December 2016

Molecular and Functional Analysis of the Pixb Gene in Xenorhabdus Nematophila

John Lucas
University of Wisconsin-Milwaukee

MOLECULAR AND FUNCTIONAL ANALYSIS OF THE *pixB* GENE IN *XENORHABDUS NEMATOPHILA*

by

John Lucas

A Thesis Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Master of Science

in Biological Sciences

at

The University of Wisconsin-Milwaukee

December 2016

ABSTRACT

MOLECULAR AND FUNCTIONAL ANALYSIS OF THE *pixB* GENE IN *XENORHABDUS NEMATOPHILA*

by

John Lucas

The University of Wisconsin-Milwaukee, 2016
Under the Supervision of Dr. Steven Forst

Xenorhabdus nematophila and the entomopathogenic nematode *Steinernema carpocapsae* form a mutualistic relationship facilitating the infection, death and consumption of an insect host. The infective juvenile (IJ) form of *S. carpocapsae* invades the insect host through natural openings and proceeds to the hemocoel where exposure to hemolymph stimulates the release of *X. nematophila* from the anterior vesicle. Excreted *X. nematophila* releases immunosuppressive compounds and insect toxins into the insect hemolymph that facilitates death of the host. As *X. nematophila* reaches high cell density it secretes exoenzymes that degrade insect tissues and produces antibiotics that reduce microbial competition. *S. carpocapsae* utilizes the insect remains as a nutrient source and reproductive site. When grown in nutrient-rich laboratory broth *X. nematophila* produces two distinct crystalline inclusion bodies. One crystalline inclusion is composed of the methionine-rich PixA protein. The other inclusion body is composed of IP2 (PixB), a 22 kDa protein that is not methionine-rich. While the *pixA* gene had been studied previously the *pixB* gene had not been identified. To better understand the function of PixB, the *pixB* gene

was identified and its expression was analyzed by GFP promoter reporter constructs in cells grown under biologically relevant conditions. In addition, comparative genomic and phylogenetic analyses were conducted to assess the prevalence and distribution of *pixB* in other microbial species.

A *pixA*-minus strain of *X. nematophila* grown in LB broth produced crystalline inclusion bodies as visualized by transmission electron microscopy. In contrast, inclusion bodies were rarely detected in cells grown in insect cell culture medium (Grace's medium) that mimics insect hemolymph. Furthermore, inclusion bodies were not present when bacteria were grown directly in the insect host. These findings indicate that crystalline inclusions form when *X. nematophila* grows in nutrient-rich medium but not under more biologically relevant conditions. The *pixB* gene was found to possess highly conserved σ70 promoter consensus sequences while the promoter of *pixA* was less well conserved. Expression analysis in Grace's insect medium revealed that initial expression of *pixB* occurred during late log phase growth while *pixA* was activated during mid log phase. Unexpectedly, *pixB* was expressed at 2.8-fold higher levels in Grace's insect medium than in LB broth. The reason that inclusion bodies form in LB broth even though the apparent expression of *pixB* is lower than in cells grown in Grace's medium is not presently understood. Consistent with results in Grace's medium the expression of *pixB* was higher than *pixA* in cells grown for 48 hours in the insect host. The *pixB* gene was found to be flanked by mobile genetic elements and present in only 3 of 18 sequenced *Xenorhabdus* genomes and was widespread in different classes in proteobacteria. Sequence comparison of the PixB-type proteins revealed two regions of sequence conservation (CD1 and CD2) separated by a linker region of similar length in all proteins. Phylogenetic analysis

indicated that *pixA* and *pixB* belonged to the same clade and provided evidence of extensive horizontal gene transfer of the *pixB* gene. In addition, another Pix-type protein group was identified in gammaproteobacteria that possessed glycine-rich regions. We refer to this group as the PixC-type proteins. Together, these findings represent a new family of genes that had not been previously recognized.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my family and friends for their support and understanding during this journey. I thank my fellow lab mates past and present for all their thoughtful contributions and suggestions, including Kristin, Swati, Dongjin, Molly, Matt, Michael, Lupita, Mariah and Feriel. Special thanks to Jian for his patience, teaching and kindling a newfound love for research and Mary for all her help and advice. Thank you to my committee, Sonia Bardy, Charles Wimpee, and Steven Forst for all their help and advice. I thank Heather Owen for her patient teaching with EM, Photoshop and helpful suggestions, Doug Steeber for his assistance with FACS, and Christopher Quinn for advice and nematodes. Above all, I thank Steven Forst for suggestions, help, and advice, he is truly the most patient and decent person I have ever come to know.

Chapter 1

Introduction

Xenorhabdus nematophila and the entomopathogenic nematode (EPN) *Steinernema carpocapsae* form a mutualistic relationship facilitating the infection, death and consumption of an insect host (Forst et al., 1997, Herbert and Goodrich-Blair, 2007). *X. nematophila* is a gram-negative mutualistic symbiont carried in the anterior gut vesicle of *S. carpocapsae*. The infective juvenile (IJ) form of *S. carpocapsae* invades the insect host through natural openings and penetrates the gut wall to proceed to the hemocoel where exposure to hemolymph stimulates the releases of *X. nematophila* from the anterior vesicle (Snyder et al., 2007). Excreted *X. nematophila* releases immunosuppressive compounds (Shrestha and Kim, 2007) and insect toxins (Vigneux et al., 2007) into the insect hemolymph that facilitates its death. As *X. nematophila* reaches high cell density it begins secreting exoenzymes and antibiotics that serve to degrade insect tissue and reduce competition (Singh et al., 2015). *S. carpocapsae* is then able to utilize the insect remains as a nutrient source and reproductive site (Forst et al., 1997, Herbert and Goodrich-Blair, 2007). Upon depletion of nutrients within the insect cadaver *X. nematophila* reassociates with *S. carpocapsae* which then proceeds to develop into the non-feeding IJ stage. This EPN-bacterial complex then exits the insect cadaver and proceeds through the soil in search of a new insect host (Morales-Soto et al., 2009). A free-living species of *Xenorhabdus* has not been described.

X. nematophila produces two distinct crystalline inclusion bodies when grown in nutrient-rich medium. One crystalline inclusion is composed of the cigar shaped PixA protein (26 kDa) that contains 185 residues and has an unusually high methionine content

of 8% (Goetsch et al., 2006, Couche and Gregson, 1987). To date, PixA has only been identified in *Xenorhabdus spp.* The other inclusion body is ovoid shaped and composed of IP2 (PixB), a 22 kDa protein that is not methionine-rich. It was shown to occur in both a disulfide-linked 44 kDa dimer and a 22 kDa monomer form (Couche and Gregson, 1987). PixA and PixB proteins can represent up to 55% of total cytoplasmic protein and inclusion bodies composed of these proteins can be visualized by phase contrast microscopy. Previous studies showed that PixA and PixB production occurred during late exponential phase. PixA production began at 16 hours and PixB appeared at 28 hours. Additionally, an increase in production of both proteins from 28 to 40 hours with self-associating crystal formation occurring concurrently has been characterized (Couche and Gregson, 1987). While the *pixA* gene had been studied previously (Goetsch et al., 2006) the *pixB* gene had not been identified.

At total of 23 species and strains of *Xenorhabdus* have been sequenced to date. Each species of *Xenorhabdus* is associated with its own species of *Steinernema* nematode. All four strains of *X. nematophila*, five of nine strains of *X. bovienii*, *X. innexi*, *X. doucetiae* and *X. szentirmaii* contain *pixA*-type genes (Lucas, unpublished). To what extent PixA proteins are expressed and function in the tripartite relationship is not well understood. It was proposed that PixA may serve as a nutrient source for nematode growth and/or reproduction. To address this question the *pixA* gene was identified and a *pixA*-minus strain (NMI1 was created (Goetsch et al., 2006). It was shown that PixA was not essential for in vitro growth, survival or reproduction of the nematode. Furthermore, when grown individually in the insect host colonization of the nematode was comparable with the mutant and wild-type strains. However, when the insect was co-infected with nematodes

colonized with either the wild-type or mutant strain the nematode progeny that emerged from the insect cadaver were preferentially colonized with the mutant strain. The ability of a *pixA* mutant to outcompete wt in an *in vitro* co-culture is still being studied. Whether these results reflect a competitive advantage for the *pixA*-minus strain during growth in the insect or for colonization of the nematode remains to be determined. In the present study, the *pixB* gene was isolated and its expression under biologically relevant conditions was analyzed.

Statement of Purpose

While previous *in vitro* characterization of the *pixA* gene has been performed, limited characterization of the *pixB* gene has been done and as such it is unknown what the primary function of either Pix protein is or if they are related. To better understand the function of Pix proteins, genomic analysis as well as the temporal expression profile of the *pix* genes needs to be explored. Since the initial characterization of *pixA*, many more genomes have been sequenced which may provide insight into the function and origin of the *pix* genes. Utilizing GFP reporter constructs and the 5' UTR of either *pixA* or *pixB* genes to drive GFP expression may provide valuable information into the function and or relationship between them. Expression profiles in LB, Grace's insect culture medium and insect hemolymph will need to be studied.

Chapter 2

Materials and Methods

Bacterial strains, media and growth conditions

Table S1 lists strains and plasmids and Table S2 lists primers used in this study. Liquid media consisted of either Lysogeny broth (LB), Grace's Insect cell culture (Gibco), nematode growth media (NGM), peptone glucose sorbitol (PGS) or insect hemolymph. LB media consisted of 1.0% tryptone (Bacto), 0.5% yeast extract (Difco), 0.5% NaCl (JT Baker) and 1.0 mM MgS04 (Fisher). Plate media was solidified with 1.5% agar (Difco). Grace's media (GM) was prepared per directions with pH adjusted to 6.1 or as indicated. NGM media was prepared with 0.3% NaCl, 0.25% peptone (Bacto) and 1.7% agar (Difco) 2.5% 1M KPO4 buffer (pH 6.0) 1 mM CaCl2, 13.0 mM cholesterol, 1.0 mM MgSO4. PGS media was prepared with 1% peptone (Bacto), 1% NaCl (JT Baker), 1% Glucose (Mallinckrodt Chemicals) and 0.15M Sorbitol (VWR). Hemolymph was obtained from *Manduca sexta* 5th instar larvae, Briefly, insects were anesthetized on ice for ~10 minutes, sterilized with 70% ethyl alcohol and incised anterior to the horn structure, hemolymph was drained into 1.5 ml micro centrifuge tubes and stored at -80◦C. Initial bacterial cultures were streaked from -80◦C glycerol stocks onto either LB plates supplemented with 0.0025% bromothymol blue and 0.004% triphenyltetrazolium chloride (Sigma Chemical Co.) (*Xenorhabdus)* or LB plates (*E. coli*). Media was supplemented with kanamycin (30 μg/ml), ampicillin (50 μg/ml) chloramphenicol (10 μg/ml) as required. Plate cultures were grown at 28◦C (*Xenorhabdus*) or 36◦C (*E. coli*) for 24-48 hours and stored at room temperature (22◦C) for up to one week. Liquid cultures were incubated on a gyratory shaker at 30◦C (*Xenorhabdus*) or 37◦C (*E. coli*). Long term cultures were stored in LB supplemented with

20% glycerol in -80°C. All media storage and incubations were done in the dark (Xu and Hurlbert, 1990).

Identification of pixB gene

Prior work was performed to identify the *pixB* gene by purification of PixB protein and N-terminal sequencing. PixB inclusions were purified as follows: 100 ml of stationary phase culture was pelleted by centrifugation (10,000 x g 15 minutes), resuspension in 10 ml K_2HPO_4/NaCl pH 7.3, followed by sonication on ice (2 cycles for 10 minutes output 2.5) and layering on 20 ml 30% glycerol cushions. Centrifugation was performed (450 x g 5 minutes) and glycerol fraction centrifuged again (120,000 x g 5 minutes). Pellets were suspended in 5 ml of saline and layered onto a 70/60/30% glycerol gradient (10 ml each). Gradients were centrifuged (1,500 x g 15 minutes), the 60% layer was removed and diluted with saline 1:1, and centrifuged (120,000 x g 5 minutes). Pellets were washed two times in saline, suspended in 80ul of 1X SDS loading buffer (0.08M Tris-HCl pH 6.8, 10% glycerol, 1.55% beta-mercaptoethanol, and 0.006% bromophenol blue) and run on a 16.5% polyacrylamide gel (15mA 8 hours). Proteins were transferred to Immobilon-P membrane (Millipore) in a Bio-Rad Trans-Blot Cell (90 mA for 10 hours at 4°C), stained in 40% methanol, 1% acetic acid, and 0.1% coomassie blue followed by destaining in 50% methanol, rinsing in ddH2O and air dried. Membrane bands containing the PixB protein were excised and used to perform N-terminal sequence analysis (Medical College of Wisconsin Protein Analysis Facility). The N-terminal sequence of 21 amino acids was used to identify the *pixB* gene in the *X. nematophila* genome.

Genomic Analysis of homologous pix genes

Genomic analysis was conducted utilizing the NCBI BLAST database (Altschul et al. 1997). Sequence alignments and percent identity matrix were performed with Clustal Omega (Larkin et al. 2007). Secondary structure predictions and alignments were achieved with the RaptorX alignment tool (Källberg et al. 2012; Ma et al. 2013; Peng and Xu 2011). Model assisted binding site predictions were performed with RaptorX (Källberg et al. 2012). Molecular weights were calculated using protein calculator v3.4 (http://protcalc.sourceforge.net). Residue analysis was performed at the species/strain level for each unique sequence and averages calculated for each genus and class. Each residue deviation from domain consensus sequences was also annotated and compiled for each class.

Phylogenetic Analysis of homologous pix genes

Table S3 lists sequences used for phylogenetic analysis. % ID and E-value cutoff scores used were 25% and 8.00E-04 respectively. Molecular phylogenetic analysis by maximum likelihood method was performed using the MEGA7 Molecular Evolutionary Genetics Analysis platform (Kumar et al., 2016). The evolutionary history was inferred by using the maximum likelihood method based on the Poisson correction model (Zuckerkandl et al., 1965). The tree with the highest log likelihood was generated. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete gamma distribution was used to model evolutionary rate differences among sites. The coding data was

translated assuming a standard genetic code table. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed (Felsenstein and Joseph, 1985).

Transmission Electron Microscopy analysis of X. nematophila NMI1

Cultures of *X. nematophila* NMI1 in LB, GM and hemolymph were analyzed by transmission electron microscopy. LB and GM cultures were grown at 30°C to stationary phase at which time OD600 values for LB and GM samples were 1.5 and 0.27 respectively. Hemolymph samples were obtained as follows: Overnight cultures were subcultured 1:20 in 5.0 ml LB and incubated for an additional 4 hours at 30°C. Subcultures were then serially diluted in GM. 50 µl of a 1:1000 dilution was injected behind the left first proleg of 4th instar *M. sexta* insects that were anesthetized on ice for 15-20 minutes and cleansed with 70% ethanol. Injections were carried out using a 1.0 mL, 0.45-mm by 16-mm Sub-Q syringe (Becton Dickinson Co.) that was mounted on a Stepper repetitive pipette (Dymax Corp.). Nine insects were injected. After 48 hours hemolymph was collected from anesthetized insects cleansed with 70% ethanol and incised beneath the last proleg. Insects were exsanguinated into 1.5 ml microcentrifuge tubes. Hemolymph was pooled together in 1.5 ml tubes on ice and pipetted into 15 ml polypropylene centrifuge tubes. 0.01 M HEPES pH 7.21 and 2.5% glutaraldehyde was added to either 4 ml hemolymph (1 mL), 8 mL of LB cultures (2 mL) or 40 mL of GM cultures (10 mL) and cells

fixed for 30 minutes. Cells were centrifuged (IEC HN-s2 swinging bucket centrifuge) 700 rcf for 4 minutes (LB) or 1500 rcf for 10 minutes (GM) and supernatant disposed. Cells in GM culture were transferred to a 15 mL graduated tube and both cultures rinsed two times with 0.01 M HEPES pH 7.21 and 2.5% glutaraldehyde and then fixed in 1% Osmium Tetroxide for 1 hour. Cells were centrifuged 700 rcf for 4 minutes, supernatant disposed, and rinsed two times with 0.01 M HEPES pH 7.21 and 2.5% glutaraldehyde. Cells were pelleted again as previously described, supernatant removed and replaced with 2.5 mL 20% acetone. Dehydration was carried out for 30 minutes. Samples were centrifuged again as previously described and 20% acetone replaced with 2.5 ml 40% acetone. Pellets were suspended and allowed to dehydrate for 20 minutes. Changes into 60%, 80%, and 100% acetone were performed in this manner at 20 minute intervals. Samples were centrifuged 700 rcf for two minutes; supernatant removed and replaced with 4 ml 20% Spurr's Quetol resin (7g Quetol 651, 11.1g ERL 4221, 31.9g NSA, 7.15g DER 736, 0.25g DMAE) in acetone. Cells were allowed to infiltrate for 15 minutes. Samples were centrifuged again and 20% resin replaced with 40% resin and allowed to infiltrate for 15 minutes. Changes into 60%, 80%, and 100% resin were carried out in this manner every 15 minutes. 100% resin was removed and replaced with a second 100% resin change. BEEM capsules were filled with 500 µl of each sample, 1.5 ml Eppendorf microcentrifuge tubes were filled with modeling clay and lids were removed. BEEM capsules were seated inside the Eppendorf tubes. Capsules were centrifuged 4 at a time in a Fischer marathon km starting at 1000 rpm and gradually increasing to 13600 rpm for 10 minutes centrifuge time. After ten minutes at 13600 rpm capsules were topped off with 100% Spurr's Quetol resin and placed in a 60°C oven for 22 hours to polymerize. Polymerized blocks were removed from the BEEM

capsules. A 0.5mm trapezoid was cut into the face of the block using a razor blade. 0.7 micron semi-thin sections were taken with a RMC MT-7 microtome using a glass knife. Semi-thin sections were lifted onto a clean glass slide with a drop of acetone using a mounted beard hair. Sections were stained with toluene blue and heated gently on a hot plate. Glass slide was gently rinsed under water and visualized under a 400X light microscope to confirm the presence of cells. Block faces were trimmed as necessary and ultra-thin (70 nm) sections were obtained and transferred to 200 mesh copper grids. A glass petri dish with parafilm adhered to the bottom was prepared and grids containing ultrathin sections were placed on the parafilm section side up. One drop of 1% uranyl acetate was placed on each grid and allowed to stain for seven minutes. Grids were rinsed for 60 seconds each with sterile filtered water. Several grains of NaOH were added to a covered petri dish with adhered parafilm. Grids were placed section side up on the parafilm and one drop of Reynolds' Lead Citrate (1.33g Pb(NO3)2, 1.76g Na(C6H5O7) *2H2O, 30 ml CO2 - free H2O) was added to each grid. Dish was quickly covered and allowed to stain for seven minutes. Grids were then rinsed in sterile filtered water for 60 seconds each. Stained grids were viewed with a Hitachi H-600 aligned using a 3mm Holey carbon film grid (TED PELLA). Representative micrographs of cells were taken.

Insertional inactivation of pixB

Prior work was performed to create the *pixB* mutant NMI2. N-terminal sequence analysis was used to identify the *pixB* gene sequence within the X. *nematophila*. Internal sense primer M100 (5'GCAATCTATCATGAGAGATTTTGGC3') and antisense primer M101 (5'CCCAATAGCAATTAACTGTTGGTTGG3') were designed for PCR amplification of a 356 bp

internal fragment from isolated genomic DNA (Edge Biosystems kit). The amplified fragment was ligated into pSTBlue-1 and transformed into XL1-Blue MRF' (Stratagene protocol). Transformed colonies were verified by colony PCR. Purified pSTBlue-1 containing the *pixB* internal fragment and pKnock-Cm (Alexeyev, 1999) were then subjected to a double digest with Xba1 and Pst1. The gel purified digested insert was then ligated into pKnock-Cm (pMG006) and electroporated into *E. coli* S17-1 λ pir. *E. coli* S17-1 λ pir containing pMG006 was then conjugated with *X. nematophila* 19061. Successful conjugates were confirmed based on growth on chloramphenicol selection with insertional inactivation of the *pixB* (NMI2) confirmed with PCR screening.

SDS-PAGE confirmation of Pix protein expression in X. nematophila, NMI1 and NMI2

Overnight cultures of *X. nematophila* 19061 (wt), NMI1 and NMI2, were collected on ice and transferred to preweighed and chilled 1.5 ml Eppendorf tubes. Cultures were pelleted in a Beckman Microfuge E at 10,000 x g for five minutes at 4◦C. Supernatant was discarded and pellet weight calculated. Pellets were then suspended in 1.0ml of chilled 20mM NaPO4 pH 7.2. Cultures were centrifuged in a Beckman Microfuge E at 10,000 x g for thirty seconds at 4◦C. Supernatant was discarded and replaced with 1.0ml of chilled NaPO4 buffer. Tubes were sonicated for four cycles of 2.5 min bursts at 80% duty and output of 10. Samples were pelleted in a Beckman Microfuge E as previously described after each cycle. Cultures were then centrifuged once more, supernatant discarded and pellets suspended in 20 µl of 1x SDS sample buffer (0.08M Tris-HCl pH 6.8, 10% glycerol, 1.55% beta-mercaptoethanol, and 0.006% bromophenol blue) per mg of pellet weight. Samples were run on an 8-16% gradient SDS-PAGE gel in MOPS running buffer (Genscript)

at 140 volts. Gel was stained in 40% methanol, 1% acetic acid, and 0.1% coomassie blue followed by destaining in 50% methanol, and rinsing in ddH2O.

Phenotypic analysis of X. nematophila, NMI1 and NMI2

Phenotypic characterization of *X. nematophila* 19061 (wt), NMI1 and NMI2 was conducted. Tests for Catalase, Oxidase, H2S Production, Amylase, Gelatinase, Lactose fermentation, dye binding on LBTA, O2 requirements, Glucose fermentation, Urease Production, DNase production, Nitrate reduction, Indole production, and Sucrose fermentation were performed as previously described (Goetsch et al., 2006, Volgyi et al., 1998).

In vitro competition

5.0 ml overnight cultures of wt, NMI1 and NMI2 were subcultured 1:20 in LB broth and incubated at 30^{o}C for 4 hours. OD600 values of 4 hour subcultures were obtained with a Tecan M200 plate reader and all cultures normalized to 0.30. Competition cultures were created by 1:1 normalized mixtures of wt and NMI1 or NMI2. Competition cultures and each individual strain was serially diluted in GM and spread on LB plates containing either ampicillin or chloramphenicol selection with sterile glass beads. Plates were incubated at 30^{o}C for 24 hours. The experiment was performed six times.

Rearing of Manduca sexta larvae

M. sexta larvae were raised from eggs obtained from Carolina Biological Supply Company. Eggs were incubated in plastic cups at room temperature in an insect incubator

with a 16-hour light/8-hour dark photoperiod. After hatching, larvae were moved to clean boxes and provided with fresh diet. The boxes were cleaned daily, and larvae were fed regularly. The diet was prepared from individual ingredients as previously described (Singh et al., 2015). Fourth-instar-stage larvae were used for all experiments.

