (BGC) analysis, with an emphasis on ribosomal natural product (RiPP) discovery. University of Illinois at Urbana-Champaign. http://ripp.rodeo/advanced.html.

28. Schmoldt A, Benthe HF, Haberland G. 1975. Signal peptide hydrophobicity modulates interaction with the twin-arginine translocase. Biochem Pharmacol 24:1639–1641. https://doi.org/10.1128/mBio.00909-17.

29. Latham RD, Torrado M, Atto B, Walshe JL, Wilson R, Guss JM, Mackay JP, Tristram S, Gell DA. 2020. A heme-binding protein produced by Haemophilus haemolyticus inhibits non-typeable Haemophilus influenzae. Mol Microbiol 113:381–398. https://doi.org/10.1111/mmi.14426.

30. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389.

31. Kelley LA, Sternberg MJ. 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc 4:363–371. https://doi.org/10.1038/nprot.2009.2.

32. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10:845–858. https://doi.org/10.1038/nprot.2015.053.

33. Einhauer A, Jungbauer A. 2001. The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins. J Biochem Biophys Methods 49:455–465. https://doi.org/10.1016/s0165-022x(01)00213-5.

34. Fantappie L, Irene C, De Santis M, Armini A, Gagliardi A, Tomasi M, Parri M, Cafardi V, Bonomi S, Ganfini L, Zerbini F, Zanella I, Carnemolla C, Bini L, Grandi A, Grandi G. 2017. Some Gram-negative lipoproteins keep their surface topology when transplanted from one species to another and deliver foreign polypeptides to the bacterial surface. Mol Cell Proteomics 16:1348–1364. https://doi.org/10.1074/mcp.M116.065094.

35. Konar M, Rossi R, Walter H, Pajon R, Beernink PT. 2015. A mutant library approach to identify improved meningococcal factor H binding protein vaccine antigens. PLoS One 10:e0128185. https://doi.org/10.1371/journal.pone.0128185.

36. Kohl P, Zingl FG, Eichmann TO, Schild S. 2018. Isolation of outer membrane vesicles including their quantitative and qualitative analyses. Methods Mol Biol 1839:117–134. https://doi.org/10.1007/978-1-4939-8685-9_11.

37. Chaston JM, Murfin KE, Heath-Heckman EA, Goodrich-Blair H. 2013. Previously unrecognized stages of species-specific colonization in the mutualism between Xenorhabdus bacteria and Steinernema nematodes. Cell Microbiol 15:1545–1559. https://doi.org/10.1111/cmi.12134.

38. Kampfer P, Tobias NJ, Ke LP, Bode HB, Glaeser SP. 2017. Xenorhabdus thuongxuanensis sp. nov. and Xenorhabdus eapokensis sp. nov., isolated from Steinernema species. Int J Syst Evol Microbiol 67:1107–1114. https://doi.org/10.1099/ijsem.0.001770.

39. Lee MM, Stock SP. 2010. A multigene approach for assessing evolutionary relationships of Xenorhabdus spp. (gamma-Proteobacteria), the bacterial symbionts of entomopathogenic Steinernema nematodes. J Invertebr Pathol 104:67–74. https://doi.org/10.1016/j.jip.2010.01.005.

40. Snyder HA, Stock SP, Kim SK, Flores-Lara Y, Forst S. 2007. New insights into the colonization and release process of Xenorhabdus nematophila and the morphology and ultrastructure of the bacterial receptacle of its nematode host, Steinernema carpocapsae. Appl Environ Microbiol 73:5338–5346. https://doi.org/10.1128/AEM.02947-06.

41. Spiridonov SE, Reid AP, Podrucka K, Subbotin SA, Moens M. 2004. Phylogenetic relationships within the genus Steinernema (Nematoda: Rhabditida) as inferred from analyses of sequences of the ITS1-5.8S-ITS2 region of rDNA and morphological features. Nematology 6:547–566. https://doi.org/10.1163/1568541042665304.

42. Vallenet D, Calteau A, Cruveiller S, Gachet M, Lajus A, Josso A, Mercier J, Renaux A, Rollin J, Rouy Z, Roche D, Scarpelli C, Medigue C. 2017. MicroScope in 2017: an expanding and evolving integrated resource for community expertise of microbial genomes. Nucleic Acids Res 45:D517–D528. https://doi.org/10.1093/nar/gkw1101.

43. McGinnis S, Madden TL. 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucleic Acids Res 32:W20–W25. https://doi.org/10.1093/nar/gkh435.

44. Murfin KE, Chaston J, Goodrich-Blair H. 2012. Visualizing bacteria in nematodes using fluorescence microscopy. J Vis Exp 68:e4298. https://doi.org/10.3791/4298.

45. Bendtsen JD, Kiemer L, Fausboll A, Brunak S. 2005. Non-classical protein secretion in bacteria. BMC Microbiol 5:58. https://doi.org/10.1186/1471-2180-5-58.

46. Cohen J, Powderly W, Opal S. 2017. Preface to the fourth edition, p xiv. In Infectious diseases, 4th ed. Elsevier, New York, NY.

47. Dent AT, Mouriño S, Huang W, Wilks A. 2019. Post-transcriptional regulation of the Pseudomonas aeruginosa heme assimilation system (Has) fine-tunes extracellular heme sensing. J Biol Chem 294:2771–2785. https://doi.org/10.1074/jbc.RA118.006185.

48. Yukl ET, Jepkorir G, Alontaga AY, Pautsch L, Rodriguez JC, Rivera M, Moënne-Loccoz P. 2010. Kinetic and spectroscopic studies of hemin acquisition in the hemophore HasAp from Pseudomonas aeruginosa. Biochemistry 49:6646–6654. https://doi.org/10.1021/bi100692f.