In vivo Competition

5.0 ml overnight cultures of wt, NMI1 and NMI2 were subcultured 1:20 in LB broth and incubated at 30°C for 4 hours. OD600 values of 4 hour subcultures were obtained with a Tecan M200 plate reader and all cultures normalized to 0.30. Competition cultures were created by 1:1 normalized mixtures of wt and NMI1 or NMI2. Competition cultures and each individual strain were serially diluted in GM. Fourth-instar *M. sexta* larvae were anesthetized on ice for 20 minutes. Anesthetized insects were cleansed around the horn with 70% EtOH and injected with a 1:1000 dilution of each strain. 50 µl of each diluted strain was injected with a 1.0 mL, 0.45-mm by 16-mm Sub-Q syringe (Becton Dickinson Co.) that was mounted on a Stepper repetitive pipette (Dymax Corp.). Injected insects were moved to plastic cups until hemolymph was collected. Hemolymph was collected from anesthetized insects cleansed with 70% ethanol and incised beneath the last proleg. Insects were exsanguinated into 1.5 ml micro centrifuge tubes individually and later pooled together. Pooled hemolymph was serially diluted in GM and plated on LB agar plates with appropriate selection. Three to nine insects per strain per experiment were used, and the experiment performed four times.

Viability Assay

To assess the relative long term viability of wt, NMI1, and NMI2 strains, 24 hour cultures (3 per strain) were diluted 1:500 in 20 ml of either LB or GM and grown for 72 hours at 30°C. 1.0 ml samples were collected at 5.5, 10.5, 16, 20, 24, 48, and 72 hours post inoculation. Samples and standard curve were prepared per LIVE/DEAD BacLight bacterial viability kit microplate protocol (Invitrogen L7012). 200 µl of each stained sample was read in a 96 well microplate (Falcon 351172). OD600, and fluorescence excitation and emission readings (485/500 nm (live) and 485/635 nm (dead)) were recorded in a Tecan M200pro microplate reader. Live/dead ratios for each time point were converted to percent viable cells using a standard curve. Briefly, 1:500 dilutions of *X. nematophila* overnight cultures were incubated at 30°C for 8 hours in either LB or GM. 1.0 mL of each culture was then pelleted and suspended in either 0.85% NaCl (Live) or 70% Isopropyl alcohol (Dead) and incubated in the dark at room temperature with gentle resuspension every 15 minutes. After incubation samples were pelleted and suspended in 0.85% NaCl. Live: Dead ratios of 0:100, 10:90, 50:50, 90:10, and 100:0 were created and 100 µl of each Live: Dead ratio was added to the wells of 96 well microplate, each ratio was plated in triplicate. 100 µl of a 2X staining solution was added to each well, plate was allowed to incubate in the dark at room temperature for 15 minutes and fluorescence readings obtained. Readings for each ratio were averaged and plotted to generate a standard curve.

Rearing and synchronization of C. elegans

C. elegans was maintained on NGM plates spotted with lawns of *E. coli* OP50 and passaged weekly to fresh plates. Synchronization of L1 populations was performed by washing 3 day old plates with M9 buffer, washings were transferred to a 15.0 ml falcon tube and worms allowed to sediment. To dissolve the nematodes M9 buffer was exchanged with 5.0 ml 20% alkaline hypochlorite solution and mixed gently for 3 minutes. Eggs were pelleted and hypochlorite solution exchanged with M9 buffer. Eggs were pelleted again and washed with M9 buffer three more times. Nematode death and egg collection was visually confirmed under 100X. Eggs in solution were allowed to hatch overnight with gentle rocking on a Nutator platform. Synchronized L1 hatchlings were visually confirmed under 100X within 24 hours.

Nematode Killing Assay

5.0 ml overnight cultures of *E. coli* BL21, and duplicate pMG006 cultures in LB were subcultured 1:20 with one pMG006 culture induced with 200 µl 100mM IPTG. Cultures were then incubated for 3 hours. Cultures were then pelleted and washed twice with 0.85% NaCl and suspended. 400 µl of each culture was then spotted on either NGM (slow killing) or PGS (fast killing) plates (X3 per strain) and incubated overnight at 36^{o}C. Approximately 10-15 L4 synchronized nematodes raised on lawns of *E. coli* BL21 were transferred to each plate. Actual number of nematodes on each plate was verified with a dissecting microscope under 50X magnification. The number of dead worms on each plate was assessed at 24 and 48 hours.

Total RNA isolation

4.0 ml of an *X. nematophila* 24-hour overnight culture was pelleted at room temperature, media was removed and cells frozen on dry ice for 10 minutes. 1.0 ml of Trizol (Invitrogen) was added and sample mixed by pipetting. Cells were then allowed to incubate at room temperature for 5 minutes. 200 µl of chloroform was added and tube vortexed for 15 seconds to mix, sample was then incubated for an additional 5 minutes at room temperature. The sample was then centrifuged at maximum speed for 15 minutes at 4°C. The aqueous layer was transferred to a new tube and 500 µl of isopropyl alcohol added. The tube was inverted several times to mix and then allowed to incubate at room temperature for 10 minutes. Sample was then centrifuged at maximum rpm at room temperature for 10 minutes. Following centrifugation, the alcohol was removed and pellet washed with 1.0 ml 75% ethyl alcohol. Sample was then centrifuged at maximum rpm for 5 minutes at 4°C, alcohol was removed and pellet was allowed to dry. The dried pellet was suspended in 50 µl nuclease free ddH20. Quality and concentration of RNA was assessed by gel electrophoresis and Nanodrop spectrophotometer.

Determination of Transcription Start Sites and 5'UTR Analysis

To obtain the transcription start sites of both Pix proteins the 5'-RACE system (Invitrogen) was used per the manufacturer's instructions. Reverse transcription was performed utilizing antisense primer RGSP3A (*pixA*) or RGSP1B (*pixB*) and 7 µg of total RNA. The first PCR reaction utilized the sense abridged anchor primer (Invitrogen) and antisense primer RGSP3B (*pixB*) or RGSP4A (*pixA*). The PCR reaction consisted of 35 cycles of 94°C for 60 seconds, 55°C (*pixB*) or 70°C (*pixA*) for 60 seconds, 72°C for 30 seconds. PCR

products were visualized on an ethidium bromide stained 2% agarose gel after which the bands were excised and eluted utilizing column purification (Geneclean). The second PCR reaction utilized sense abridged universal amplification primer (Invitrogen) and the antisense primers as used in the first PCR reaction with identical reaction conditions. The second PCR products were visualized, eluted and column purified as previously mentioned. The purified PCR product was then blunt cloned into pSTBlue-1 and sequenced (Sequetech) with vector T7 sense and SP6 antisense primers as well as RGSP3B (*pixB*) or RGSP4A (*pixA*) and abridged universal amplification primer. The 5' UTR upstream from the transcription start site of both genes was analyzed using BPROM prediction software to elucidate possible transcription factor binding sites and promoter elements (Solovyev 2011).

Construction of GFP Reporter Constructs

X. nematophila 19061 gDNA was isolated using a PurElute Bacterial Genomic isolation kit per instructions (EdgeBio). The 400bp 5'UTR that included ribosome binding sites of each *pix* gene was amplified from gDNA with primers which incorporated a 5' HindIII site and a 3' EcoR1 site. PCR products were column purified (Geneclean). pProbe-NT (addgene) and PCR products were digested with HindIII-HF and EcoR1-HF (New England Biolabs). The 5'UTR of each gene was ligated into pProbe-NT and transformed into *E. coli* DH5α. Overnight cultures of *E. coli* DH5α containing either the *pixA* reporter (pLucy), or the *pixB* reporter (pPB400) were subjected to plasmid DNA isolation (Qiagen). Isolated plasmid DNA was then used to transform *E. coli* S17 λ pir.

Transformation of E. coli

Competent *E. coli* cells were created from an overnight culture that was subcultured 1:20 and incubated for 4 hours. 500 µl of culture was transferred to 1.5 ml microfuge tubes and placed on ice for 10 minutes. Cells were then pelleted by centrifugation at 1377 x g for 5 minutes at 4°C after which the supernatant was discarded and pellets suspended in 500 µl of transformation buffer (50 mM CaCl2 25.0 ml, 10 mM Tris-HCl 5.0ml, ddH20 470 ml, pH 8.0 at -20°C). Tubes were then placed on ice for 15 minutes and pelleted as previously described. Supernatant was discarded and replaced with 50 µl fresh transformation buffer and tubes returned to ice for 15 minutes. Approximately 200 ng of DNA was added to cells and stirred with pipette tip. Cells were then returned to ice for 15 minutes and then heat shocked for 30 seconds at 42°C and then returned to ice for 5 minutes. 250 µl of 37 °C SOC was and culture then incubated at 37 °C for one hour. After incubation cells were spotted on LB plates with appropriate selection and spread with sterile beads. Plates were incubated overnight at 37 °C and colonies screened for successful transformants via cPCR.

Conjugation of plasmids into X. nematophila

Overnight cultures of wt, NMI1, NMI2 and *E. coli* S17 λ pir carrying either the *pixA* reporter construct (pLucy), *pixB* reporter (pPB400) or promoterless vector pProbe-NT were subcultured in LB and grown for four hours. Each culture was pelleted in a 1.5 ml micro centrifuge tube, washed once with 1.0 ml LB and suspended in 1.0 ml LB. 100 µl of *E. coli* S17 λ pir culture was added to 100 µl of either wt, NMI1 or NMI2. 100 µl of each mixed culture was plated on LB agar plates and incubated overnight at 28°C. Conjugation plates were washed with 2.0 ml LB broth containing appropriate antibiotics and lawns lifted with

a sterile spreader. 100 µl of each washed culture was plated on LB agar plates containing selection and incubated overnight at 28°C. Conjugated colonies of wt, NMI1 and NMI2 carrying pLucy or pPB400 were verified by, the presence of GFP fluorescence under λ 395 nm, a negative catalase test and their characteristic smell of fresh cut grass.

Gene Expression assay

pix gene expression analysis in *X. nematophila* was performed following the methods of Jubelin et al., (2011). 24 hour cultures of wt, NMI1, NMI2, or the *rpoS* mutant, carrying either pProbeNT, pLucy or pPB400 vectors were diluted 1:500 in either LB or GM with Kanamycin selection. 200 µl of each dilution was loaded in a 96-well plate (Corning). Surrounding wells were loaded with 200 µl of sterile ddH2O. Plate was sealed with parafilm and incubated at 28°C with orbital agitation for 24 hours in a Tecan M200 plate reader. OD600 and GFP fluorescence (488/520 nm) were recorded every 30 minutes. Background fluorescence was eliminated by subtracting initial fluorescent readings from all subsequent readings. Specific fluorescence was obtained by dividing the fluorescence values by absorbance values. Each assay was performed at least three times and readings averaged.

For expression assays performed in hemolymph, 24 hour cultures of wt carrying either pProbeNT, pLucy or pPB400 vectors were grown in NGM. 1.0 ml of culture was then pelleted 16,873 x g for 60 seconds media removed and pellet suspended in 1.0 ml 0.85% NaCl and pelleted as previously described, procedure was performed once more for a total of two washings. Washed pellets were suspended in 50 µl of 0.85% NaCl. 20 µl of

suspended culture was added to 180 µl of hemolymph in the plate under a blanket of CO_2 to prevent melanization. Plate was prepped and assay performed as previously described.

In vitro and in vivo flow cytometry analysis

Xenorhabdus strains carrying either pLucy, pPB400 or pProbe-NT GFP reporter constructs were analyzed by flow cytometry. Serial dilutions up to 10^{-6} were analyzed using a FACSCalibur flow cytometer (BD Biosciences). Dilutions were processed at 60 µl/min for one minute and GFP expressing cells gated for evaluation of population level expression.

In vitro analysis was carried out on 24 hour cultures of each strain that were grown in triplicate under appropriate selection and then subcultured 1:500 into 20 ml of 0.2 µm filtered LB or GM with kanamycin selection. Cultures were incubated at 30°C for 24 hours. Samples were taken at various time points over 24 hours and serial dilutions created in 0.2 µm filtered LB or GM. The mean fluorescent intensity readings from each strain were averaged.

In vivo analysis was performed on fourth instar *M. sexta* hemolymph. Wt strains carrying either pLucy, pPB400 or the promoterless vector pProbeNT were grown overnight (~22hrs) at 30°C on a gyratory shaker in LB plus ampicillin and kanamycin selection. Overnight cultures were subcultured 1:20 in 5.0 ml LB and incubated for an additional 4 hours at 30°C. Subcultures were then OD normalized (600 nm) and serially diluted in GM. 50 µl of a 1:1000 dilution was injected behind the left first proleg of 4th instar *M. sexta* insects that were anesthetized on ice for 15-20 minutes and cleansed with 70% ethanol. Injections were carried out using a 1.0 mL, 0.45-mm by 16-mm Sub-Q syringe (Becton Dickinson Co.) that was mounted on a Stepper repetitive pipette (Dymax

Corp.) nine insects were injected per condition. At 48 hours hemolymph was collected from anesthetized insects cleansed with 70% ethanol and incised beneath the last proleg. Insects were exsanguinated into 1.5 ml micro centrifuge tubes. Hemolymph from at least three insects was pooled and used to create serial dilutions in 0.85% saline.

In vivo Competition

Wt, NMI1 and NMI2 2.0 ml LB cultures containing either Ampicillin and Kanamycin or Ampicillin, Kanamycin and chloramphenicol selection were started from single colonies and grown overnight (~22hrs). Overnight cultures were then subcultured 1:20 in 5.0 ml of LB and incubated for four hours. Subcultures were normalized to 0.25 (OD600). Competition cultures were created using a 1:1 normalized mixture of either wt and NMI1 or wt and NMI2. Cultures were serially diluted in GM and 50 µl of a 1:1000 dilution was injected behind the left first proleg of 4th instar *M. sexta* insects as previously described. Nine insects were injected per condition. Insect mortality was recorded at 24 and 48 hours. At 24 and 48 hours hemolymph was collected from anesthetized insects cleansed with 70% ethanol and incised beneath the last proleg followed. Insects were exsanguinated into 1.5ml micro centrifuge tubes. Hemolymph form at least three insects per condition were pooled and used to create serial dilutions in GM, dilutions were then plated and patched on LB media with selection as previously described. Plates were incubated for 48 hours at 28°C. To determine what percentage of the competition strains were composed of mutant cells, colonies on competition plates were patched onto LB plates containing ampicillin, kanamycin, and chloramphenicol selection and incubated for 48 hours at 28°C. The experiment was repeated five times.

Chapter 3

Genomic and phylogenetic analysis

Genomic analysis of pixB in Xenorhabdus strains

A BLAST search of 23 sequenced *Xenorhabdus* strains and species was performed using the *X. nematophila* 19061 PixB sequence. The search results indicate a PixB homolog does not exist in every species of *Xenorhabdus*. Homology of PixB across *Xenorhabdus* species ranges from 31% to 90% identity. A PixB homolog was found in all *X. nematophila* that had been sequenced. These include *X. nematophila* isolated from *S. carpocapsae* (19061, F1), *S. websteri, S. anatoliense,* and an unidentified *Steinernema* strain (C2-3). All sequences were identical to PixB except the F1 strain that was 90% identical. Among the other species that contained a PixB protein were *X. cabanillasii* isolated from *S. riobrave* (57% identity), X. *intermedium* from *S. intermedium* (34% identity), and finally *X. bovienii-S. Kraussii* (Quebec) (43% identical). Interestingly, another *X. bovienii-S. Kraussii* strain did not possess a PixB homolog (Table 1). The genomic location of *pixB* among the *Xenorhabdus* species that contain a PixB protein was variable. Clustal alignment of *Xenorhabdus* PixB homologs revealed two conserved domains. The sequence between G65 and S74 in *X. nematophila* GDxxRWRxx(S/T) (CD1) was conserved in all strains. A putative Protein Kinase B (PKB) like motif RxRxx(S/T) occurs within this sequence (Manning and Cantley 2007). A second conserved sequence G(Y/C)xxWDP (CD2) was located between G170 and P176 (Fig 1). Both domains are linked together by a domain (LD) consisting of 95 to 96 residues. Together the three domains (CD1-LD-CD2) compose the PixB motif that is specific to PixB homologs (Fig. 1).

TABLE 1: Clustal 2.1 percent identity matrix of *Xenorhabdus* Pix protein sequences

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1: XBKQ1_PB	100	38	43	42	31	25	25	25	25	21	23	20	23	23
2: XCR1_PB		100	57	55	35	28	28	28	28	26	28	22	23	23
3: XNC1_PB			100	90	34	36	36	36	36	33	34	26	29	31
4: XNF1_PB				100	33	34	34	34	34	32	32	24	27	28
5: XBI1_PB					100	35	36	36	35	36	35	30	29	29
6: XBFFL1_PA						100	99	99	99	54	53	48	47	46
7: XBI1_PA							100	100	99	54	53	48	47	46
8: XBKQ1_PA								100	99	54	53	48	47	46
9: XBP1_PA									100	54	53	48	47	45
10: XCR1_PA										100	70	39	37	37
11: XiS1_PA											100	42	40	39
12: XDD1_PA												100	50	45
13: XNC1_PA													100	77
14: XSR1_PA														100

Identifiers: PB = PixB (shaded), PA = PixA, XBKQ1= *X. bovienii kraussei quebec*, XCR1= *X. cabanillasii*, XNC1= *X. nematophila* 19061, XNF1= *X. nematophila* F1, XBI1= *X. bovienii intermedium*, XBFFL1= X. *bovienii feltiae florida*, XBP1= *X. bovienii puntauvense*, XiS1= *X. innexi*, XDD1= *X. doucetiae*, XSR1= *X. szentirmaii*

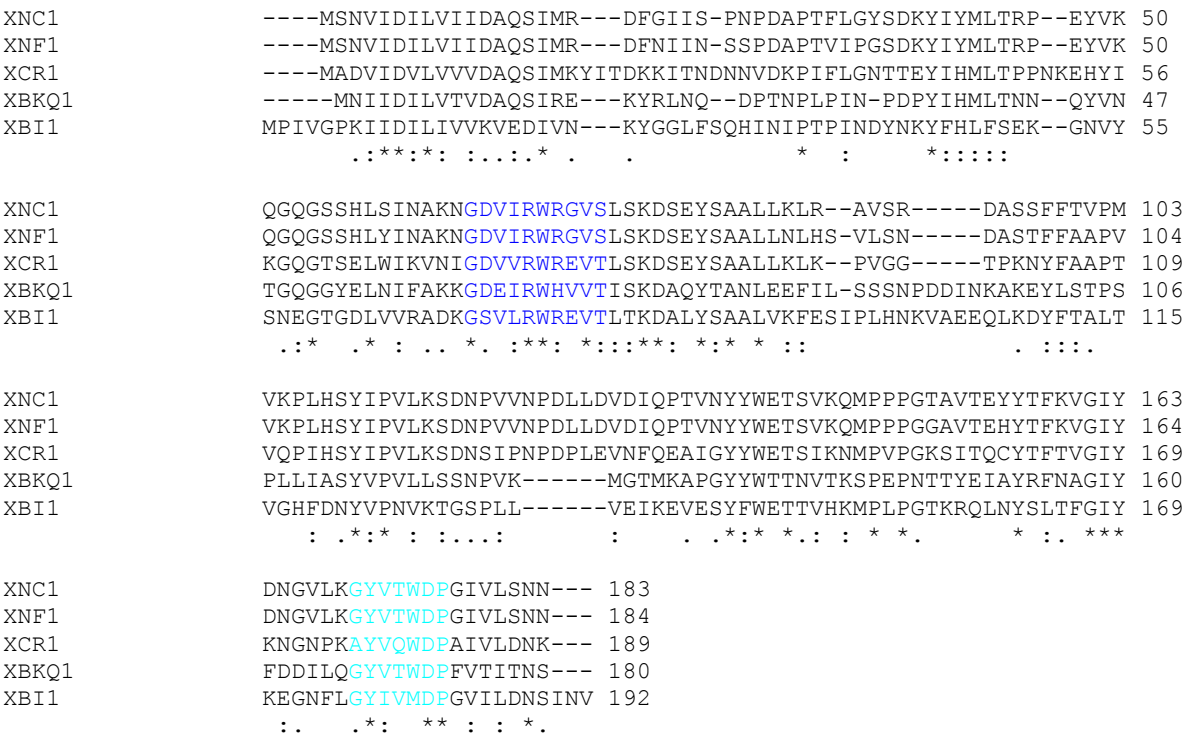


Figure 1: CLUSTAL 2.1 multiple sequence alignment of Xenorhabdus PixB homologs. Multiple sequence alignment of *Xenorhabdus* PixB homologs *X. nematophila 19061 (XNC1), X. nematophila F1 (XNF1), X. cabanillasii JM26 (XCR1), X. bovienii Intermedium (XBI1), X. bovienii Kraussei Quebec (XBKQ1)* reveal two conserved Domains. CD1 (__) GDXXRWRXX(S/T) which contains a Protein Kinase B (PKB) like sequence RXRXX(S/T) and CD2 (__) GYXXWDP.

Sequence alignment of *X. nematophila* 19061 PixB and PixA sequences reveal a 30% identity between PixA and PixB (Fig. 2). Genomic analysis of various *Xenorhabdus* species (Table 2) indicated that *pixA* is not present in every species. Furthermore, *pixA* is geographically diverse and not distinguishable based on clade analysis or genomic position (Lee and Stock, 2010). Additionally, not all species in which *pixA* is present contain the *pixB* homolog. Of those sequences analyzed, the occurrence of methionine rich PixA alone is more prevalent than those containing both a PixA and PixB protein. Among the *Xenorhabdus* species that have a *pixA* and/or *pixB* gene, protein identity ranges from 20% to 36%. Homology of PixB across *Xenorhabdus* species ranges 23% to 29% identity when compared to *X. nematophila* 19061 PixA (Table 2). The predicted secondary structure of PixB consists of an N-terminal alpha helix and two C-terminal beta sheets which contain CD1 and CD2. The beta sheets are separated by a loop structure consisting of the LD. Secondary structure alignments of *X. nematophila* PixA and PixB sequences yield a TMscore of 0.827 predictive (90% chance) that the two proteins share a similar fold (Fig. 3).

Genbank BLAST analysis of the PixB motif

The PixB motif from *X. nematophila* was used to generate BLASTp sequence results. PixB proteins contain sequence variability at the N-terminal sequence adjacent to CD1, LD and C-terminal sequence adjacent CD2. Due to the degree of sequence variability, identity values were found to be a poor indicator of PixB identity; as a result, whole FASTA sequences generated from the BLASTp search were manually scanned for the presence of the PixB motif. 339 sequences were identified as having the PixB motif. PixB homologs were present in multiple other genera, these include numerous members of the gamma

```
XNC1_PixB      MSNVIDILVIIDAQSIMRDFGIISPNPDAPTFLGY---SDKYIYMLTRPEYVKQGQGSSH     57
XNC1_PixA      -MRNIDIMLAVKAEEIMNDYGKLSKDINKPVLITPSILNKNYLRLLPEDDSVVCVDEELT     59
                 . ***:: :.*:.**.*:* :* : : *.::     ..:*: :* . : *   : .

XNC1_PixB      LSINAKNGDVIRWRGVSLSKDSEYSAALLKLRAVSRDAS---SFFTVPMVKPLHSYIPVL     114
XNC1_PixA      LSIKANIGDMIRLNVRSLNIETTYSAALVKIEPMNVMRDGVVSMVSLPVVES------VV     113
               ***:*: **:** .  **. :: *****:*:. :.   .   *:.::*:*:       *:

XNC1_PixB      KSDNPVVNPDLLDVDIQPTVNYYWETSVKQMPPPGTAVTEYYTFKVGIYDNGVLKGYVTW     174
XNC1_PixA      RSVATVDMKNTVDVDINTMKDYHWMIKVNDLPSVDRISTMYYLSTIAIYRDKMLMGYIAL     173
               :*   *   : :****:   :*:*  .*:::*      * **  .:.** : :* **::

XNC1_PixB      DPGIVLSNN---      183
XNC1_PixA      ESGLILDHTVMM      185
               : *::*...
```

Figure 2: *X. nematophila* Clustal alignment of PixA and PixB. Clustal 2.1 percent identity matrix indicates a 30.11% identity between the two sequences.

Table 2: Occurrence of *pixA* in various *Xenorhabdus* genomes

Strain	Nematode Host	Origin	PixA?	Clade	Amino Acids	Methionine Residues	Genome location	% GC
X. nematophila (19061)	*S. carpocapsae*	USA	Y	II	185	15	5'-1044134 -1043577	36
X. nematophila (Websteri)	*S. websteri*	Peru	Y	II	185	15	Shotgun 16,568 -17,125	36
X. nematophila (Anatoliense)	*S. anatoliense*	Jordan	Y	II	185	15	NA	36
X. cabanillasii (JM26)	*S. riobrave*	USA	Y	IV	203	34	5'-4275623-4276249	37
X. bovienii (Kraussei Quebec)	*S. Kraussei*	Canada	Y	III	184	16	Shotgun 40183-40737	33
X. bovienii (feltiae Florida)	*S. feltiae (Florida)*	USA	Y	III	184	16	Shotgun 2815-3369	33
X. bovienii (feltiae France)	*S. feltiae (France)*	France	Y	III	184	16	Shotgun 2815-3369	33
X. bovienii (Intermedium)	*S. intermedium*	USA	Y	III	184	16	Shotgun 40183-40737	33
X. bovienii (puntauvense)	*S. puntauvense*	Costa Rica	Y	III	184	16	Shotgun 9357-9911	33
X. doucetiae	*S. diapripesi*	USA	Y	I	197	29	5'-3125045-3125638	37
X. innexi	*S. scapterisci*	Uruguay	Y	IV	184	21	NA	NA
X. szentirmaii	*S. rarum*	Argentina	Y	II	185	10	5'-3822156-3822713	37
X. poinarii	*S. glaseri*	USA	N	I	NA	NA	NA	NA
X. bovienii (jollieti)	*S. jollieti*	USA	N	III	NA	NA	NA	NA

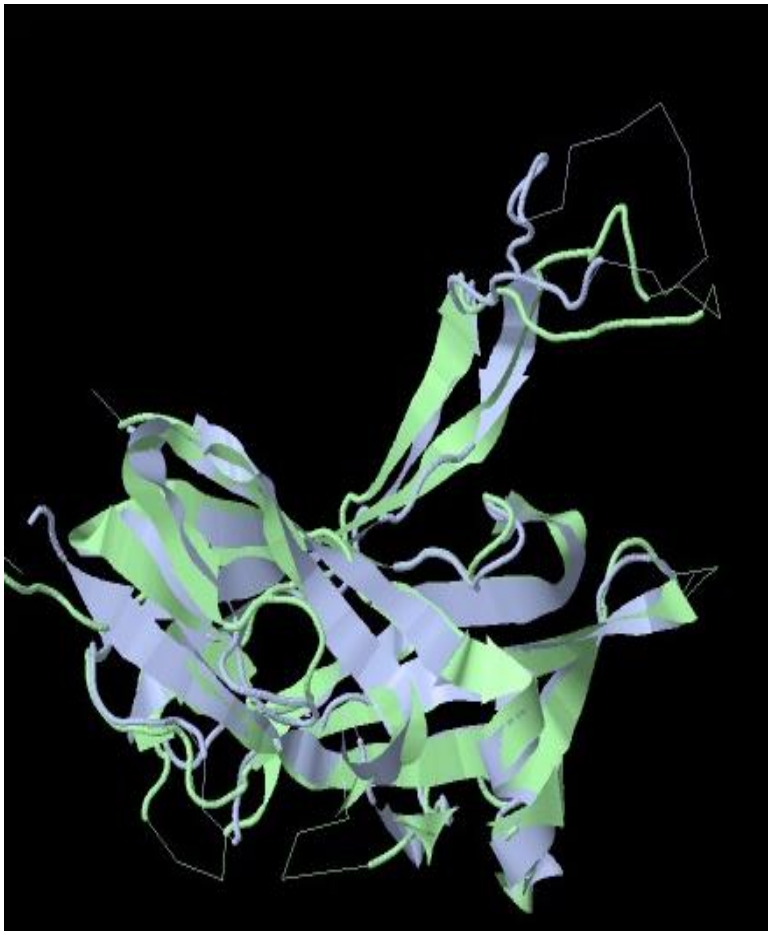


Figure 3: Secondary structure alignments of *X. nematophila* PixA and PixB. Secondary structure alignments of PixA (__) and PixB (__) yield a TMscore of 0.827 predictive (90% chance) that the two proteins share a similar fold.

proteobacteria (*Halomonas, Shewanella, Photorhabdus, Pectobacterium, Rheinheimera, and Pseudomonas*), beta proteobacteria (*Burkholderia, Streptosporangium, Chromobacterium, and Ralstonia*), alpha proteobacteria (*Rhizobium, Bradyrhizobium, and Caulocator*), delta proteobacteria (*Hylangium and Myxococcus*) and some Gram-positive organisms (*Mumia, Streptomyces, and Streptosporangium*) (Table S4). Collectively, these organisms represent four classes of proteobacteria and some Gram-positive organisms.

Phylogenetic analysis

The evolutionary history of selected PixB homologs was inferred by using the Maximum Likelihood method based on the Poisson correction model. The tree with the highest log likelihood (-12048.2154) was generated (Fig 4). The percentage of trees in which the associated taxa clustered together is shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 4.9310)). The analysis involved 41 amino acid sequences. There was a total of 255 positions in the final dataset. The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analyzed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches (Fig S1).

Analysis with a selected subset of PixB proteins (Table S3) indicated PixB and PixA proteins of *Xenorhabdus spp* formed a single clade suggesting gene duplication events and subsequent sequence divergence. PixA proteins of *X. bovienii* exhibited >90% identity and PixA in Clade II species (*X. nematophila* and *X. szentirmaii*) were most closely related.

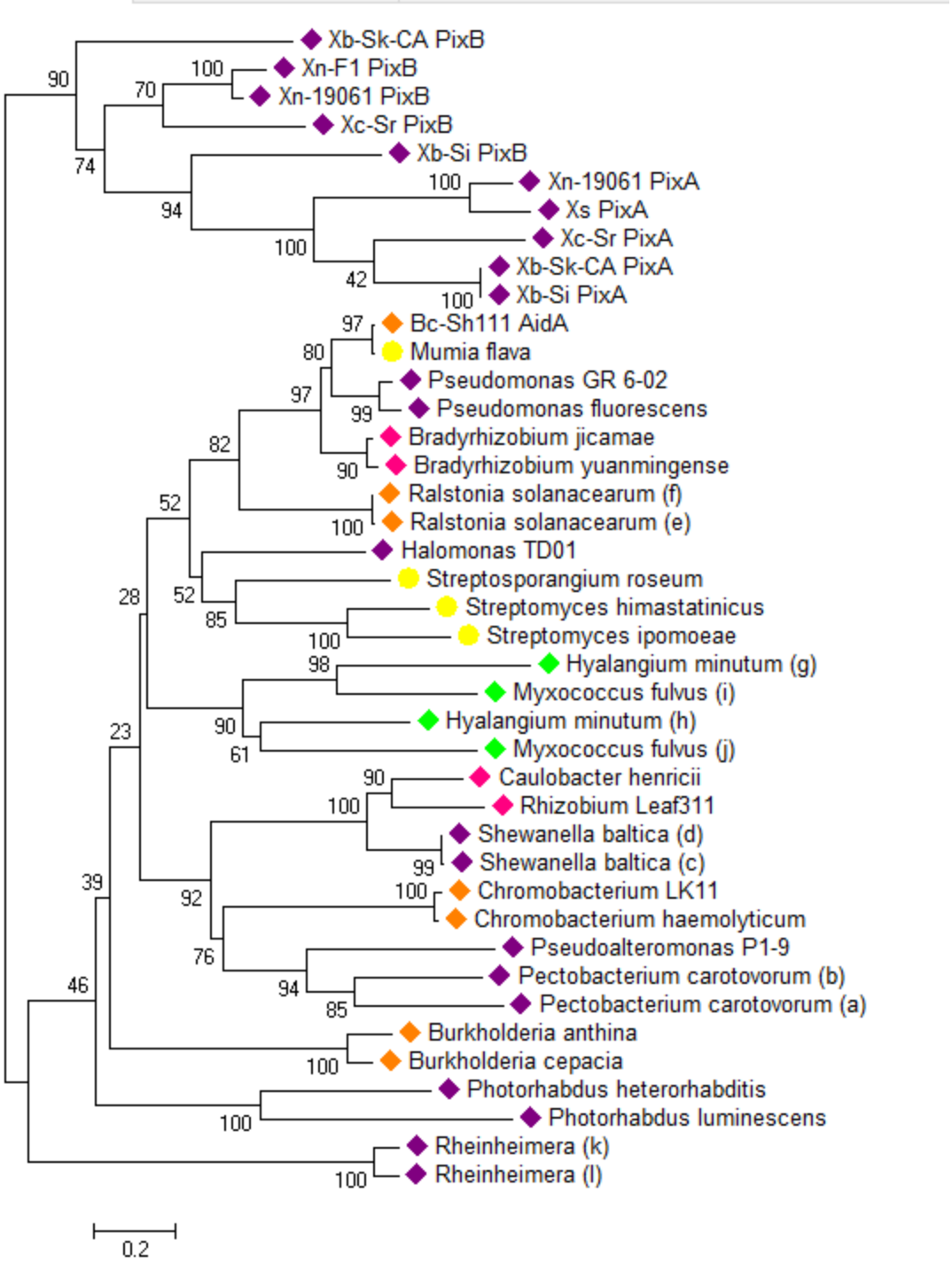


Figure 4: Molecular Phylogenetic analysis by Maximum Likelihood method of PixB homologs. Bacterial classes are distinguished with colored diamonds (Gram negative) or circles (Gram positive): Actinobacteria-yellow, alphaproteo-red, betaproteo-orange, gammaproteo-purple, deltaproteo-green. Branching is not class specific suggesting PixB is acquired via horizontal gene transfer.

PixB proteins were present in different classes of Proteobacteria being most prevalent in the γ Proteobacteria (purple diamonds). Phylogenetic relationships for the PixB proteins were not class-specific. For example, PixB in *Pseudomonas aeruginosa* was more similar to PixB in *Bradyrhizobium* (α Proteobacterium) and AidA of *Burkholderia cenocepacia* (β Proteobacterium) than to PixB proteins in other γ Proteobacteria. AidA was shown previously to possess "slow-killing" activity against *C. elegans*. AidA shared a high level of identity with a PixB-type protein from the gram-positive species, *Mumia flava* and was more distantly related to the PixB-type proteins found in other *Burkholderia* species.

Residue analysis of the PixB motif in PixB homologs

The LD of all homologs ranges from 83 to 97 residues (Table 3). The molecular weight of these proteins is 14 kDa to 22 kDa (Table 4). Multiple gene duplications as well as C-terminal truncated fragments were found in *Photorhabdus* and *Burkholderia* (Table S1). Among all classes, residue analysis indicates the greatest percentage of residues in the LD occurs with serine and threonine averaging 15% to 20% (Table 5). The average LD residue composition is 40% hydrophobic, 43% polar, 9% acidic, and 9% basic (Table 6). Residue analysis of CD1 and CD2 was performed on all genera and classes (Table S5, 7). CD analysis indicates the greatest variation in each CD sequence occurs with betaproteobacteria and gammaproteobacteria, the greatest conservation occurs in Gram (+), and alphaproteobacteria. The relative degree of conservation among all classes is Gram (+)>Alpha>Delta>Beta>Gamma.

Table 3: Average PixB motif and linker domain length by Genus and Class.

Family	Class	Genus	Length	LD Length
PixA	Gammaproteobacteria	Xenorhabdus	184-214	93-124
PixB	Actinobacteria	Mumia	169	83
		Streptosporangium	169	83
		Streptomyces	168-169	83-86
	Alphaproteobacteria	Bradyrhizobium	167-169	83
		Caulobacter	174	91
		Rhizobium/Agrobacterium group	178	95
	Betaproteobacteria	Burkholderia	115-191	83-88
		Ralstonia	167	83
		Chromobacterium	175	92
	Gammaproteobacteria	Xenorhabdus	126-192	96-97
		Pseudomonas	167-195	83-93
		Halomonas	171	83
		Shewanella	174	91
		Photorhabdus	173-198	83-97
		Rheinheimera	177	86
		Pectobacterium	184-185	87-90
		Pseudoalteromonas	173	89
	Deltaproteobacteria	Hyalangium	175	85
		Myxococcus	177	87-91
PixC	Gammaproteobacteria	Pectobacterium	200-210	119-129
		Rahnella	182	101
		Yersinia	180-190	99-109
		Salmonella	183-202	100-119
ALL			**115-214**	**83-129**

Table 4: PixB Homologs identity, similarity, and weight

XNC1_PixB Homology					
	% ID	% Similarity	Length	Linker Length	kDa
Gram (+)	30 - 35	47 - 51	168 - 169	83 - 86	18.9 - 19.7
Alphaproteobacteria	30 - 32	46 - 48	167 - 178	83 - 95	18.7 - 19.2
Betaproteobacteria	27 - 36	43 - 51	115 - 191	83 - 90	13.2 - 21.2
Gammaproteobacteria	24 - 100	41 - 100	126 - 198	83 - 97	14.0 - 22 .0
Deltaproteobacteria	27 - 31	48 - 53	175 - 177	85 - 91	19.0 - 19.6
PixA	29 - 36	46 - 53	184 - 214	95 - 124	21.2 - 24.6
PixC	30 - 36	45 - 50	180 - 210	99 - 129	19.2 - 21.7
All	24 - 100	41 - 100	115 - 214	83 - 129	13.2 - 24.6
All Without PixA/C	24 - 100	41 - 100	115 - 198	83 - 97	14.0 - 22.0

Table 5: PixB homolog LD residue analysis by class. Serine and threonine percent average residue is shaded. Percent serine and threonine is highlighted yellow.

Linker Domain Residue Analysis by Class	Gram (+)	Alphaproteo-	Betaproteo-	Gammaproteo-	Deltaproteo-	All Classes
	% Residues of Avg	% Residues of Avg	% Residues of Avg	% Residues of Avg	% Residues of Avg	Average
Ala	2	7	4	4	5	5
Arg	6	3	5	4	3	4
Asn	5	6	4	6	6	5
Asp	6	4	5	6	4	5
Gln	5	4	3	4	7	5
Glu	6	4	5	3	1	4
Gly	6	8	7	7	8	7
His	1	2	1	2	0	1
Ile	5	6	6	5	5	5
Leu	7	7	5	7	7	7
Lys	4	2	5	3	2	3
Met	1	2	0	1	1	1
Phe	5	4	5	5	5	5
Pro	7	7	8	6	7	7
Ser	9	8	7	9	8	8
Thr	7	11	8	8	12	9
Tyr	7	7	6	6	5	6
Val	8	7	7	8	8	8
Trp	3	2	4	2	2	3
Cys	1	1	3	1	1	1
S/T %	16	19	15	18	20	17
Greatest % individual Residue and Number	Serine / 9	Threonine / 11	Threonine / 8	Serine / 9	Threonine / 12	Threonine / 9

Table 6: Average percentage of LD residue type of PixB homologs by class

	Gram (+)	Alphaproteo	Betaproteo	Gammaproteo	Deltaproteo	Average
Hydrophobic	37	42	40	39	41	**40**
Polar	40	44	39	43	48	**43**
Basic	12	7	11	8	6	**9**
Acidic	12	8	10	10	6	**9**

Table 7: Conserved domain analysis of PixB homologs by class. CD1 and CD2 residue analysis. Residue conversions that occur in each class is shown. Residues that convert to similar residues are marked with an asterisk.

CD Domain Residue Analysis by Class	Gram (+)	Alphaproteo-	Betaproteo-	Gammaproteo-	Deltaproteo-	All Classes
CD1 Sequence	CD1 conservation					
G	(N*), (R*)	(E)	(N*), (D)	(L), (A), (D)	(N*), (S*)	(N*), (R*), (E),(D),(L), (A),(S*)
D				(S)	(N)	(S),(N)
X	X	X	X	X	X	X
X	X	X	X	X	X	X
R1		(H*)	(Y)			(H*),(Y)
W				(F*)		(F*)
R2			(A)	(H*), (T), (S)		(A),(H*), (T), (S)
X	X	X	X	X	X	X
X	X	X	X	X	X	X
S/T			(N*)			(N*)
CD2 Sequence	CD2 conservation					
G				(A)	(S*)	(A), (S*)
Y/C						X
X	X	X	X	X	X	X
X	X	X	X	X	X	X
W				(M*), (H)		(M*), (H)
D				(P)		(P)
P			(A*)			(A*)
# Conserved Position changes	1	2	5	7	3	10
% of Residue changes that retain Similarity	100%	50%	50%	25%	75%	43%
ORDER OF CD CONSERVATION= Gram (+)>Alpha>Delta>Beta>Gamma						

While a PixB homolog has also been found in numerous genera and classes of bacteria, no PixA homolog has been located outside the *Xenorhabdus* genus. Additionally, the majority of methionine content that PixA is noted for occurs within the LD and positioned towards the N-terminus of the domain. The LD in PixA is expanded ranging from 93 to 124 residues. This is most notable in *Xenorhabdus sp. NBAII* XenSa04 that contains methionine rich repeats (Fig 5). The CD1 and CD2 of the PixB motif is also less well conserved within PixA homologs compared to other PixB homologs.

AidA homology

Among the PixB sequences identified is the "nematicidal protein" AidA (34% identity) in both *Ralstonia* and *Burkholderia* (Huber et al., 2004) (Fig 6). AidA is the only PixB-type protein identified that had been studied previously. The CD1 and CD2 regions of AidA shared 75% and 57% identity, respectively, with CD1 and CD2 of *X. nematophila* PixB. RaptorX secondary structure predictions are based on template folding against similar proteins whose crystal structure is known. Both PixB and PixA were templated against AidA' an AidA homolog present immediately adjacent to *aidA* perhaps because of a gene duplication event. Structural alignment of *X. nematophila* PixB and the *Ralstonia* AidA yield a TMscore of 0.79 predictive (>90% chance) the proteins share a similar fold.

MSKMIDVLMGVRTEDIMNDFGKLSKDINSPVSMWSADVFKKYVTYIHEDETMTSVDSSGNLMVK
CDVGDVIRMNLITLNNNPDYSVAMMKMDPMMKMDSMMKMDSMMKMDSMMKMDSMMKMD
PMMKVMQGDMVQMLEPHKMVASTKYMPVVDVRNPMDVDMQKVKSCYLMTRVNMMPPMGT
MRVMYYECILGIYKEGMLQGYIAAEPGLV
LDNR

Figure 5: The methionine rich PixA sequence of *Xenorhabdus NBAII XenSa04*. *Xenorhabdus NBAII XenSa04* CD1 (), LD (), and CD2 () are highlighted. Methionine rich repeats occurring within the LD are underlined.

```
AidA      MSRVTDVLVSFDTETILKKYPNPSKNPAAPTLIDW--HYVYMVTNQDNVISGQAGGELDL
PixB      MSNVIDILVIIDAQSIMRDFGIISPNPDAPTFLGYSDKYIYMLTRPEYVKQGQGSSHLSI
          **.* *:** :*:::*::.:   * ** ***:: :  :*:**:*. : * .**....*.:

AidA      KAQVGDLIRWRETSLSLGFENQVVFYKFIG--NVGNDLISTPTPRVAEASIPVPNTSKPE
PixB      NAKNGDVIRWRGVSLSKDSEYSAALLKLRAVSRDASSFFTVPMVKPLHSYIPVLKSDNPV
          :*: **:**** .***   * ...: *: .  . ...::::.*  :  .: *** ::.:*

AidA      V------PTCQKVANYYWSSECLKVGR------VTYHFQFQIIDRNCQSQGCFSWDPFIS
PixB      VNPDLLDVDIQPTVNYYWETSVKQMPPPGTAVTEYYTFKVGIYD-NGVLKGYVTWDPGIV
          *         * ..****.:.  ::          * *:. * * *   :* .:*** *

AidA      IHN-
PixB      LSNN
          : *
```

Figure 6: Clustal alignment of *X. nematophila* PixB and *B. cenocepacia* AidA. The two sequences share a 34% identity. CD1 (__) GDXXRWRXX(S/T) and CD2 (__) G(Y/C)XXWDP are highlighted.

Identification of third pix variant pixC

Of note is the discovery of a third Pix variant (PixC). PixC was identified in multiple genera including *Pectobacterium, Yersinia, Salmonella,* and *Rahnella* with sequence alignments showing identity ranging from 60% to 99% (Fig. 7). Identity based on alignments to XNC1 PixB is 29% to 30% with motif identity ranging from 26% to 29% (Fig. 8, 9). These organisms have all been described as associating with nematodes (Nykyri et al., 2013; Gengler et al., 2015; Acton, 2012; Vicente et al., 2013). BLAST values indicate these variants shares 45% to 50% similarity to PixB (Table S4). Secondary structure alignments of *X. nematophila* PixB and the *P. carotovorum* PixC variant (WP_052325947) sequences yield a TMscore of 0.713 predictive (90% chance) that the two proteins share a similar fold (Fig. 10) (Wang et al. 2013; Wang et al. 2011). One PixC homolog (WP_020833590) was located on a 78193 bp plasmid isolated and sequenced from the human pathogen *Salmonella enterica* subsp. enterica serovar Bareilly str. CFSAN000189 (Timme et al., 2013). This homolog was flanked by transposase and inverse autotransporter beta-barrel genes. The plasmid shares a 69% identity (2% coverage) with the *X. nematophila* 155327 bp plasmid XNC1_p (FN667743.1). Like PixA, PixC is thought to be specific to gammaproteobacteria and demonstrates significant residue enrichment within the LD. Unlike PixA, the PixC variant LD is enriched with glycine instead of methionine. These glycine rich repeats are positioned towards the C-terminal end of the LD. Like PixA the LD is expanded ranging from 99 to 129 residues with varying degrees of enrichment. Both CD1 and CD2 are well conserved in this variant apart from a (x) deletion in CD2 resulting in the PixC CD2 consensus GYxWDP.