49. Jackson LA, Ducey TF, Day MW, Zaitshik JB, Orvis J, Dyer DW. 2010. Transcriptional and functional analysis of the Neisseria gonorrhoeae Fur regulon. J Bacteriol 192:77–85. https://doi.org/10.1128/JB.00741-09.

50. Quillin SJ, Hockenberry AJ, Jewett MC, Seiferta HS. 2018. Neisseria gonorrhoeae exposed to sublethal levels of hydrogen peroxide mounts a complex transcriptional response. mSystems 3:e00156-18. https://doi.org/10.1128/mSystems.00156-18.

51. Stohl EA, Criss AK, Seifert HS. 2005. The transcriptome response of Neisseria gonorrhoeae to hydrogen peroxide reveals genes with previously uncharacterized roles in oxidative damage protection. Mol Microbiol 58:520–532. https://doi.org/10.1111/j.1365-2958.2005.04839.x.

52. Gómez-Santos N, Glatter T, Koebnik R, Świątek-Połatyńska MA, Søgaard-Andersen L. 2019. A TonB-dependent transporter is required for secretion of protease PopC across the bacterial outer membrane. Nat Commun 10:1360. https://doi.org/10.1038/s41467-019-09366-9.

53. Low KE, Howell PL. 2018. Gram-negative synthase-dependent exopolysaccharide biosynthetic machines. Curr Opin Struct Biol 53:32–44. https://doi.org/10.1016/j.sbi.2018.05.001.

54. Wang Y, Andole Pannuri A, Ni D, Zhou H, Cao X, Lu X, Romeo T, Huang Y. 2016. Structural basis for translocation of a biofilm-supporting exopolysaccharide across the bacterial outer membrane. J Biol Chem 291:10046–10057. https://doi.org/10.1074/jbc.M115.711762.

55. Whitney JC, Howell PL. 2013. Synthase-dependent exopolysaccharide secretion in Gram-negative bacteria. Trends Microbiol 21:63–72. https://doi.org/10.1016/j.tim.2012.10.001.

56. Marmont LS, Whitfield GB, Rich JD, Yip P, Giesbrecht LB, Stremick CA, Whitney JC, Parsek MR, Harrison JJ, Howell PL. 2017. PelA and PelB proteins form a modification and secretion complex essential for Pel polysaccharide-dependent biofilm formation in Pseudomonas aeruginosa. J Biol Chem 292:19411–19422. https://doi.org/10.1074/jbc.M117.812842.

57. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504. https://doi.org/10.1101/gr.1239303.

58. Bastian M, Heymann S, Jacomy M. 2009. Gephi: an open source software for exploring and manipulating networks, p 361–362. In Proceedings of the Third International AAAI Conference on Weblogs and Social Media. Association for the Advancement of Artificial Intelligence, Menlo Park, CA.

59. Fruchterman TMJ, Reingold EM. 1991. Graph drawing by force-directed placement. Softw Pract Exp 21:1129–1164. https://doi.org/10.1002/spe.4380211102.

60. Almagro Armenteros JJ, Tsirigos KD, Sonderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol 37:420–423. https://doi.org/10.1038/s41587-019-0036-z.

61. Nielsen H, Engelbrecht J, Brunak S, von Heijne G. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10:1–6. https://doi.org/10.1093/protein/10.1.1.

62. Orchard SS, Goodrich-Blair H. 2004. Identification and functional characterization of a Xenorhabdus nematophila oligopeptide permease. Appl Environ Microbiol 70:5621–5627. https://doi.org/10.1128/AEM.70.9.5621-5627.2004.

63. Koontz L. 2014. TCA precipitation. Methods Enzymol 541:3–10. https://doi.org/10.1016/B978-0-12-420119-4.00001-X.

64. Murfin KE, Lee MM, Klassen JL, McDonald BR, Larget B, Forst S, Stock SP, Currie CR, Goodrich-Blair H. 2015. Xenorhabdus bovienii strain diversity impacts coevolution and symbiotic maintenance with Steinernema spp. nematode hosts. mBio 6:e00076. https://doi.org/10.1128/mBio.00076-15.

65. Cock PJ, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B, de Hoon MJ. 2009. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics 25:1422–1423. https://doi.org/10.1093/bioinformatics/btp163.

66. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340.

67. Vaidya G, Lohman DJ, Meier R. 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27:171–180. https://doi.org/10.1111/j.1096-0031.2010.00329.x.

68. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

69. Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9:772. https://doi.org/10.1038/nmeth.2109.

70. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

71. Huson DH, Scornavacca C. 2012. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. Syst Biol 61:1061–1067. https://doi.org/10.1093/sysbio/sys062.

72. Altekar G, Dwarkadas S, Huelsenbeck JP, Ronquist F. 2004. Parallel metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics 20:407–415. https://doi.org/10.1093/bioinformatics/btg427.

73. Ayres DL, Darling A, Zwickl DJ, Beerli P, Holder MT, Lewis PO, Huelsenbeck JO, Ronquist F, Swofford DL, Cummings MP, Rambaut A, Suchard MA. 2012. BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. Syst Biol 61:170–173. https://doi.org/10.1093/sysbio/syr100.

74. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539–542. https://doi.org/10.1093/sysbio/sys029.

75. Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees, p 1–8. *In* Proceedings of the Gateway Computing Environments Workshop (GCE). Institute of Electrical and Electronics Engineers (IEEE), Piscataway, NJ.

76. Rambaut A. 2018. FigTree v1.4.4. http://tree.bio.ed.ac.uk/software/figtree/.