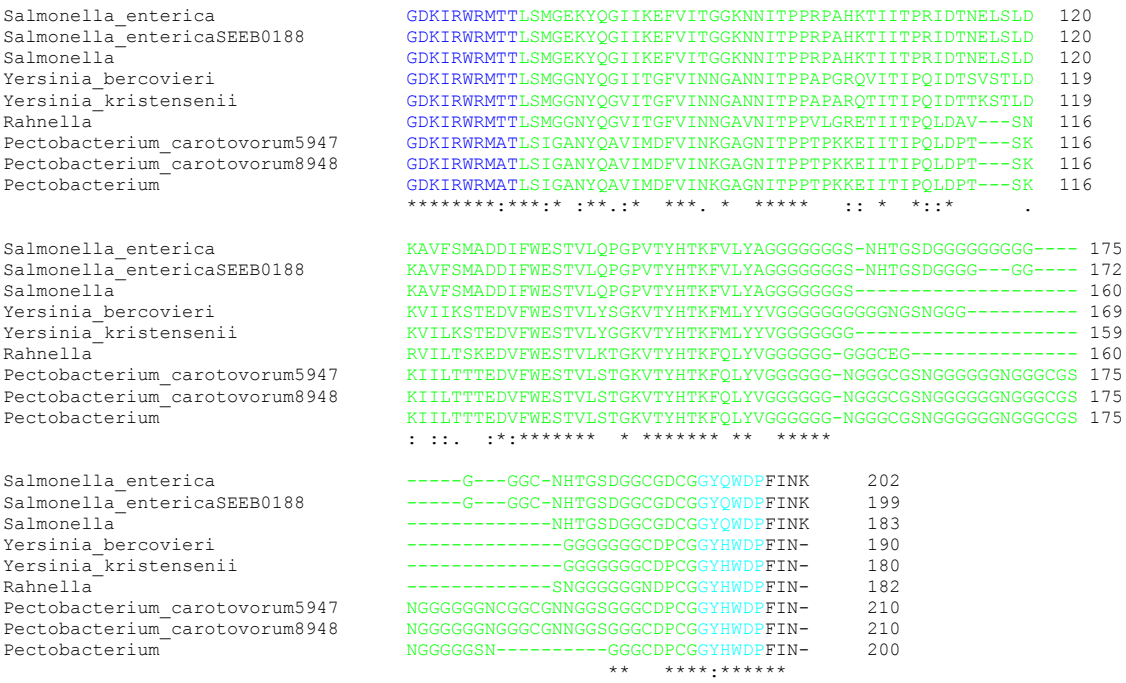


Figure 7: clustal 2.1 sequence alignments of the PixB motif in all PixC variants identified. CD1 (■), LD (■), and CD2 (■) regions are highlighted below alignment. All PixC sequences contain a (x) deletion in CD2. Sequence Alignments reveal a 60% to 99% identity.

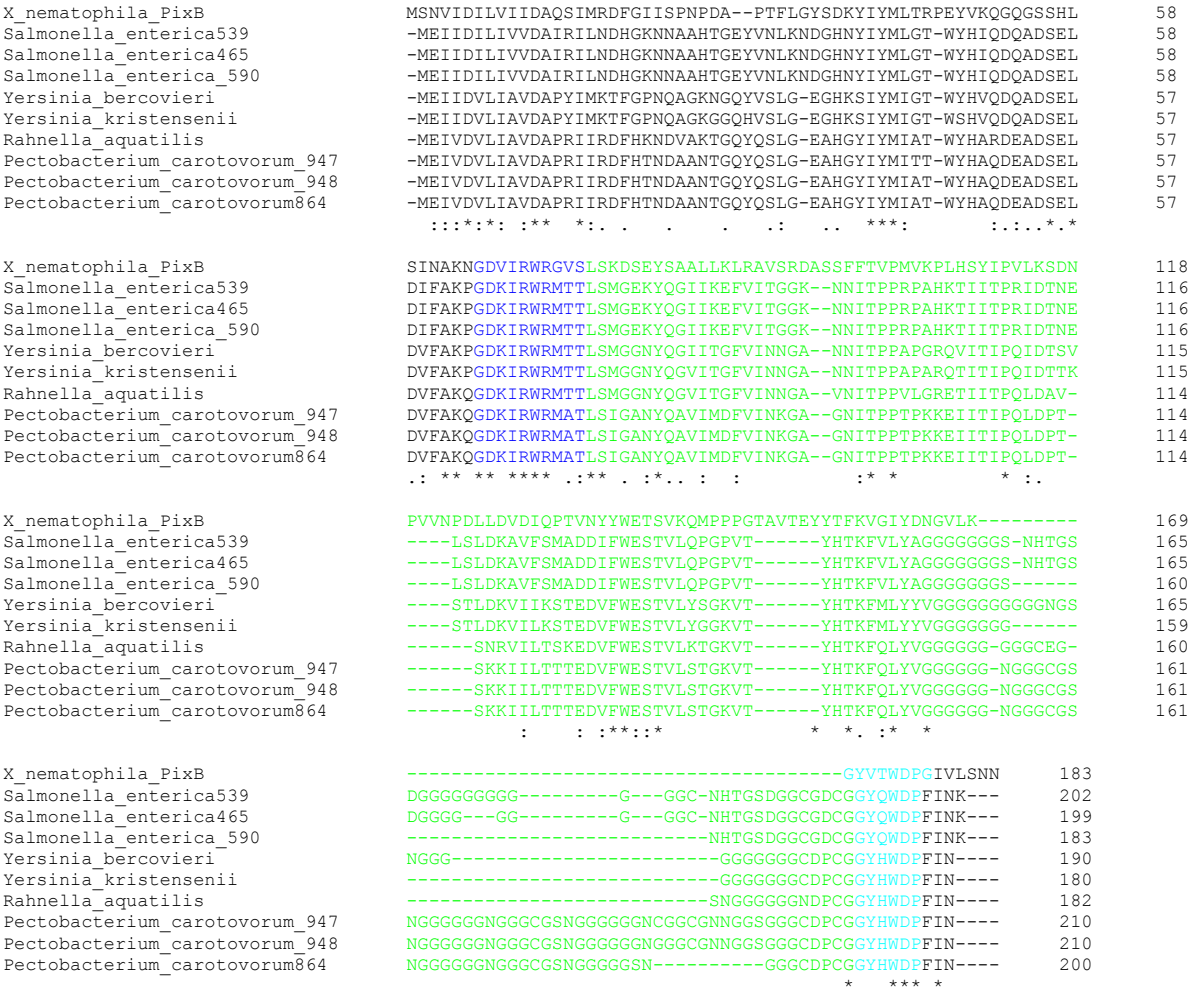


Figure 8: Clustal 2.1 sequence alignments of PixC and *X. nematophila* PixB. CD1 (), LD (), and CD2 () regions are highlighted below alignment. Sequence Alignments reveal a 29% to 30% identity.

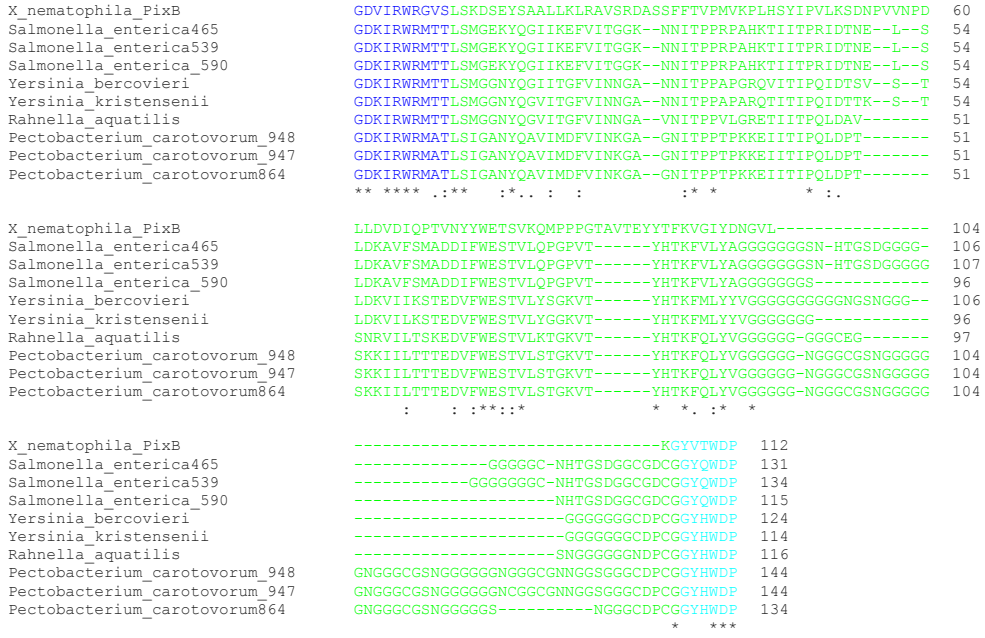


Figure 9: Clustal 2.1 sequence alignments of the PixB motif in PixC variants and *X. nematophila* PixB. The motif domains CD1 (), LD (), and CD2 () are colored. Sequence Alignments reveal a 26% to 29% identity.

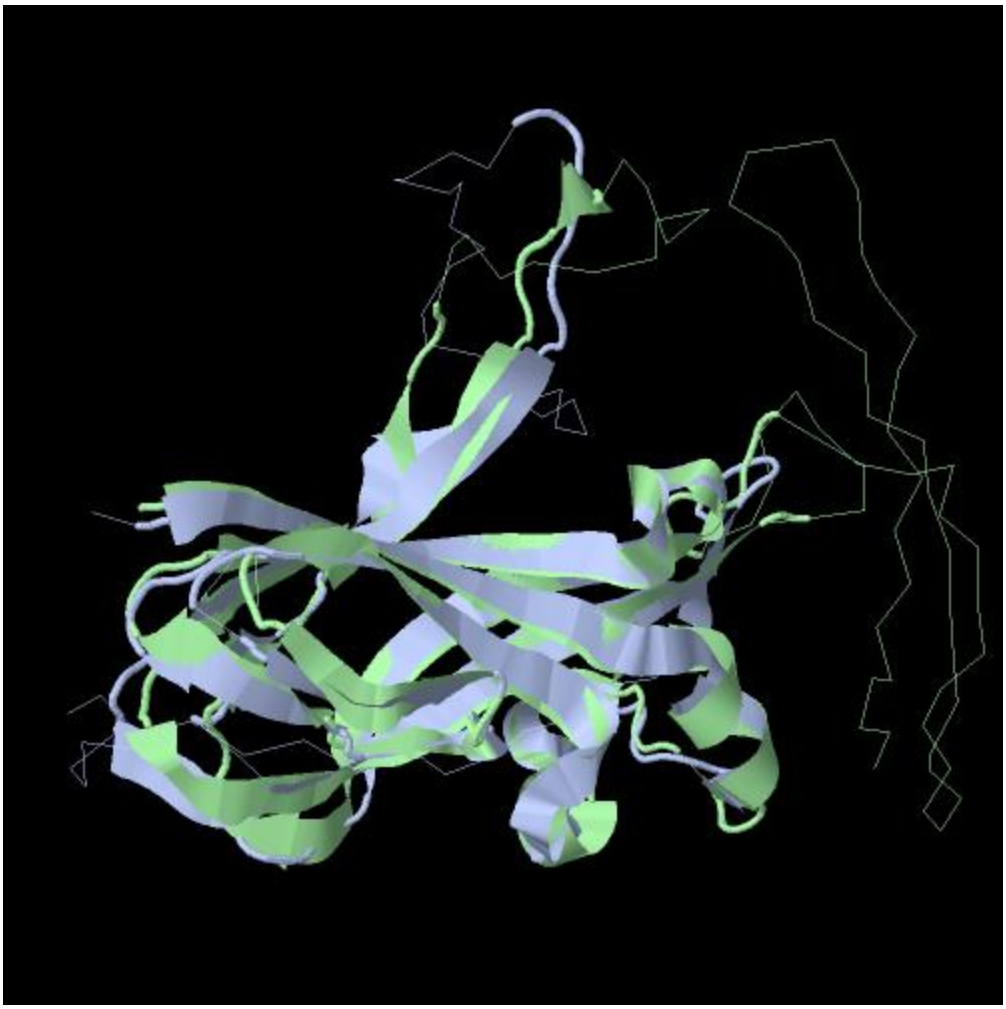


Figure 10: Secondary structure alignment of PixB and PixC. Secondary structure alignment of *X. nematophila* PixB (__) and the *P. carotovorum* PixC (__) variant (WP_052325947). Alignment yields a TMscore of 0.713 which is predictive (>90% chance) that the two proteins share a similar fold.

Glycine rich proteins have been described as adhesins and calcium binding proteins (Espitia et al. 1999, Linhartová et al., 2010). The predicted binding properties of the glycine rich PixC variant as well as the PixB and PixA proteins of *Xenorhabdus* were analyzed to determine if any binding properties are shared among these variants (Källberg et al., 2012). *Xenorhabdus* PixB, PixA and PixC variant proteins were all predicted to bind calcium as well as chlorine and glycerol. Additionally, some proteins were predicted to bind beta-D-glucose, heme, copper, iron, N-acetyl-glucosamine, isopropyl alcohol, sodium and zinc (Table 8, S6, S7, S8).

BLAST analysis of PixB also indicates 41% to 49% similarity to NCBI annotated "DNA directed RNA polymerase subunit beta" (DDRP) proteins in six classes of bacteria including, alphaproteobacteria, betaproteobacteria, gammaproteobacteria, flavobacteria, sphingobacteria, and cytophagia. These proteins bear close resemblance to other PixB homologs with sequence lengths ranging from 182 to 193 residues and molecular weights of 19.7 kDa to 21.2 kDA. The PixB motif is present with less conservation about CD1 (R_1) and (R_2) residues and an aromatic conversion of tryptophan to phenylalanine. The LD range is 88 to 94 residues with CD2 well conserved (Table 9). Clustal alignment of the PixB motif in these proteins results in identity values of 34% to 79% to each other, 20% to 28% to XNC1 PixB and 16% to 18% to XNC1 PixA (Fig. 11, 12). Residue analysis conducted on each of the nine DDRP homologs identified indicates the greatest average LD residue count occurred with asparagine (11%) glycine (8%) and valine (8%). 45% of the LD is polar (Table 10).

Table 8: Predicted binding properties of PixA, PixB, and PixC proteins.

Binding Predictions											
Strain	**Chlorine**	**Calcium**	**Glycerol**	**BETA-D-GLUCOSE**	**Heme**	**Copper**	**Iron**	**N-ACETYL-D--GLUCOSAMINE**	**ISOPROPYL ALCOHOL**	**Sodium**	**Zinc**
>Pectobacterium_carotovorum_947_PixC	x	X	x								
>Pectobacterium_carotovorum_948_PixC	x	X	x	x							
>Pectobacterium_carotovorum864_PixC	x	X	x								
>Rahnella_aquatilis_PixC	x	X	x	x							
>Yersinia_bercovieri_PixC	x	X	x								
>Yersinia_kristensenii_PixC	x	X	x								
>Salmonella_enterica539_PixC	x	X	x								
>Salmonella_enterica465_PixC	x	X	x								
>Salmonella_enterica_590_PixC	x	X	x	x							
>XNC1_PixB	x	X	x						x		
>XBI1_PixB	x	X	x		x	x				x	
>XNFF1_PixB	x	X	x			x			x		
>XCR1_PixB	x	X	x	x							
>XBKQ1_PixB	x	X	x				x				
>XCR1_PixA	x	X	x		x						
>XNC1_PixA	x	X	x	x	x						
>XSR1_PixA	x	X	x			x					
>XDD1_PixA	x	X	x			x					
>XBI1_PixA	x	X	x			x	x	x			
>XBKQ1_PixA	x	X	x			x	x	x			
>XBP1_PixA	x	X	x		x	x					
>XiS1v1_PixA	x	X	x			x	x				
>XBFFL1v2_PixA	x	X	x			x			x		

Table 9: PixB homologs designated DNA directed RNA polymerases. Homologs with 95% or greater coverage were selected for analysis

Class	Genus	Species	ACCESSION #	% ID	% Similarity	E-value	Length	Linker Length	kDa	NCBI annotations and CD Notes:
Flavobacteria	Chryseobacterium	UNC8MFCol	WP_027371878.1	27	41	1.00E-07	188	94	20.9	GDRVSFRGSS// SYCCWDP
Gammaproteobacteria	Pseudoalteromonas	flavipulchra	WP_010605470.1	19	43	1.00E-05	183	93	19.7	GDNVSFSGTS // GYYQWDP
Sphingobacteriia	Mucilaginibacter	PAMC 26640	WP_067189497.1	25	47	1.00E-13	183	91	20.2	GDTVWFRGTS// GYYFWDP
Cytophagia	Cytophagales	bacterium B6	WP_022829892.1	26	41	2.00E-11	182	88	20.2	annotated PixA // GDVIRWYAVS// GYYYWDP
Gammaproteobacteria	Rhodanobacter	115	WP_037099989.1	24	46	2.00E-10	184	91	20.3	annotated PixA // GDLVAFTGVS // GYYYWDP
Betaproteobacteria	Collimonas	arenae	WP_061533258.1	26	49	2.00E-09	184	91	20.2	GDTVAFTGVS//GYYYWDP
Alphaproteobacteria	Rhizobium	leguminosarum	WP_027690943.1	21	43	1.00E-05	193	91	21.2	GDNVSFTGVS // GYYCWDP
Betaproteobacteria	Hydrogenophaga	taeniospiralis	WP_068174534.1	24	43	5.00E-07	185	93	20.1	GDYASFTGTS// GYFYWDP
Betaproteobacteria	Burkholderia	cepacia complex	WP_006485654.1	24	46	4.00E-08	185	92	20.3	Burkholderia cenocepacia J2315 chromosome 2// GDRVQFTGVS // GYYWWDP

Figure 11: Clustal % ID matrix of the PixB motif in DDRP homologs and *X. nematophila* PixA and PixB. PixB and PixA are shaded.

	1	2	3	4	5	6	7	8	9	10	11
1: Cytophagales	100.00	30.77	33.65	35.92	36.54	31.73	33.65	36.54	33.65	27.66	18.28
2: Chryseobacterium		100.00	44.86	42.45	40.19	46.73	49.53	44.86	44.86	28.28	16.33
3: Pseudoalteromonas			100.00	57.41	56.88	57.94	60.75	61.68	59.81	20.20	17.35
4: Burkholderia				100.00	52.78	54.72	66.04	61.32	64.15	24.24	18.37
5: Hydrogenophaga					100.00	64.49	62.62	57.01	61.68	23.23	16.33
6: Mucilaginibacter						100.00	61.68	60.75	63.55	22.68	17.71
7: Rhizobium							100.00	65.42	66.36	21.65	17.71
8: Collimonas								100.00	79.44	26.80	15.62
9: Rhodanobacter									100.00	21.65	17.71
10: XNC1_PixB										100.00	32.08
11: XNC1_PixA											100.00

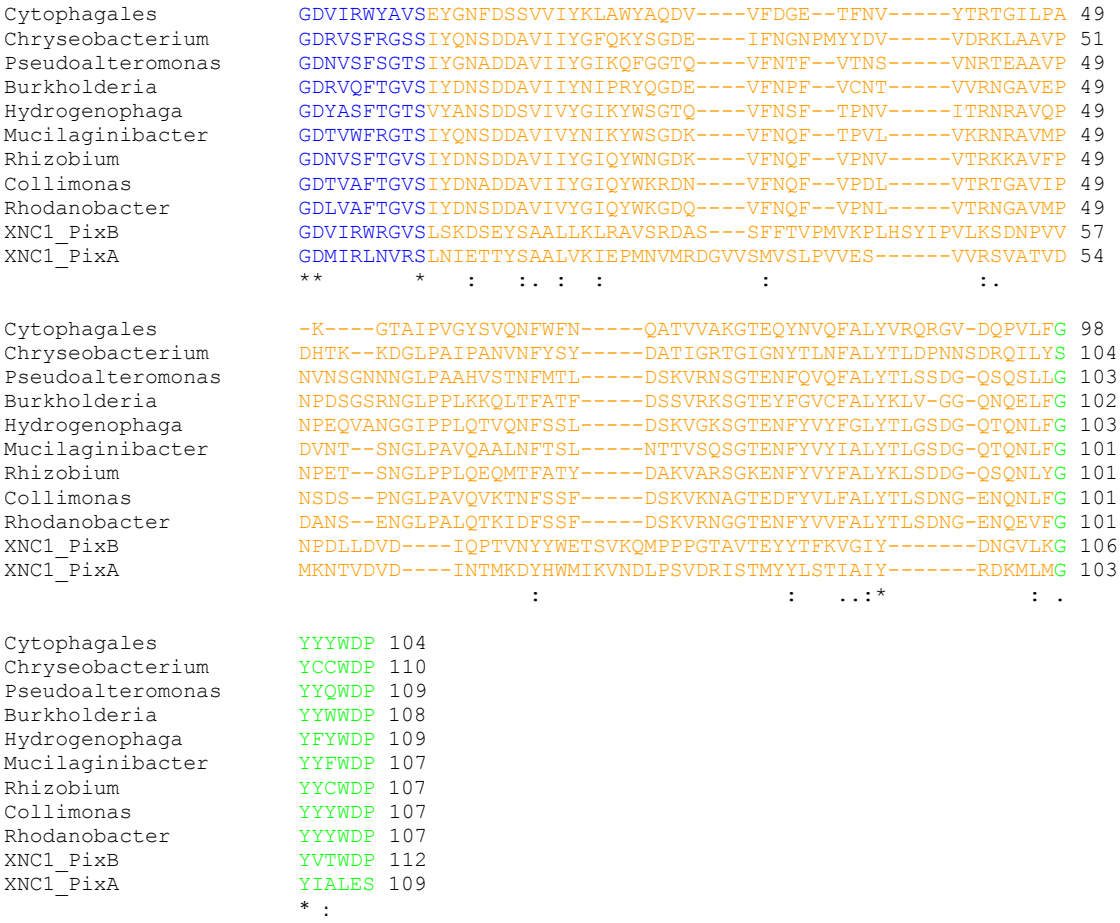


Figure 12: Clutstal alignment of the PixB motif in DDRP strains and *Xenorhabdus* PixA and PixB. Domain regions are highlighted CD1 blue, LD orange and CD2 green.

Table 10: Residue analysis of the PixB motif in DDRP homologs

	Burkholderia cepacia complex	*Chryseobacterium UNC8MFCol*	*Pseudoalteromonas flavipulchra*	*Mucilaginibacter sp. PAMC 26640*	*Cytophagales bacterium B6*	*Rhodanobacter sp. 115*	*Collimonas arenae*	*Rhizobium leguminosarum*	*Hydrogenophaga taeniospiralis*	Avg.	% Residues of Avg	CD1 conservation
Ala	4	7	7	6	7	5	6	6	3	6	7	**G**
Arg	4	3	2	2	3	2	2	2	2	3	3	**D**
Asn	8	9	12	10	5	10	10	8	9	10	11	**X**
Asp	6	10	4	5	4	8	9	7	4	7	7	**X**
Gln	4	3	6	6	7	5	4	6	7	4	5	**R1 to (S) (W) (A) (Q)**
Glu	4	1	2	1	3	4	2	3	2	2	3	**W to (F)**
Gly	8	7	8	5	7	7	5	5	9	8	8	**R2 to (S) (Y) (T)**
His	0	1	1	0	0	0	0	0	0	1	1	**X**
Ile	4	8	4	4	3	4	5	4	4	5	6	**X**
Leu	4	6	6	7	4	5	6	5	5	5	6	**S/T**
Lys	6	4	2	3	3	3	4	6	3	4	4	**CD2 conservation**
Met	0	1	1	1	0	1	0	1	0	1	1	
Phe	7	4	6	5	7	7	7	6	6	6	6	**G to (S)**
Pro	6	5	2	3	3	3	4	5	5	4	5	**Y/C**
Ser	6	4	9	7	3	6	6	5	9	6	7	**X**
Thr	4	5	8	8	6	4	5	4	7	6	6	**X**
Tyr	5	10	3	6	7	5	5	8	6	6	7	**W**
Val	9	5	9	10	13	10	9	8	10	8	8	**D**
Trp	0	0	0	1	2	1	1	1	1	0	0	**P**
Cys	2	0	0	0	0	0	0	0	0	1	1	**D**
Residue#	91	93	92	90	87	90	90	90	92	92		**P**

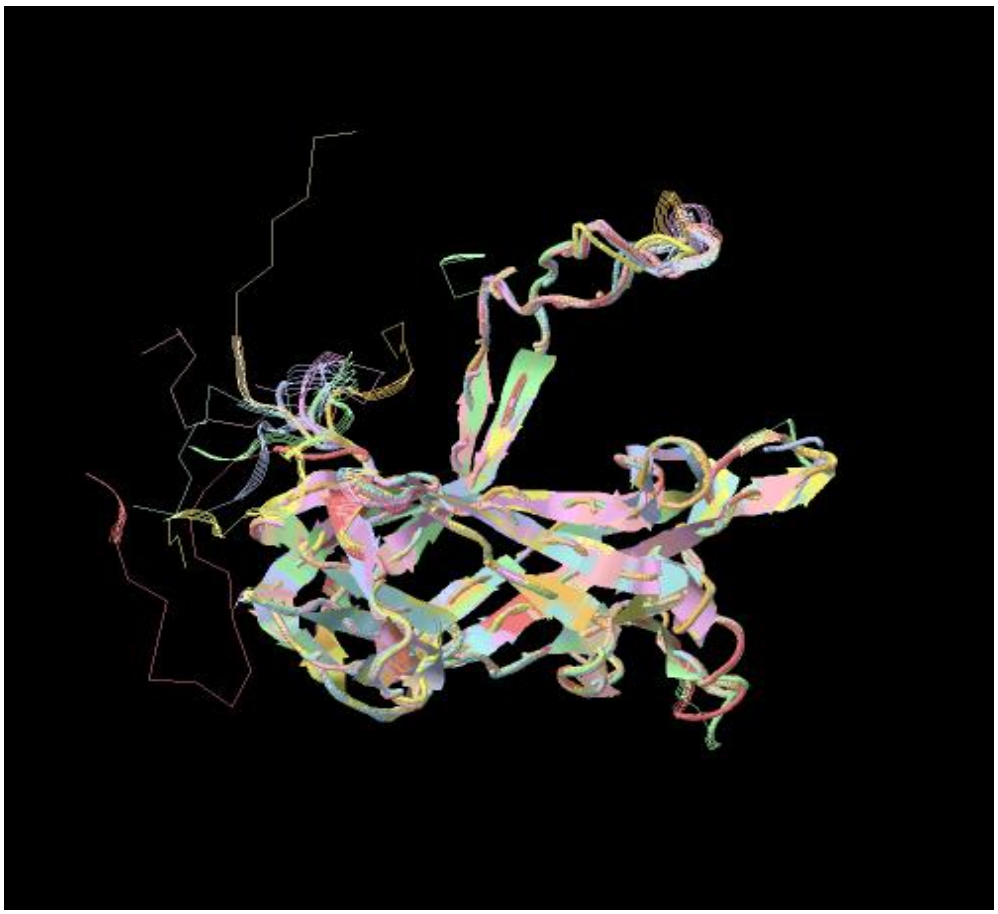


Figure 13: Secondary structure alignments of DDRP PixB homologs identified in table 9. Alignments result in a Tm score of 0.935, predictive (>90% chance) that the proteins share a similar fold.

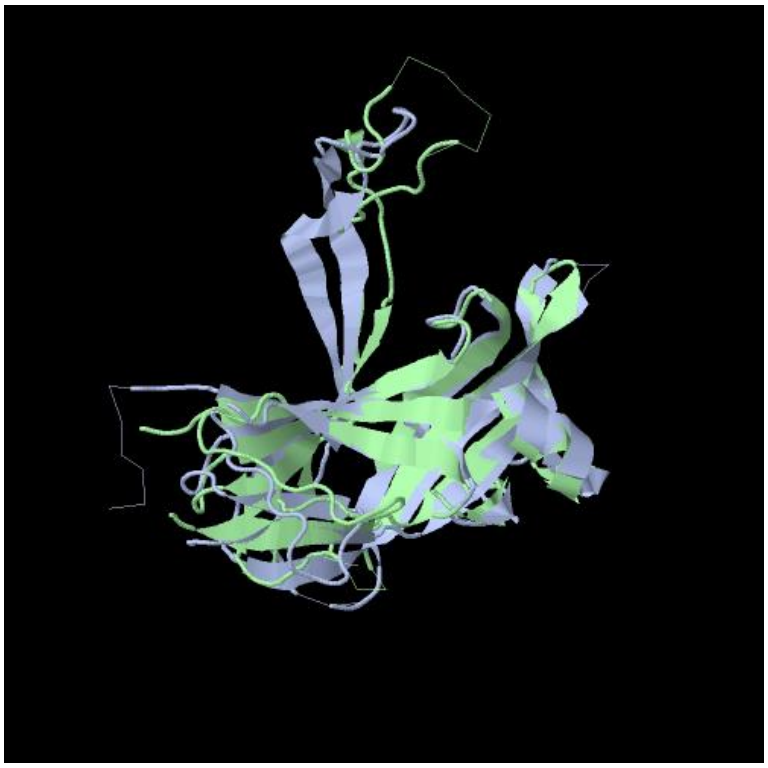


Figure 14: Secondary structure alignment of PixB and homologous DDRP protein. Secondary structure alignment of *X. nematophila* PixB (▇) and the *Mucilaginibacter* DDRP protein WP_067189497.1 (▇). Alignment yields a TMscore of 0.787 which is predictive (>90% chance) that the two proteins share a similar fold.

Table 11: Secondary structure alignment scores of DDRP homologs to *X. nematophila* PixB. If the TMscore is greater than 0.6, it is considered very likely (90% of chance) that the proteins share a similar fold.

Strain	Alignment to XNC1 PixB
>Burkholderia cepacia complex_DDRP	TMscore: 0.785
>Chryseobacterium_DDRP	TMscore: 0.769
>Collimonas arenae_DDRP	TMscore: 0.780
>Cytophagales bacterium B6_DDRP	TMscore: 0.776
>Hydrogenophaga taeniospiralis_DDRP	TMscore: 0.783
>Mucilaginibacter sp. PAMC 26640_DDRP	TMscore: 0.787
>Pseudoalteromonas flavipulchra_DDRP	TMscore: 0.776
>Rhizobium leguminosarum_DDRP	TMscore: 0.780
>Rhodanobacter sp. 115_DDRP	TMscore: 0.780

Secondary structure alignments of XNC1 PixB and the DDRP homologs results in Tm scores ranging 0.769 to 0.787 and a score of 0.935 when aligned to each other, all scores are predictive (>90% chance) that the proteins share a similar fold (Fig 13, 14) (Table 11). Binding analysis also indicates predictions for chlorine, calcium and glycerol in all homologs as well as beta-D-glucose in all but two DDRP's as well as N-acetyl-D-glucosamine, sodium, and zinc in at least one homolog (Table 12, S9). All DDRP's except Cytophagia demonstrate an aromatic substitution in CD1 of (W) to (F). This substitution in also present in some gammaproteobacteria. CD1 comparisons of DDRP, Gram positive and proteobacteria PixB homologs also indicate the greatest amount of substitutions occurs with (G) followed by (R_1) and (R_2) indicating the most evolutionarily conserved CD1 sequence to be DxxxΩxxx(S/T) were Ω = any aromatic residue (Aasland et al., 2002).

Given the occurrence of PixB across such a broad range of prokaryotic organisms the homology search was extended to the NCBI eukaryotic yeast taxid using the most broadly conserved domain sequence CD2 and a position-specific iterated BLAST search. A putative PixB homolog was identified in the fungal parasite *Metschnikowia bicuspidata*. *M. bicuspidata* is an aquatic pathogen of crustaceans that gains access to the host via the gut and migrates to the hemolymph where it reproduces (Ebert et al., 2000). The protein (GenBank: OBA19327.1) is designated a GPI mannosyltransferase 3 and contains a perfect C-terminal CD2 consensus sequence as well as an N-terminal DxxxΩxxx(S/T) CD1 sequence. The LD is 370 residues long and is enriched (33%) with leucine (14%), phenylalanine (11%), and serine (8%). GPI proteins undergo post-translational cleavage of the C-terminus by GPI transamidase and are subsequently covalently attached to a membrane bound GPI anchor.

Table 12: Binding Predictions for PixB DDRP homologs

Binding Predictions							
Strain	**Chlorine**	**Calcium**	**Glycerol**	**BETA-D-GLUCOSE**	**N-ACETYL-D-GLUCOSAMINE**	**Sodium**	**Zinc**
>Burkholderia cepacia complex_DDRP	x	x	x				
>Chryseobacterium_DDRP	x	x	x	x			
>Collimonas arenae_DDRP	x	x	x				
>Cytophagales bacterium B6_DDRP	x	x	x	x			
>Hydrogenophaga taeniospiralis_DDRP	x	x	x	x		x	X
>Mucilaginibacter sp. PAMC 26640_DDRP	x	x	x	x	x		
>Pseudoalteromonas flavipulchra_DDRP	x	x	x	x			
>Rhizobium leguminosarum_DDRP	x	x	x				
>Rhodanobacter sp. 115_DDRP	x	x	x	x			

The site at which the protein is cleaved is known as the omega site. GPI anchored proteins can function as enzymes, membrane receptors, surface antigens and adhesion molecules and can also be involved in signaling processes, immunomodulation and host-pathogen responses. PredGPI predicted omega cleavage site for protein OBA19327.1 is located immediately adjacent to N-terminal side of CD2 (Fig. S2) (Pierleoni, et al., 2008).

Phylogenetic analysis of PixC and DDRP Proteins

The evolutionary history of selected PixB homologs including PixC and DDRP homologs was inferred by using the Maximum Likelihood method based on the Poisson correction model. The tree with the highest log likelihood (-29121.5002) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 5.1383)). The analysis involved 64 amino acid sequences. There were a total of 322 positions in the final dataset (Fig 15). The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The analysis involved 64 amino acid sequences. There were a total of 322 positions in the final dataset (Fig S3). Results indicate that like PixB and PixA, PixC homologs branch together. As in Figure 13 branching is not class specific indicative of horizontal gene transfer.

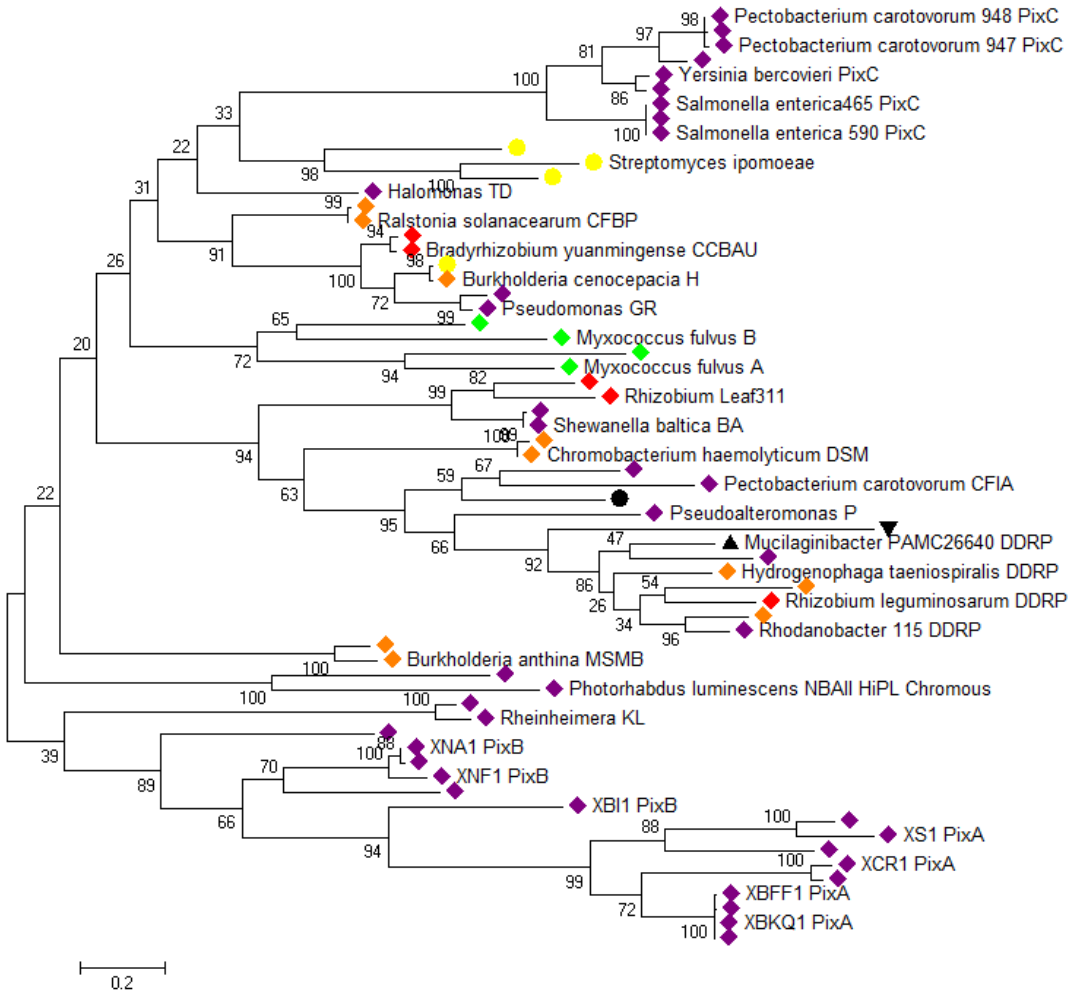


Figure 15: Molecular Phylogenetic analysis by Maximum Likelihood method of PixB homologs including the PixC and DNA directed RNA polymerase (DDRP) homologs. Bacterial classes are distinguished with colored diamonds: Gram positive-Yellow, alphaproteo-red, betaproteo-orange, gammaproteo-purple, deltaproteo-green, Flavobacteria- ▼, Sphingobacteria-▲, Cytophagia-●. Branching is not class specific suggesting PixB is acquired via horizontal gene transfer.

Chapter 4

Functional analysis of PixB

Transmission Electron Microscopy analysis of X. nematophila

Transmission electron micrographs of *X. nematophila* NMI1 cells cultured in LB, GM, or hemolymph reveal diverse changes in membrane structure. Inclusion bodies could be visualized in LB cultured cells; however, much fewer GM cultured cells seemed to contain inclusion bodies. Of those cells in GM cultures which could be found to contain inclusion bodies, inclusions appeared more diffuse and less organized than that seen in LB cultured cells. No inclusion bodies could be visualized in cells obtained from *M. sexta* hemolymph (Fig. 16).

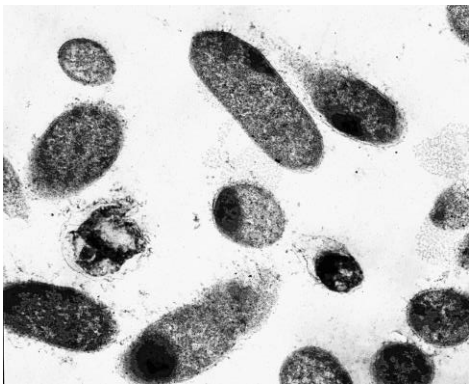
Protein expression

To confirm that both mutants were no longer producing a Pix protein, SDS-PAGE was performed on overnight cultures of both wt and mutants NMI1 and NMI2. Bands corresponding to the PixA and PixB proteins were visualized in the wt sample. As expected, the NMI1 sample lacked a 26 kDa band corresponding to PixA and NMI2 lacked a 22 kDa band corresponding to the PixB protein. However, in both mutant lanes protein expression of the remaining Pix protein seemed greater than wt (Fig. 17).

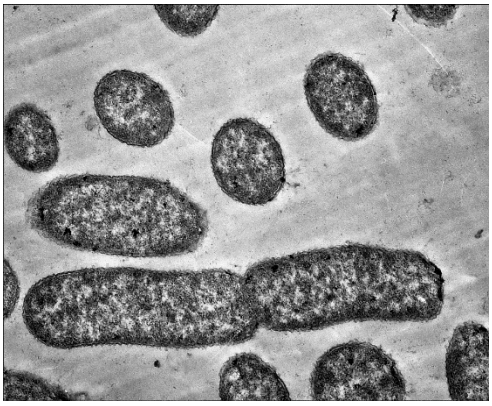
Phenotypic analysis

Results from phenotypic assays of both NMI1 and NMI2 were consistent with wt results (Table S10).

A. LB



B. GM



C. Hemolymph

Figure 16: Transmission electron micrographs of *X. nematophila* NMI1. Representative ultrathin sections of *X. nematophila* NMI1 cultured in either (A) LB, (B) GM, or (C) *M. sexta* hemolymph. Inclusion bodies are frequently visible in cells cultured in LB, very few inclusion bodies visualized in cells cultured in GM, no inclusion bodies were observed in cells cultured in hemolymph. All micrographs are 15000X magnification.

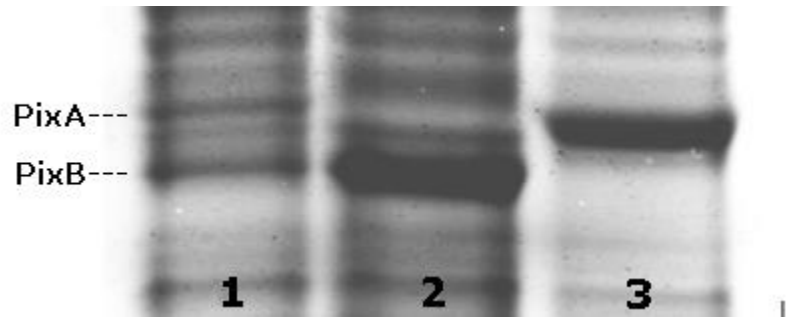


Figure 17. SDS-PAGE of *X. nematophila*, NMI1 and NMI2. SDS-PAGE was performed on overnight cultures of both wt (lane 1) and mutants NMI1 (lane 2) and NMI2 (lane 3). Bands corresponding to the PixA and PixB proteins were visualized in the wt sample. NMI1 sample lacked a corresponding PixA band at 26 kDa while NMI2 lacked corresponding PixB band at 22 kDa.

In vivo and in vitro competition

The previously reported variability and poor plating efficiency of *Xenorhabdus* was successfully reproduced leading to inconclusive results (Xu et al., 1989; Gotz et al., 1981).

Viability assay

To assess whether the expression of either *pixA* or *pixB* contributes to the long-term viability of *X. nematophila* live /dead staining of wt, NMI1 and NMI2 was performed over the course of 72 hours in both LB and GM. No significant difference in the percent viability of cells was observed in strains cultured in GM over the course of 72 hours (Fig. 18B). When the experiment was repeated with strains grown in LB media, a significant temporal increase in stationary phase viability was observed with the NMI2 mutant compared to wt and NMI1 (Fig. 18A). Results from three replicates were averaged. Mean LB viability of NMI2 compared to wt viability from hours 16 to 24 was statistically significant ($P < 0.05$ by Student's t-test). Overall viability of all strains increased in GM when compared to LB.

Nematode killing assay

No differences in growth or mortality of *C. elegans* between slow killing (NGM) and fast killing (PGS) plates were seen over the course of 48 hours.

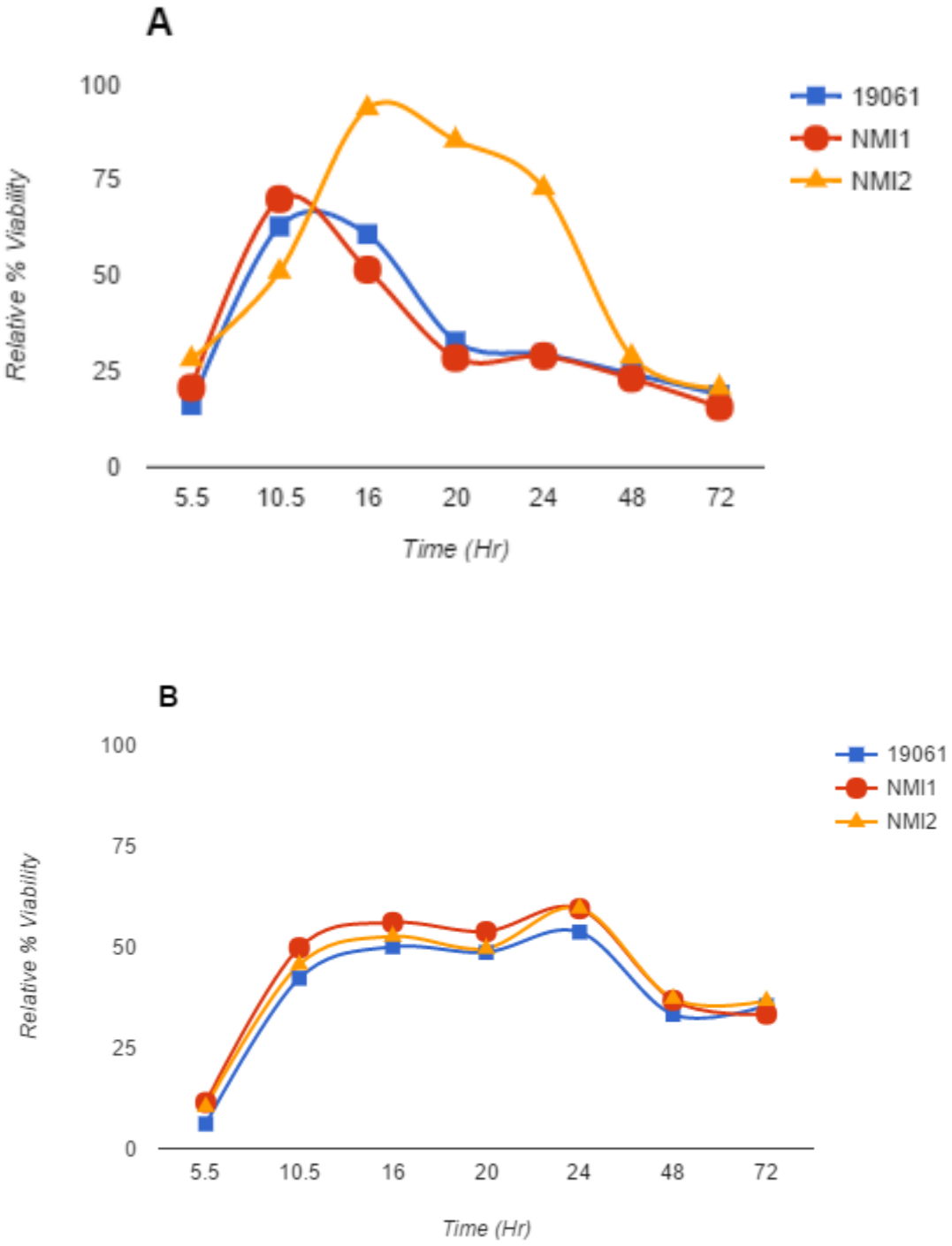


Figure 18. **Cell viability assay.** Effect of growth media on relative percent cell viability, assay was performed in triplicate for each condition and results averaged. (■) 19061 (●) NMI1 (▲) NMI2. (A) Cells grown in LB media. NMI2 cell viability hours 16 to 24 statistically significant ($P < 0.05$ by Student's t-test) (B) Cells grown in GM.

Determination of transcription start sites and 5'UTR analysis

Sequencing results from the *pixA* 5'RACE product reveal a transcription start site 90-91 bp upstream from the AUG start codon. The unspecificity of transcription start site residues is due to a native (G) residue present 91 bp upstream of the pixA 5'UTR that overlaps with the poly(G) tail of the 5'RACE product (Fig. 19A). Sequencing results from the *pixB* 5'RACE product yielded a plus one site approximately 89 bp upstream from the AUG start codon (Fig. 19B).

The 5'UTR of both genes was analyzed for potential promoter consensus sequences and transcription factor binding sites. No promoter consensus sequence was identified in the *pixA* -1 to -40 region. However, a well-conserved sigma sequence TTTACA_16_TATATT was identified 6 bp upstream of the *pixB* +1 transcriptional start, a strong indicator of either a sigma 70 or sigma 38 promoter sequence. Additionally, analysis of the -3 to +7 region indicates *pixA* has a GC content of 50%-60% while *pixB* has a GC content of 40%. BPROM analysis of the *pixB* 5' UTR indicates a potential OmpR binding sequence TCATATTT located in the -63 to -71 region (Solovyev, 2011).

Gene expression assay

Gene expression was analyzed in wt, NMI1 and NMI2 strains carrying either the *pixA* or *pixB* GFP reporter vector or the promoterless vector pProbeNT in both LB and GM. No significant difference in mean absorbance was observed between wt, NMI1, and NMI2 over the course of 24 hours in either LB or GM, suggesting that the expression of either protein does not cause a significant growth defect (Fig. 20).

A. *

TATTTAACTTTCATATTTTTATATATGTTATGACATTATTTTTGATATTTACATTTTATTTTATTTTGATATATTTAAATCT**T**GTAAGCA

-35 -10

B. **

ATATTTATCATGATAAATTTATTTTTCGGCGAGACTCTTTTTGTGGTTTCGGTTCAATAATTTTTTATTATTATATATTCC**GT**AAGCAGC

Figure 19: Transcription start sites determined by 5'RACE. 5' UTR of both *pixB* (A) and *pixA* (B). Transcription start sites (*) were determined by sequencing of 5'RACE products. A sigma 70 consensus sequence (TTGACA_TATAAT) was identified 6 bp upstream of the *pixB* transcription start site indicative of a strong promoter, the -10 and -35 sequences are underlined. No consensus for *pixA* -35 region could be identified. A potential -10 region for *pixA* is underlined.

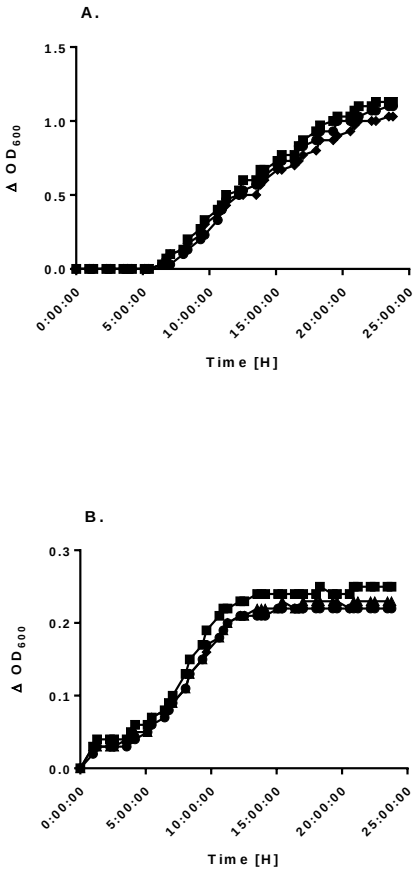


Figure 20: Mean absorbance of *X. nematophila*, NMI1, and NMI2. Mean absorbance of *X. nematophila* (●), NMI1 (■), and NMI1 (◆) in either LB (A) or GM (B). No signifigant difference in growth of either mutant compared to wt is seen in either LB or GM.

Doubling times for *X. nematophila* in LB is 4.12 hours and 3.02 hours in GM. Expression of *pixB* occurs during early mid log growth in LB and in GM expression begins during the transition to stationary phase. Expression of *pixA* occurs during early log growth in LB and during late log phase in GM (Fig. 21). By 16 hours of growth, *pixB* expression was significantly higher (by Student's t-test) in those cells grown in GM compared to those grown in LB (Fig. 22). Expression of *pixB* in NMI1 and NMI2 is also greater in GM than LB (Fig. 23). To determine if either gene was under the control of a RpoS promoter, both reporter constructs were transformed into a *rpoS* mutant strain of *X. nematophila* and the gene expression assay repeated. Results indicate no significant deviation in expression profiles for either *pixA* or *pixB,* suggesting neither gene is controlled by a sigma 38 promoter.

Doubling time for cells grown in hemolymph is 5.28 hours. Expression profiles of *pixA* and *pixB* in hemolymph indicate both genes are activated during mid log phase. Like LB and GM expression of *pixB* is greater than *pixA* in hemolymph (Fig. 24).

In vitro and in vivo flow cytometry analysis

To assess the relative *in vitro* population level expression of *pixA* and *pixB*, 24 hour wt cultures carrying the *pixA* or *pixB* reporters and the promoterless vector were diluted 1:500 in 0.2 μm filtered LB or GM and sampled at various time points. Gated cells were assessed for relative percent expression of either *pixA* or *pixB* and concentration (events/ml). Results indicate *pixA* expression occurs before *pixB* expression with fewer overall cells expressing *pixB* than *pixA* until late stationary phase (24 hours) where the

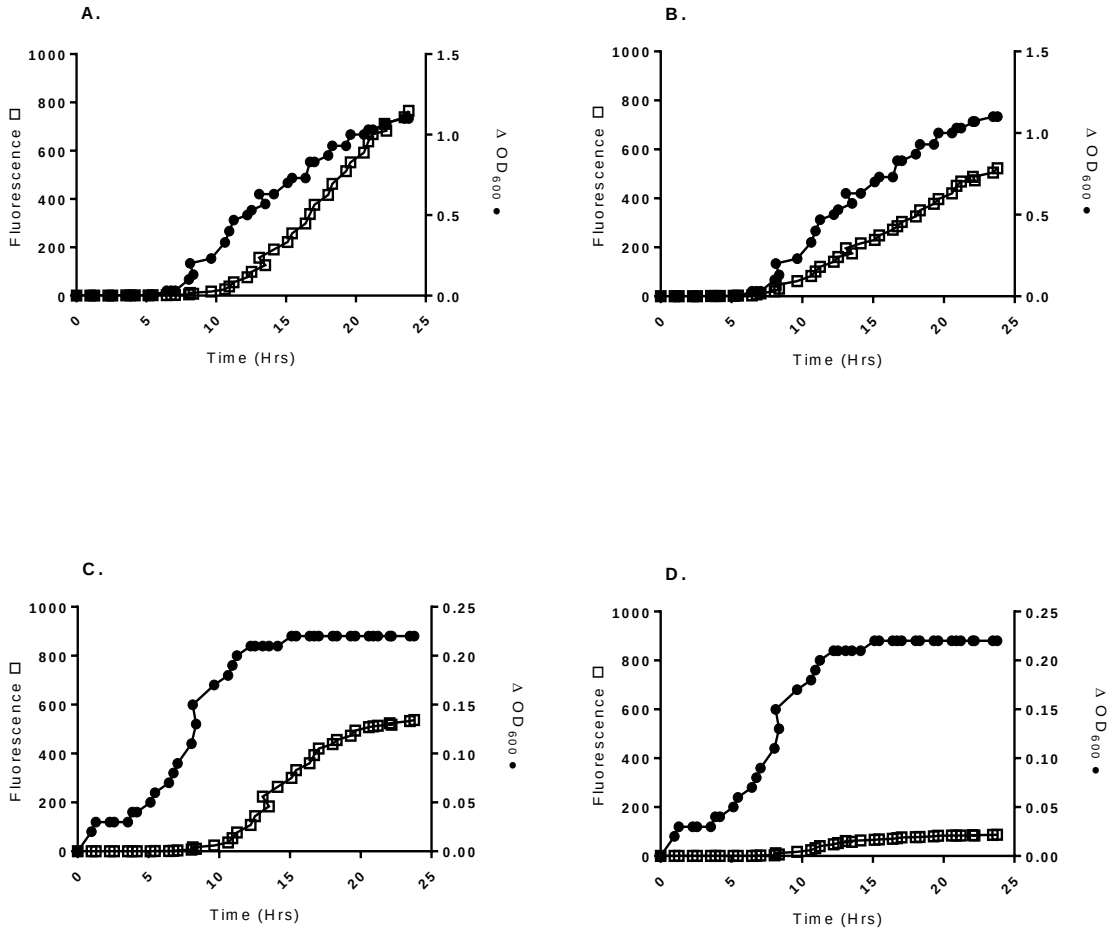
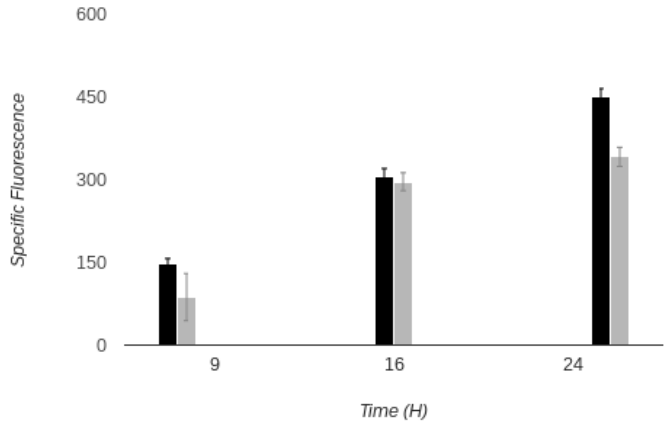


Figure 21: *X. nematophila pixB* (A,C) and *pixA* (B,D) expression in either LB (A,B) or GM (C, D). Both genes begin expressing during early log growth in LB culture and late log in GM culture.

A.



B.

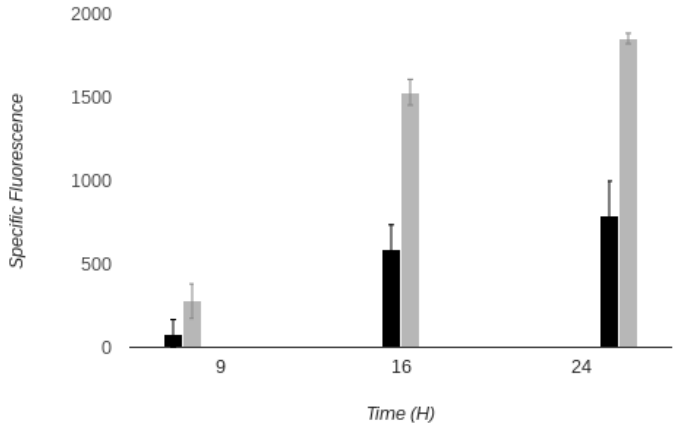


Figure 22. *X. nematophila* expression of *pixA* and *pixB* in either LB or GM. *X. nematophila* expression of *pixA* and *pixB* in either LB (dark columns) or GM (light columns). By nine hours no significant difference in expression occurs. After 16 hours of growth, *pixA* expression (A) in LB begins to increase relative to that in GM until a significant difference occurs by 24 hours of growth. By 16 hours of growth, *pixB* expression (B) was significantly higher in those cells grown in GM compared to those grown in LB. Error bars represent standard error of means. Statistical significance determined by Student's t-test.

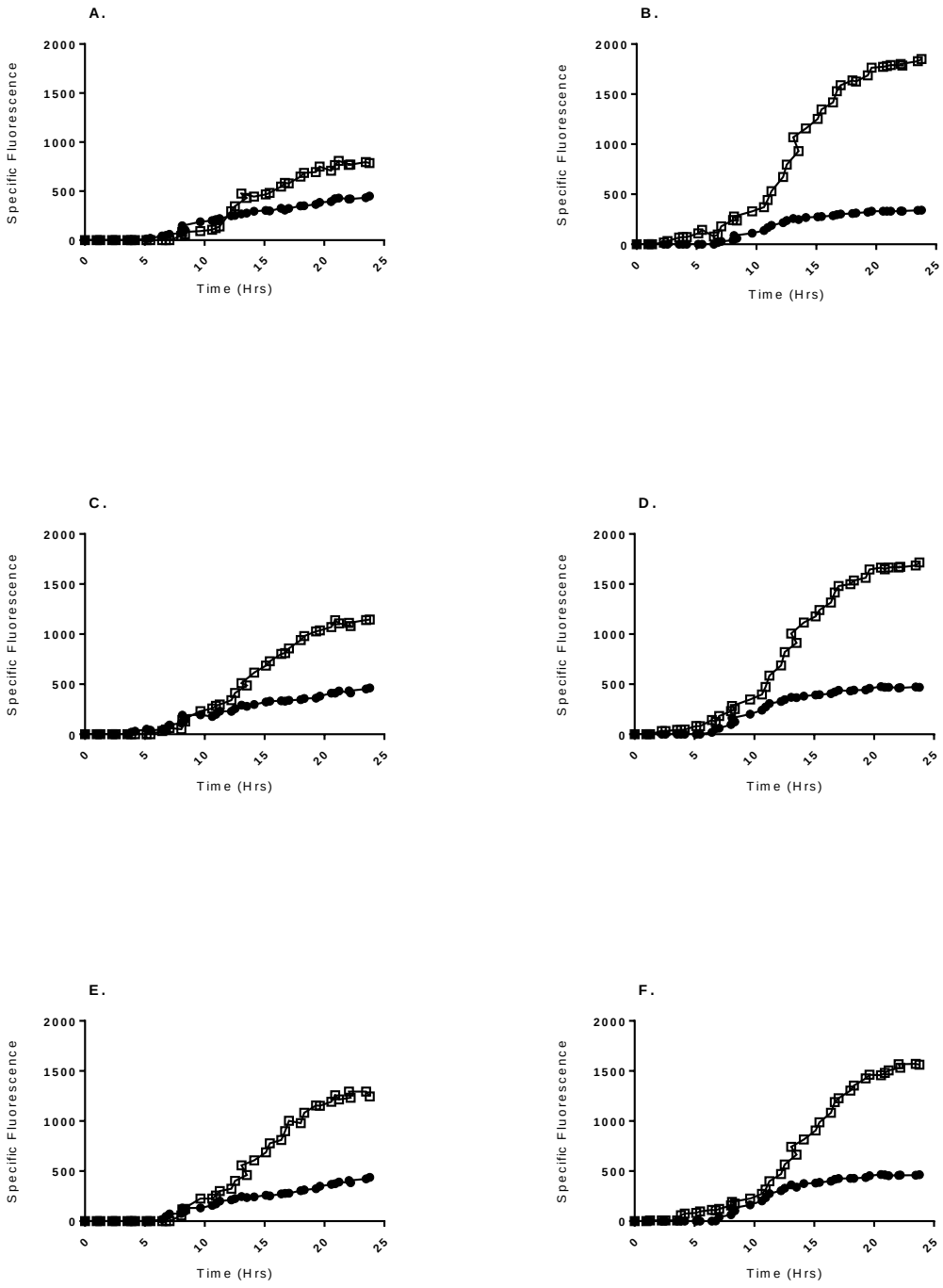


Figure 23: ***X. nematophila*, NMI1, and NMI2 specific fluorescence expression over 24 hours.** *X. nematophila* (A,B), NMI1 (C,D), and NMI2 (E,F) specific fluorescence expression over 24 hours of *pixA* (●) and *pixB* (□) in LB (A,C,E) and GM (B,D,F). Expression of *pixA* and *pixB* is lower in LB (A) than GM (B). Expression of NMI2 *pixB* is greater than wt in LB.

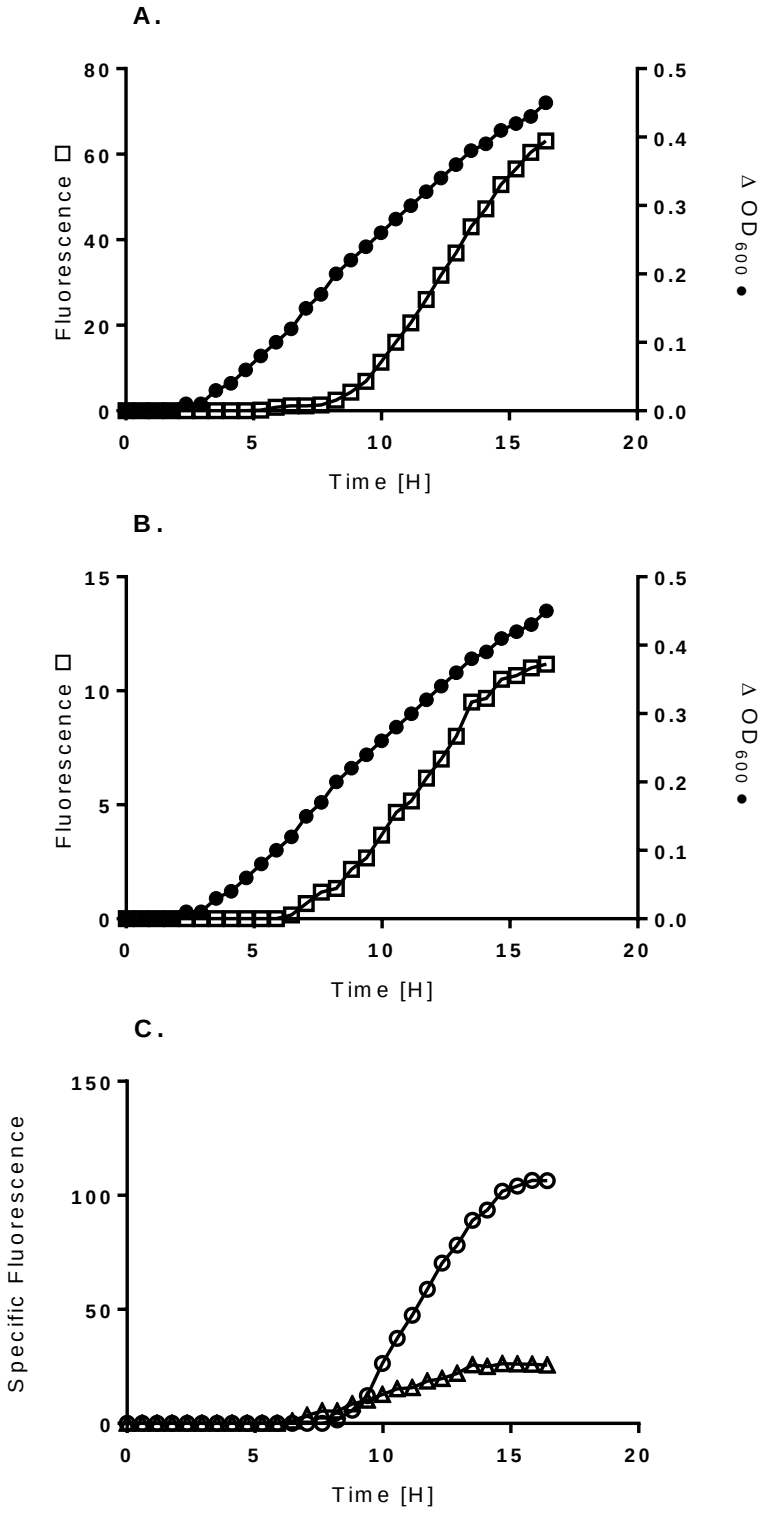
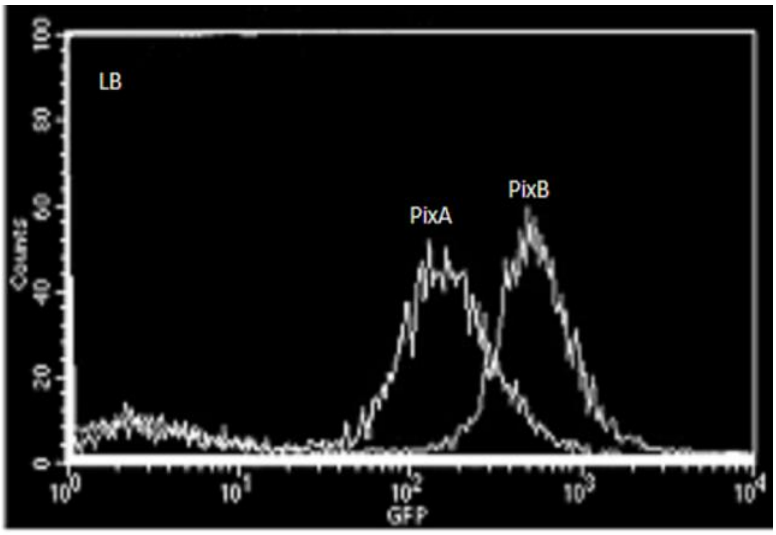


Figure 24: *X. nematophila pix* gene expression in hemolymph. *pixB* (A) and *pixA* (B) are activated during mid log growth. *pixB* (o) is expressed at higher levels than *pixA* (Δ) (C).

percentage of cells expressing either protein is comparable. The relative amount of expression of *pixB* per cell is greater than that of *pixA* in both LB and GM cultures. Expression of both *pixA* and *pixB* is greater in GM than LB (Fig. 25). Significant differences occur in mean fluorescent intensity of both *pixB* and *pixA* expression between LB and GM cultures. *pixB* expression is 3.2X that of *pixA* in LB and 6.1X that of *pixA* in GM (Fig. 26). Relative expression profiles of *pixA* and *pixB* were ascertained from 24 and 48-hour hemolymph samples of 4th instar *M. sexta* insects infected with *X. nematophila* harboring the GFP reporters or the promoterless pProbeNT vector. Gated cells were assessed for relative percent expression of either *pixA* or *pixB*. At 24 hours post injection hemolymph samples revealed only 1% of cells expressing *pixA* and 10% of cells expressing *pixB*, however by 48 hours the percent of cells expressing *pixA* increased to 77% while those expressing *pixB* increased to 23%. Relative expression levels of *pixB* is greater than *pixA* at both 24 hour and 48 hours post inoculation (Fig. 27). Mean fluorescent intensity of *pixB* expression is 2X that of *pixA* at 24 hours and 3.5X that of *pixA* at 48 hours in insect hemolymph (Fig. 28).

A.



B.

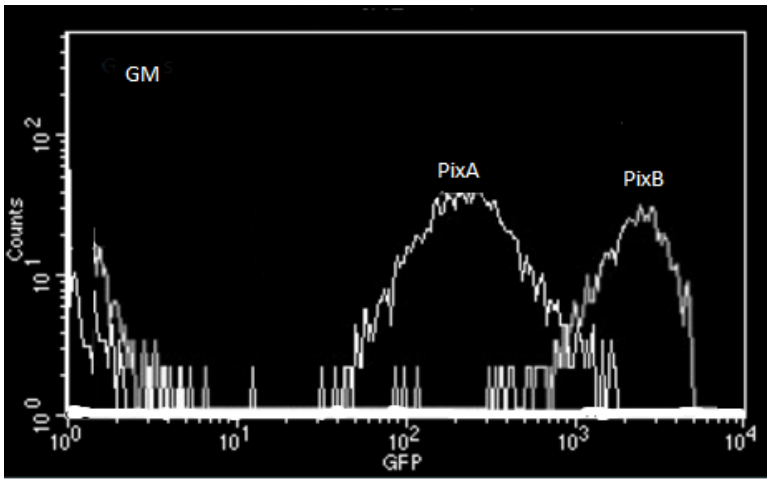


Figure 25: The relative amount of expression of *pixA* and *pixB* in LB and GM. Relative expression of *pixB* is greater than that of *pixA* in both LB (A) and GM cultures (B). Expression of *pixA* and *pixB* is greater in GM than LB.

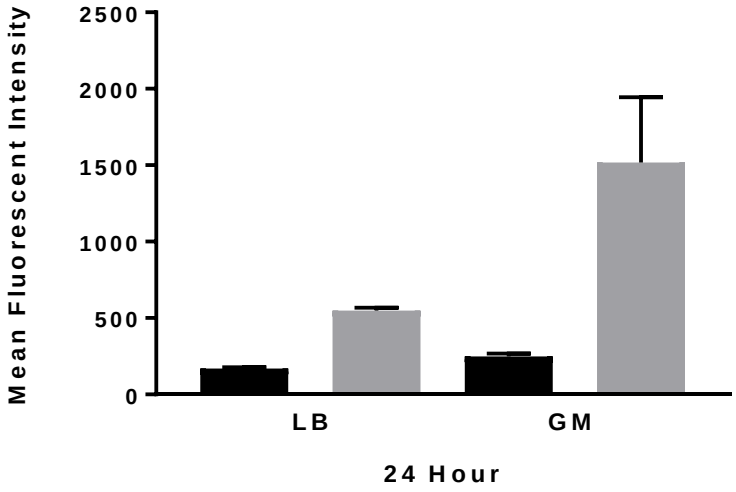


Figure 26: *X. nematophila* Mean Fluorescent Intensity of 24 hour cultures in either LB or GM. Significant differences occur in mean fluorescent intensity of *pixA* expression (dark columns) in LB and GM cultures (p=0.0019 by t test). Significant differences occur in mean fluorescent intensity of *pixB* expression (light columns) occurs between LB and GM cultures (p=0.0122 by t test). *pixB* expression is 3.2X that of *pixA* in LB and 6.1X that of *pixA* in GM.

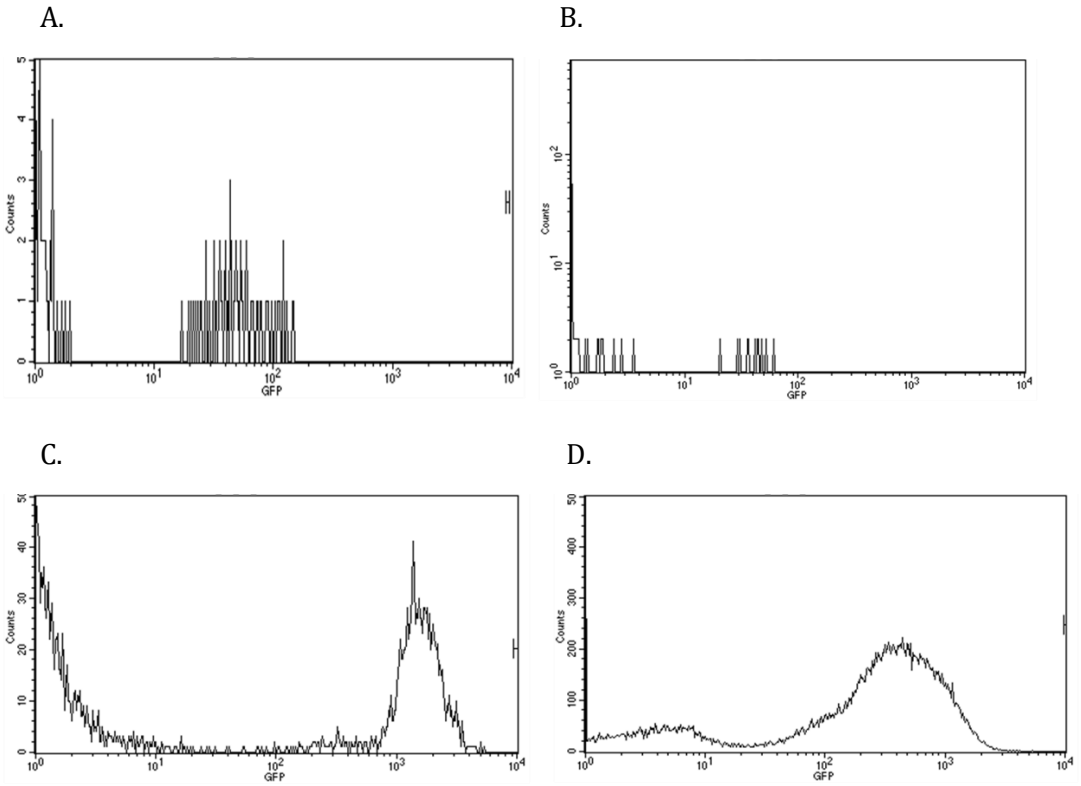


Figure 27: The relative amount of expression of *pixA* and *pixB* in hemolymph. Relative expression of *pixB* (A,C) is greater than that of *pixA* (B,D) at both 24 hour (A,B) and 48 hours post inoculation (C,D).

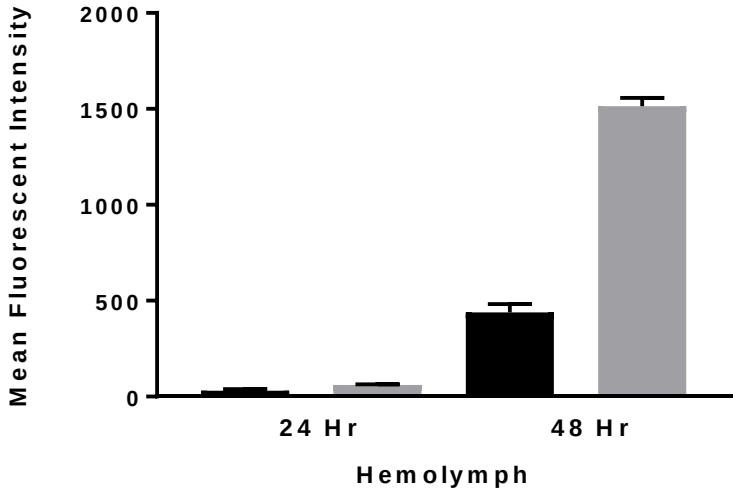


Figure 28: *X. nematophila* Mean Fluorescent Intensity of 24 and 48 hour cells grown in insect hemolymph. *pixB* expression (light columns) is 2X that of *pixA* (dark columns) at 24 hours and 3.5X that of *pixA* at 48 hours.

Chapter 5

Discussion

Neither *pixA* nor *pixB* is present in every species of *Xenorhabdus* which indicate the proteins are not essential to the survival of the bacterium. Those *Xenorhabdus* species which contain a *pix* gene are not isolated to a particular clade nor is the gene restricted to a specific genomic location which suggests the possibility of incorporation via horizontal gene transfer (HGT). Evidence for HGT is further supported based on genomic analysis and phylogenetic results illustrating branches of PixB homologs from multiple classes. Phylogenetic results also indicate that the *Xenorhabdus* PixB species is the most recent evolutionary ancestor of PixA. Those organisms in which a PixB homolog has been identified are most representative of soil dwelling and clinical microbes. HGT between clinical pathogens and soil dwelling microbes has been described previously (Forsberg et al., 2012).

PixA and PixB of *X. nematophila* share a 30% identity, however, the PixA found in some species contains a higher identity to the PixB sequence of other species than to other PixA homologs. Moreover, secondary structure alignments predict the proteins share similar folding. The widespread occurrence of PixB across classes coupled with the identification of now two homologous variants suggests *pixB* represents a family of proteins of which *pixA* and *pixC* appear to be subfamily paralogs. The LD of the PixB motif in PixB homologs was found to be rich in serine and threonine while in PixA richness favors methionine and in PixC glycine is favored.

The significance of the LD and richness of (S/T) residues in the PixB motif is unclear. However, natural linker domains act as spacers between two functionally distinct domains

and are typically enriched with serine, glycine, threonine, and alanine (Reddy Chichili et al., 2013). Serine rich protein domains in proteobacteria have been described and found to be involved in carbohydrate binding and degradation. Characteristic of these proteins is a polyserine linker domain that separates a catalytic and binding domain. The secondary structure of this region is predicted to form an extended flexible loop, the PixA and PixC LD secondary structure predictions also show this region to form a loop. Like in PixA and PixC, the length of this region is variable, depending on the number of enriched residues present, creating a "flexible spacer" which has been postulated to enhance the targeting of substrates when associated to the outer membrane (Howard et al., 2004).

The LD of PixA variants is enriched with methionine and is relatively more variable in length depending on the degree of enrichment. Coupled with the degradation of CD1/CD2 consensus, it is unclear whether PixA is the result of functional divergence or functional adaptation favorable for maintenance within *Xenorhabdus*. PixA may play a role as a nutritional storage protein with residue variability tuned to the nutritional requirements of the bacterium's nematode host. *Photorhabdus* is a close relative to *Xenorhabdus* that also produces a methionine rich inclusion protein CipA. Previous nematode feeding experiments with CipA have demonstrated that the presence or absence of the protein influences the development of the nematode (You et al., 2006). SDS-PAGE results seem to show a subjective increase in one Pix protein when the other ceases to be expressed. If the Pix proteins are functioning as a nutritional storage protein the loss of one protein may result in increased availability of nutrients to be sequestered by the remaining protein resulting in increased accumulation.

PixC has a LD with variable polyglycine enhancement with good conservation about CD1 and CD2. Like PixA the LD in PixC is expanded to varying degrees depending on the amount of glycine incorporation. Glycine rich repeat domains have been described in various secreted virulence factors including the phage GP38 adhesins (Trojet et al., 2011). Present in *Mycobacterium tuberculosis* are PE_PGRS genes encoding RTX toxins with C-terminal glycine rich repeat sequences (GGxGx(D/N)xUx). This repeat sequence closely resembles the glycine rich LD sequences found in PixC. These proteins are described as calcium binding beta-roll or beta-helix structures that are thought to function in evasion of host immune responses (Bachhawat and Singh, 2007).

Nematode killing experiments performed with the PixB homolog AidA demonstrate that the protein plays a role in bacterial accumulation within the nematode gut resulting in nematicidal activity on NGM media (Huber et al., 2004), however the exact mechanism of accumulation is unclear. The PixC variants identified were sequenced from organisms that have also been found associated with nematodes. However, it is unclear if PixC plays a role in this association. When the AidA nematode experiment was performed with *E. coli* containing the XNC1 PixB expression vector, no significant differences were noted on NGM, indeed in all circumstances the nematodes could grow and reproduce normally. It may be that PixB needs to be exported to produce this nematicidal effect and *E. coli* simply does not naturally possess the proper mechanisms to accomplish this task.

AidA, mannose-fucose binding lectin LecM, and the quorum quenching AidC protein are co-expressed in *Ralstonia.* These factors contribute to improved virulence while the *lecM* mutant results in deficient biofilm formation (Meng et al., 2015). The identification of a possible eukaryotic PixB homolog was striking both in the type of organism, one which is

pathogenic and associates with hemolymph, and in its identification as a mannose transferase protein. It is also interesting that the predicted omega cleavage site is immediately adjacent to CD2. It may be that PixB is anchored to the membrane with an analogous prokaryotic mechanism. Several PixB homologs annotated as "DNA dependent RNA polymerases" were also identified. RNAP have also been found to associate with cell membranes (Ding and Winkler, 1990).

Ultrathin sections of *Xenorhabdus* cultured in different types of media reveal curious changes in membrane structure along with a subjective decrease in organized inclusions and possibly the complete loss of inclusions once *in vivo* conditions are met. These results suggest that *in vitro* expression and accumulation may be the result of '*confused regulation*' in the presence of non-endogenous growth conditions. The majority of cells cultured in LB appeared to contain well defined inclusion bodies while those cells cultured in GM seemed to contain very few diffuse inclusions. LB is a complex medium while GM is a defined medium containing various salts, amino acids, sugars, vitamins and organic acids designed to mimic insect hemolymph. In hemolymph the protein may be effectively exported from within the cell thereby resulting in the loss of inclusion bodies.

Viability analysis indicates increased cellular viability in the absence of *pixB* expression when cultured in LB but not when cultured in GM. Moreover, expression data indicates higher levels of expression in GM than LB, which is consistent with FACS data indicating higher levels of cellular *pixB* expression. If PixB is detrimental to cell survival, the greater levels of expression occurring in GM might lead one to expect at least comparable viability results. The decrease in viability during LB culture may be due to a toxic intercellular build up of protein that might otherwise be secreted in GM culture conditions.

The higher levels of *pixB* expression are most probably a result of a strong sigma 70 promoter and lower percentage of G/C content (13.3%) in the +8 to -82 region compared to *pixA* (26.7%).

The current function of PixB is unclear but it may be involved in adhesion and/or biofilm formation, possibly via a sugar binding property that may contribute to an organism's virulence *in vivo* and/or ability to associate with a nematode host that could transport the organism to sources of fresh nutrients. Future localization studies of *in vivo* cultures will need to be conducted to ascertain whether PixB is naturally exported, associating with hemolymph cells and/or involved in biofilm formation.

References

1. **Aasland R**, **Abrams C**, **Ampe C**, **Ball LJ**, **Bedford MT**, **Cesareni G**, **Gimona M**, **Hurley JH**, **Jarchau T**, **Lehto V-P**, **Lemmon MA**, **Linding R**, **Mayer BJ**, **Nagai M**, **Sudol M**, **Walter U**, **Winder SJ**. 2002. Normalization of nomenclature for peptide motifs as ligands of modular protein domains. FEBS Lett **513**:141–144.

2. **Acton QA**. 2012. Advances in Salmonella enterica Research and Application: 2012 Edition. ScholarlyEditions.

3. **Alexeyev MF**. 1999. The pKNOCK series of broad-host-range mobilizable suicide vectors for gene knockout and targeted DNA insertion into the chromosome of gram-negative bacteria. Biotechniques **26**:824–6, 828.

4. **Altschul SF**, **Madden TL**, **Schäffer 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.

5. **Bachhawat N**, **Singh B**. 2007. Mycobacterial PE_PGRS proteins contain calcium-binding motifs with parallel beta-roll folds. Genomics Proteomics Bioinformatics **5**:236–241.

6. **Couche GA**, **Gregson RP**. 1987. Protein inclusions produced by the entomopathogenic bacterium Xenorhabdus nematophilus subsp. nematophilus. J Bacteriol **169**:5279–5288.

7. **Ding HF**, **Winkler HH**. 1990. Purification and partial characterization of the DNA-dependent RNA polymerase from Rickettsia prowazekii. J Bacteriol **172**:5624–5630.

8. **Ebert D**, **Lipsitch M**, **Mangin KL**. 2000. The Effect of Parasites on Host Population Density and Extinction: Experimental Epidemiology with Daphnia and Six Microparasites. Am Nat **156**:459–477.

9. **Espitia C**, **Laclette JP**, **Mondragón-Palomino M**, **Amador A**, **Campuzano J**, **Martens A**, **Singh M**, **Cicero R**, **Zhang Y**, **Moreno C**. 1999. The PE-PGRS glycine-rich proteins of Mycobacterium tuberculosis: a new family of fibronectin-binding proteins? Microbiology **145 (Pt 12)**:3487–3495.

10. **Felsenstein J**. 1985. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. Evolution **39**:783.

11. **Forsberg KJ**, **Reyes A**, **Wang B**, **Selleck EM**, **Sommer MOA**, **Dantas G**. 2012. The shared antibiotic resistome of soil bacteria and human pathogens. Science **337**:1107–1111.

12. **Forst S**, **Dowds B**, **Boemare N**, **Stackebrandt E**. 1997. Xenorhabdus and Photorhabdus spp.: bugs that kill bugs. Annu Rev Microbiol **51**:47–72.

13. **Gengler S**, **Laudisoit A**, **Batoko H**, **Wattiau P**. 2015. Long-term persistence of Yersinia pseudotuberculosis in entomopathogenic nematodes. PLoS One **10**:e0116818.

14. **Goetsch M**, **Owen H**, **Goldman B**, **Forst S**. 2006. Analysis of the PixA inclusion body protein of Xenorhabdus nematophila. J Bacteriol **188**:2706–2710.

15. **Gotz P**, **Boman A**, **Boman HG**. 1981. Interactions between Insect Immunity and an Insect-Pathogenic Nematode with Symbiotic Bacteria. Proceedings of the Royal Society B: Biological Sciences **212**:333–350.

16. **Herbert EE**, **Goodrich-Blair H**. 2007. Friend and foe: the two faces of Xenorhabdus nematophila. Nat Rev Microbiol **5**:634–646.

17. **Howard MB**, **Ekborg NA**, **Taylor LE**, **Hutcheson SW**, **Weiner RM**. 2004. Identification and analysis of polyserine linker domains in prokaryotic proteins with emphasis on the marine bacterium Microbulbifer degradans. Protein Sci **13**:1422–1425.

18. **Huber B**, **Feldmann F**, **Kothe M**, **Vandamme P**, **Wopperer J**, **Riedel K**, **Eberl L**. 2004. Identification of a Novel Virulence Factor in Burkholderia cenocepacia H111 Required for Efficient Slow Killing of Caenorhabditis elegans. Infect Immun **72**:7220–7230.

19. **Jubelin G**, **Pagès S**, **Lanois A**, **Boyer M-H**, **Gaudriault S**, **Ferdy J-B**, **Givaudan A**. 2011. Studies of the dynamic expression of the Xenorhabdus FliAZ regulon reveal atypical iron-dependent regulation of the flagellin and haemolysin genes during insect infection. Environ Microbiol **13**:1271–1284.

20. **Källberg M**, **Wang H**, **Wang S**, **Peng J**, **Wang Z**, **Lu H**, **Xu J**. 2012. Template-based protein structure modeling using the RaptorX web server. Nat Protoc **7**:1511–1522.

21. **Kumar S**, **Stecher G**, **Tamura K**. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol **33**:1870–1874.

22. **Larkin MA**, **Blackshields G**, **Brown NP**, **Chenna R**, **McGettigan PA**, **McWilliam H**, **Valentin F**, **Wallace IM**, **Wilm A**, **Lopez R**, **Thompson JD**, **Gibson TJ**, **Higgins DG**. 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23**:2947–2948.

23. **Lee M-M**, **Stock SP**. 2010. A multilocus approach to assessing co-evolutionary relationships between Steinernema spp. (Nematoda: Steinernematidae) and their bacterial symbionts Xenorhabdus spp. (gamma-Proteobacteria: Enterobacteriaceae). Syst Parasitol **77**:1–12.

24. **Linhartová I**, **Bumba L**, **Mašín J**, **Basler M**, **Osička R**, **Kamanová J**, **Procházková K**, **Adkins I**, **Hejnová-Holubová J**, **Sadílková L**, **Morová J**, **Sebo P**. 2010. RTX proteins: a highly diverse family secreted by a common mechanism. FEMS Microbiol Rev **34**:1076–1112.

25. **Ma J**, **Wang S**, **Zhao F**, **Xu J**. 2013. Protein threading using context-specific alignment potential. Bioinformatics **29**:i257–i265.

26. **Manning BD**, **Cantley LC**. 2007. AKT/PKB signaling: navigating downstream. Cell **129**:1261–1274.

27. **Meng F**, **Babujee L**, **Jacobs JM**, **Allen C**. 2015. Comparative Transcriptome Analysis Reveals Cool Virulence Factors of Ralstonia solanacearum Race 3 Biovar 2. PLoS One **10**:e0139090.

28. **Morales-Soto N**, **Snyder H**, **Forst S**. 2009. Interspecies Competition in a Bacteria-Nematode Mutualism. Defensive Mutualism in Microbial Symbiosis, Vol. 27 (White JF & Torres MS, eds), pp. 117–127. Taylor and Francis Group, LLC, Boca Raton, FL.

29. **Nykyri J**, **Fang X**, **Dorati F**, **Bakr R**, **Pasanen M**, **Niemi O**, **Palva ET**, **Jackson RW**, **Pirhonen M**. 2013. Evidence that nematodes may vector the soft rot-causing enterobacterial phytopathogens. Plant Pathol **63**:747–757.

30. **Peng J**, **Xu J**. 2011. Raptorx: Exploiting structure information for protein alignment by statistical inference. Proteins: Struct Funct Bioinf **79**:161–171.

31. **Pierleoni A**, **Martelli P**, **Casadio R**. 2008. PredGPI: a GPI-anchor predictor. BMC Bioinformatics **9**:392.

32. **Reddy Chichili VP**, **Chichili VPR**, **Kumar V**, **Sivaraman J**. 2013. Linkers in the structural biology of protein-protein interactions. Protein Sci **22**:153–167.

33. **Shrestha S**, **Kim Y**. 2007. An entomopathogenic bacterium, Xenorhabdus nematophila, inhibits hemocyte phagocytosis of Spodoptera exigua by inhibiting phospholipase A2. J Invertebr Pathol **96**:64–70.

34. **Singh S**, **Orr D**, **Divinagracia E**, **McGraw J**, **Dorff K**, **Forst S**. 2015. Role of secondary metabolites in establishment of the mutualistic partnership between Xenorhabdus nematophila and the entomopathogenic nematode Steinernema carpocapsae. Appl Environ Microbiol **81**:754–764.

35. **Snyder H**, **Stock SP**, **Kim S-K**, **Flores-Lara Y**, **Forst S**. 2007. New Insights into the Colonization and Release Processes of Xenorhabdus nematophila and the Morphology and Ultrastructure of the Bacterial Receptacle of Its Nematode Host, Steinernema carpocapsae. Appl Environ Microbiol **73**:5338–5346.

36. **Timme RE**, **Pettengill JB**, **Allard MW**, **Strain E**, **Barrangou R**, **Wehnes C**, **Van Kessel JS**, **Karns JS**, **Musser SM**, **Brown EW**. 2013. Phylogenetic diversity of the enteric pathogen Salmonella enterica subsp. enterica inferred from genome-wide reference-free SNP characters. Genome Biol Evol **5**:2109–2123.

37. **Trojet SN**, **Caumont-Sarcos A**, **Perrody E**, **Comeau AM**, **Krisch HM**. 2011. The gp38 adhesins of the T4 superfamily: a complex modular determinant of the phage's host specificity. Genome Biol Evol **3**:674–686.

38. **Vicente CSL**, **Nascimento FX**, **Espada M**, **Barbosa P**, **Hasegawa K**, **Mota M**, **Oliveira S**. 2013. Characterization of bacterial communities associated with the pine sawyer beetle Monochamus galloprovincialis, the insect vector of the pinewood nematode Bursaphelenchus xylophilus. FEMS Microbiol Lett **347**:130–139.

39. **Vigneux F**, **Zumbihl R**, **Jubelin G**, **Ribeiro C**, **Poncet J**, **Baghdiguian S**, **Givaudan A**, **Brehélin M**. 2007. The xaxAB genes encoding a new apoptotic toxin from the insect pathogen Xenorhabdus nematophila are present in plant and human pathogens. J Biol Chem **282**:9571–9580.

40. **Volgyi, A., Fodor, A., Szentirmai, A., & Forst, S.** 1998. Phase Variation in Xenorhabdus nematophilus. Applied and Environmental Microbiology, **64**:1188–1193.

41. **Wang S**, **Peng J**, **Xu J**. 2011. Alignment of distantly-related protein structures: algorithm, bound and implications to homology modeling. Bioinformatics.

42. **Wang S**, **Ma J**, **Peng J**, **Xu J**. 2013. Protein structure alignment beyond spatial proximity. Sci Rep **3**:1448.

43. **Xu, J., & Hurlbert, R. E.** 1990. Toxicity of Irradiated Media for Xenorhabdus spp. Applied and Environmental Microbiology, **56**:815–818.

44. **Xu J**, **Lohrke S**, **Hurlbert IM**, **Hurlbert RE**. 1989. Transformation of Xenorhabdus nematophilus. Appl Environ Microbiol **55**:806–812.

45. **You J**, **Liang S**, **Cao L**, **Liu X**, **Han R**. 2006. Nutritive significance of crystalline inclusion proteins of Photorhabdus luminescens in Steinernema nematodes. FEMS Microbiol Ecol **55**:178–185.

46. **Zuckerkandl E**, **Pauling L**. 1965. Evolutionary Divergence and Convergence in Proteins, p. 97–166. *In* Evolving Genes and Proteins.

Appendices

Appendix A:
Table S1: Bacterial strains and plasmids used in this study

BACTERIAL STRAIN OR PLASMID	RELEVANT GENOTYPE	SOURCE
X. nematophila 19061	Wild type, Amp^r	ATCC
NMI1	19061 *pixA* mutant	Lab stock
NMI2	19061 *pixB* mutant	Lab stock
E. coli BL21DE3	*fhuA2* [lon] *ompT* (λ DE3)	Lab stock
E. coli DH5a	*recA1 endA1 hsdR17 relA1*	Lab stock
E. coli S17 λpir	*recA, thi, pro, hsdR*-M+. RP4-2Tc::Mu Km::Tn7	Lab stock
pACYCDuet-1	Cm^r	Lab stock
pPROBE-NT	GFP, Kan^r	Addgene
pLUCY	*pixA_UTR*:GFP, Kan^r	This study
pPB400	*pixB_UTR*:GFP, Kan^r	This study
pMG006	pACYCDuet-1 IPTG inducible PixB expression, Cm^r	Lab stock

Appendix B:
Table S2: Primers used in this study

Primer	Description	Sequence (Restriction Site underlined)	Direction
FA5	*pixA* forward primer incorporating Nco1 Restriction site	5'-AGT AAC ATG CCA TGG GGA TGA GAA ATA TCG ATA TTA TGT TAG CA-3'	Forward
RA5	*pixA* Reverse primer incorporating HindIII Restriction site	5'-AGT AAC CCA AGC TTT TAC ATC ATA ACT GTG TGA TCC AGA ATA AG-3'	Reverse
FGFP3A	*pixA* 5'UTR forward primer for amplification from gDNA	5'-CCC TAA GTA ATT GAT CAG GTA A-3'	Forward
RGFP1A	*pixA* 5'UTR reverse primer for amplification from gDNA	5'-TGT AGT ACC CTT AAT TAG TTA ACG-3'	Reverse
FPA5UH3	*pixA* 5'UTR forward primer incorporating HindIII Restriction site	5'-GAT CCC AAG CTT CAT TAA TCA GAA TTA GAA AAA AAA TAG AAA CTG-3'	Forward
RPA5UE1	*pixA* 5'UTR Reverse primer incorporating EcoR1 Restriction site	5'-GAT CCG GAA TTC TGT AGT ACC CTT AAT TAG TTA ACG TTA TGA-3'	Reverse
AAP	5' RACE Abridged Anchor Primer	5'-GGC CAC GCG TCG ACT AGT ACG GGI IGG GII GGG IIG-3'	Forward
AUAP	5' RACE Abridged universal amplification primer	5'-GGC CAC GCG TCG ACT AGT AC-3'	Forward
T7P	T7 Promoter	5'-CTA ATA CGA CTC ACT ATA GGG-3'	Forward
SP6	SP6 reverse	5'-ATT TAG GTG ACA CTA TAG-3'	Reverse
RGSP3A	5' RACE *pixA* nested primer	5'-GAT AAT CTT TCA TTG TAT TGA TAT CTA-3'	Reverse
RGSP4A	5' RACE *pixA* nested primer	5'-CAA CAC CAT CTC TCA TTA CAT TCA TCG GTT CAA TTT TGA CCA AAG CAG CAG AG-3'	Reverse
pPROBE-F	pProbeNT forward primer	5'-CCA GGA ATT GGG GAT CGG AAG-3'	Forward
pPROBE-R	pProbeNT reverse primer	5'-AAT TGT GCC CAT TAA CAT CAC CAT C-3'	Reverse
FPB5UH3	*pixB* 5'UTR forward primer incorporating HindIII Restriction site	5'-GAT CCC AAG CTT AAT AAC AGC ACT GGA GTA ATT TTT G-3'	Forward
RPB5UE1	*pixB* 5'UTR Reverse primer incorporating EcoR1 Restriction site	5'-GAT CCT GAA TTC TAT AGT ACC CTC ACT AGT TAA TGG TA-3'	Reverse
RGSP1B	5' RACE *pixB* nested primer	5'-CAG ATT TCA ATA CGG GGA TA-3'	Reverse
RGSP2B	5' RACE *pixB* nested primer	5'-CTC GCA TCC CTC GAA ACT GCG CGT AGC-3'	Reverse

Appendix C: Table S3: Sequences used for phylogenetic analysis. %ID and E-value cutoff scores used were 25% and 8.00E-04 respectively.

Designation	Genus	Species	Strain	ACCESSION #	% ID	% Similarity	E-value
PixA	Xenorhabdus	bovienii	Intermedium	CDH34691	36	53	4.00E-28
PixA	Xenorhabdus	bovienii	kraussei Quebec	CDH21217	36	53	4.00E-28
PixA	Xenorhabdus	cabanillasii	JM26	CDL86366	30	47	3.00E-16
PixA	Xenorhabdus	szentirmaii	DSM 16338	CDL83749	30	51	1.00E-18
PixA	Xenorhabdus	nematophila	19061	WP_010845455.1	30	50	1.00E-24
PixB	Xenorhabdus	bovienii	Intermedium	CDH31982	32	55	8.00E-30
PixB	Xenorhabdus	cabanillasii	JM26	CDL86734	57	69	5.00E-59
PixB	Xenorhabdus	nematophila	19061	WP_013184504.1	100	100	3.00E-130
PixB	Xenorhabdus	nematophila	F1	CCW32634	89	92	1.00E-112
PixB	Xenorhabdus	bovienii	kraussei Quebec	CDH22132	40	55	1.00E-32
(a)	Pectobacterium	carotovorum	YC D46	KHS91609	33	44	2.00E-10
(b)	Pectobacterium	carotovorum	CFIA1033	KFF65531	25	41	8.00E-04
(c)	Shewanella	baltica	OS155	ABN60197	33	50	4.00E-15
(d)	Shewanella	baltica	BA175	AEG12786	33	50	6.00E-15
(e)	Ralstonia	solanacearum	AW1	AAC45946	35	46	1.00E-14
(f)	Ralstonia	solanacearum	CFBP2957	CBJ41502	34	45	2.00E-13
(g)	Hyalangium	minutum	DSM 14724	KFE61976	30	49	9.00E-13
(h)	Hyalangium	minutum	minutum	KFE61977	31	49	2.00E-16
(i)	Myxococcus	fulvus	124B02	AKF85240	30	53	2.00E-10
(j)	Myxococcus	fulvus	124B02	AKF85239	27	48	5.00E-08
(k)	Rheinheimera	KL1	KL1	KOO59823	28	45	3.00E-08
(l)	Rheinheimera	A13L	A13L	EGM79634	28	46	2.00E-08
	Photorhabdus	luminescens	luminescens	WP_040154142	27	44	3.00E-08
	Photorhabdus	heterorhabditis	heterorhabditis	WP_054479578	30	50	4.00E-12
	Burkholderia	cepacia	JBK9	ALX14620	35	50	1.00E-20
	Burkholderia	anthina	MSMB0849	WP_059581353.1	35	50	7.00E-20
	Pseudoalteromonas	P1-9	P1-9	KPV95834	28	44	8.00E-06
	Chromobacterium	haemolyticum	haemolyticum	WP_043641426	31	50	2.00E-13
	Chromobacterium	LK11	LK11	KMN77888	31	50	4.00E-14
	Rhizobium	Rhizobium	Leaf311	KQQ58627	32	47	6.00E-11
	Caulobacter	henricii	henricii	WP_035070873	31	48	4.00E-14
	Streptomyces	ipomoeae	91-03	EKX60943	30	51	6.00E-14
	Streptomyces	himastatinicus	ATCC 53653	EFL20450	35	51	5.00E-13
	Streptosporangium	roseum	DSM 43021	WP_012886853.1	34	53	7.00E-16
	Halomonas	TD01	TD01	EGP19391	31	53	2.00E-19
	Bradyrhizobium	yuanmingense	yuanmingense	WP_036032112	30	46	5.00E-16
	Bradyrhizobium	jicamae	PAC68	KRQ92913	31	46	1.00E-16
	Pseudomonas	fluorescens	FW300-N2E3	ALI02782	31	51	1.00E-19
	Pseudomonas	GR 6-02	GR 6-02	WP_019580873	31	50	2.00E-19
	Mumia	flava	MUSC 201	KHL13532	31	47	6.00E-18
	Burkholderia	cenocepacia	H111	CDN64866	31	47	1.00E-17

Appendix D:
Table S4: Identified homologous PixB proteins from NCBI BLAST results

	Genus	Species	Strain	ACCESSION #	% ID	% Similarity	E-value	Notes/NCBI Annotations "-" (CD1 = GDXXRWRXX(S/T) CD2 = G(Y/C)XXWDP)
1	Xenorhabdus	nematophila	Anatoliense	CEE93243	100	100	3.00E-130	PixB
2	Xenorhabdus	nematophila	AN6	CEK23445	100	100	3.00E-130	PixB
3	Xenorhabdus	nematophila	Websteri	CEF30909	99	100	2.00E-85	PixB
4	Xenorhabdus	nematophila	F1	CCW32634	89	92	1.00E-112	PixB
5	Xenorhabdus	cabanillasii	JM26	CDL86734	57	69	5.00E-59	PixB CD2 (G to A)
6	Xenorhabdus	bovienii	kraussei Quebec	CDH22132	40	55	1.00E-32	PixB CD1 (R to H)
7	Xenorhabdus	bovienii	Intermedium	CDH31982	32	55	8.00E-30	PixB CD2 (W to M)
8	Xenorhabdus	bovienii	kraussei Quebec	CDH21217	36	53	4.00E-28	True PixA, partial conservation of CD1 (W to F, R to N) and CD2 (W to I, D to E)
9	Xenorhabdus	bovienii	Intermedium	CDH34691	36	53	4.00E-28	True PixA, partial conservation of CD1 (W to F, R to N) and CD2 (W to I, D to E)
10	Xenorhabdus	bovienii	feltiae France	CDG87156	36	53	5.00E-28	True PixA, partial conservation of CD1 (W to F, R to N) and CD2 (W to I, D to E)
11	Xenorhabdus	bovienii	feltiae Florida	CDG94507	36	53	5.00E-28	True PixA, partial conservation of CD1 (W to F, R to N) and CD2 (W to I, D to E)
12	Xenorhabdus	bovienii	puntauvense	CDG95144	36	52	9.00E-28	True PixA, partial conservation of CD1 (W to F, R to N) and CD2 (W to I, D to E)

13	Burkholderia	cepacia	JBK9	ALX14620	35	50	1.00E-20	"PixA" CD1 (G to N), Gene duplication Chromosome 2
14	Burkholderia	anthina	MSMB0849	KVH07160	35	50	7.00E-20	"PixA" CD1 (G to N)
15	Burkholderia	anthina	MSMB0848	KVH08084	35	50	7.00E-20	"PixA" CD1 (G to N)
16	Burkholderia	anthina	MSMB1497	KVM90525	35	50	7.00E-20	"PixA" CD1 (G to N)
17	Burkholderia	anthina	MSMB1496	KVN51712	35	50	7.00E-20	"PixA" CD1 (G to N)
18	Burkholderia	anthina	MSMB1506WGS	KVX35308	35	50	7.00E-20	"PixA" CD1 (G to N)
19	Pseudomonas	fluorescens	FW300-N2E3	ALI02782	31	51	1.00E-19	"PixA" CD1 (G to L)
20	Halomonas	TD01	TD01	EGP19391	31	53	2.00E-19	"PixA"
21	Pseudomonas	MultiSpecies	NA	WP_019580873	31	50	2.00E-19	"PixA"
22	Pseudomonas	chlororaphis	piscium	KZO50600	31	50	6.00E-19	"PixA"
23	Pseudomonas	fluorescens	HK44	EXF91114	31	50	8.00E-19	"PixA"
24	Burkholderia	ubonensis	RF32-BP3	KUZ80147	34	48	1.00E-18	"PixA" CD1 (G to N)
25	Xenorhabdus	szentirmaii	DSM 16338	CDL83749	30	51	1.00E-18	True PixA, partial conservation of CD1 (W to L, R to N); Loss of CD2 (W to V, D to E, P to T)
26	Pseudomonas	syringae	Riq4	KNH27756	31	51	1.00E-18	"PixA"
27	Pseudomonas	GM41(2012)	GM41(2012)	EUB72119	31	51	2.00E-18	AidA
28	Pseudomonas	MRSN12121	MRSN12121	AJO77056	31	51	3.00E-18	"PixA" CD2 (G to A)
29	Pseudomonas	chlororaphis	PCL1606	AKA26152	31	51	3.00E-18	"PixA" CD2 (G to A)
30	Pseudomonas	fluorescens	C1	KJZ38182	30	51	3.00E-18	"PixA"
31	Burkholderia	cepacia	JBK9	ALX16764	31	48	3.00E-18	"PixA" Gene duplication Chromosome 3
32	Pseudomonas	CFII68	CFII68	EPJ88298	30	51	4.00E-18	"PixA" CD2 (G to A)
33	Pseudomonas	GM79	GM79	EJN19656	30	51	4.00E-18	"PixA"

34	Pseudomonas	fluorescens	FW300-N1B4	KZN18539	30	51	4.00E-18	"PixA"
35	Pseudomonas	MultiSpecies	NA	WP_018926817	30	51	4.00E-18	"PixA"
36	Pseudomonas	GM102	GM102	EJL98347	32	48	5.00E-18	"PixA"
37	Pseudomonas	GM50	GM50	EJM59492	32	48	5.00E-18	"PixA"
38	Pseudomonas	lini	DSM 16768	KMM94062	32	48	5.00E-18	"PixA"
39	Pseudomonas	lini	ZBG1	KNH45576	32	48	5.00E-18	"PixA"
40	Pseudomonas	Root329	Root329	KQV16104	30	51	6.00E-18	"PixA"
41	Burkholderia	lata	383	ABB06256	31	47	6.00E-18	"PixA"
42	Burkholderia	contaminans	FFH2055	KKL35787	31	47	6.00E-18	"PixA"
43	Burkholderia	contaminans	LMG 23361	KKL42170	31	47	6.00E-18	"PixA"
44	Burkholderia	contaminans	MS14	AKM45145	31	47	6.00E-18	"PixA"
45	Mumia	flava	MUSC 201	KHL13532	31	47	6.00E-18	"PixA", High G+C content, Gram +
46	Pseudomonas	655	655	KOY02858	30	51	7.00E-18	"PixA"
47	Pseudomonas	In5	In5	KPN92316	30	51	7.00E-18	"PixA"
48	Pseudomonas	mandelii	PD30	KDD65828	30	51	7.00E-18	"PixA"
49	Burkholderia	cepacia	LK29	KML43855	31	47	1.00E-17	"PixA"
50	Burkholderia	cenocepacia	CEIB S5-2	KWU27628	31	47	1.00E-17	"PixA"
51	Pseudomonas	fluorescens	C3	KJZ45612	30	51	1.00E-17	"PixA"
52	Burkholderia	cepacia	MSMB0859	KVH35186	31	47	1.00E-17	AidA
53	Burkholderia	cepacia	MSMB1302	KVK71923	31	47	1.00E-17	AidA
54	Burkholderia	cepacia	MSMB1820	KVK97762	31	47	1.00E-17	AidA
55	Burkholderia	cepacia	MSMB1906	KVL56302	31	47	1.00E-17	AidA
56	Burkholderia	cepacia	AU 1054	ABF80382	31	47	1.00E-17	AidA
57	Burkholderia	cepacia	HI2424	ABK12607	31	47	1.00E-17	AidA

58	Burkholderia	cepacia	MC0-3	ACA95498	31	47	1.00E-17	AidA
59	Burkholderia	cepacia	J2315	CAR57225	31	47	1.00E-17	AidA
60	Burkholderia	cepacia	K56-2Valvano	EPZ85253	31	47	1.00E-17	AidA
61	Burkholderia	cepacia	BC7	ERI26507	31	47	1.00E-17	AidA
62	Burkholderia	cepacia	H111	CDN64866	31	47	1.00E-17	AidA
63	Burkholderia	cepacia	869T2	KEA54850	31	47	1.00E-17	AidA
64	Burkholderia	cepacia	DDS 7H-2	AIO43676	31	47	1.00E-17	AidA
65	Burkholderia	cepacia	DWS 16B-4	KGC05564	31	47	1.00E-17	AidA
66	Burkholderia	cepacia	FDAARGOS_82	KGY43805	31	47	1.00E-17	AidA
67	Burkholderia	cepacia	LMG 16656	KIS50362	31	47	1.00E-17	AidA
68	Burkholderia	cepacia	K56-2	KKI82503	31	47	1.00E-17	AidA
69	Burkholderia	cepacia	ST32	ALV61348	31	47	1.00E-17	AidA
70	Burkholderia	seminalis	FL-5-5-10-S1-D0	KVF43682	31	47	1.00E-17	AidA
71	Burkholderia	cenocepacia	FL-6-2-30-S1-D2	KVF49181	31	47	1.00E-17	AidA
72	Burkholderia	cenocepacia	MSMB364WGS	KWF16077	31	47	1.00E-17	AidA
73	Burkholderia	cenocepacia	MSMB383WGS	KWF54437	31	47	1.00E-17	AidA
74	Burkholderia	cenocepacia	842	AMU11294	31	47	1.00E-17	AidA
75	Burkholderia	cenocepacia	895	AMU18992	31	47	1.00E-17	AidA
76	Pseudomonas	fluorescens	NCIMB 11764	AKV07160	30	51	2.00E-17	"PixA"
77	Burkholderia	cepacia	ATCC 25416	AIO26808	30	48	2.00E-17	"PixA"
78	Burkholderia	lata	LK13	KML22166	30	48	2.00E-17	"PixA"
79	Burkholderia	LK4	LK4	KMN56259	30	48	2.00E-17	"PixA"
80	Burkholderia	cepacia	UCB 717	ALK23144	30	48	2.00E-17	"PixA"
81	Burkholderia	cepacia	INT1-BP233	KVA30721	30	48	2.00E-17	"PixA"

82	Burkholderia	cepacia	INT1-BP232	KVA36964	30	48	2.00E-17	"PixA"
83	Burkholderia	cepacia	FL-4-2-10-S1-D0	KVF12427	30	48	2.00E-17	"PixA"
84	Burkholderia	cepacia	MSMB1129	KVH66724	30	48	2.00E-17	"PixA"
85	Burkholderia	cepacia	MSMB1224	KVH69728	30	48	2.00E-17	"PixA"
86	Burkholderia	cepacia	MSMB1824WGS	KVQ27173	30	48	2.00E-17	"PixA"
87	Burkholderia	cepacia	MSMB1831WGS	KVQ46004	30	48	2.00E-17	"PixA"
88	Burkholderia	cepacia	MSMB1055WGS	KVS30374	30	48	2.00E-17	"PixA"
89	Burkholderia	cepacia	MSMB1062WGS	KVS56827	30	48	2.00E-17	"PixA"
90	Burkholderia	cepacia	MSMB1062WGS	KVS56827	30	48	2.00E-17	"PixA"
91	Burkholderia	cepacia	MSMB1213WGS	KVU54366	30	48	2.00E-17	"PixA"
92	Burkholderia	cepacia	MSMB1486WGS	KVW85737	30	48	2.00E-17	"PixA"
93	Burkholderia	cepacia	MSMB1510WGS	KVX50114	30	48	2.00E-17	"PixA"
94	Burkholderia	cepacia	MSMB1513WGS	KVX64366	30	48	2.00E-17	"PixA"
95	Burkholderia	cepacia	MSMB1819WGS	KVZ98718	30	48	2.00E-17	"PixA"
96	Burkholderia	cepacia	MSMB1224WGS	KWC57625	30	48	2.00E-17	"PixA"
97	Burkholderia	cepacia	MSMB1228WGS	KWC76292	30	48	2.00E-17	"PixA"
98	Burkholderia	cepacia	MSMB2164WGS	KWD61810	30	48	2.00E-17	"PixA"
99	Burkholderia	cepacia	MSMB2165WGS	KWD85864	30	48	2.00E-17	"PixA"
100	Burkholderia	cepacia	MSMB2177WGS	KWE24593	30	48	2.00E-17	"PixA"
101	Burkholderia	cepacia	MSMB616WGS	KWH28447	30	48	2.00E-17	"PixA"
102	Burkholderia	cepacia	MSMB648WGS	KWH48359	30	48	2.00E-17	"PixA"
103	Pseudomonas	fluorescens	H24	KJZ63571	30	50	2.00E-17	"PixA"
104	Xenorhabdus	NBAII XenSa04	NBAII XenSa04	WP_047681901	29	46	3.00E-17	True PixA, partial conservation of CD1 (W to M, R to N); and CD2

								(W to A, D to E)
105	Burkholderia	MSh2	MSh2	KEZ03662	30	48	3.00E-17	"PixA"
106	Burkholderia	MSh1	MSh1	KFG97921	30	48	3.00E-17	"PixA"
107	Burkholderia	pyrrocinia	Lyc2	KFL51991	31	47	4.00E-17	"PixA"
108	Burkholderia	cenocepacia	KC-01	ESS40820	31	47	4.00E-17	AidA
109	Burkholderia	anthina	AZ-4-2-30-S1-D7	KVE10228	31	47	5.00E-17	"PixA"
110	Burkholderia	anthina	MSMB0849	KVH10857	31	47	5.00E-17	"PixA"
111	Burkholderia	anthina	MSMB0848	KVH11093	31	47	5.00E-17	"PixA"
112	Burkholderia	anthina	MSMB1497	KVM86453	31	47	5.00E-17	"PixA"
113	Burkholderia	anthina	MSMB1496	KVN57654	31	47	5.00E-17	"PixA"
114	Burkholderia	anthina	MSMB1506WGS	KVX38030	31	47	5.00E-17	"PixA"
115	Burkholderia	anthina	MSMB649WGS	KWH59010	31	47	5.00E-17	"PixA"
116	Burkholderia	anthina	AZ-4-2-10-S1-D7	KWZ32831	31	47	5.00E-17	"PixA"
117	Burkholderia	9120	9120	WP_035559998	32	47	5.00E-17	"PixA"
118	Pseudomonas	GM16	GM16	EJM05219	31	48	6.00E-17	"PixA"
119	Pseudomonas	GM24	GM24	EJM25417	31	48	6.00E-17	"PixA"
120	Burkholderia	ubonensis	MSMB2187WGS	KWD12304	36	49	6.00E-17	"PixA"
121	Burkholderia	ubonensis	MSMB2188WGS	KWD16904	36	49	6.00E-17	"PixA"
122	Burkholderia	ubonensis	MSMB1941WGS	KWQ02218	36	49	6.00E-17	"PixA"
123	Pseudomonas	GM67	GM67	EJM94480	30	50	8.00E-17	"PixA"
124	Pseudomonas	GM60	GM60	EJM86080	30	50	9.00E-17	"PixA"
125	Pseudomonas	fluorescens	H16	KII35730	31	48	9.00E-17	"PixA"
126	Pseudomonas	RIT-PI-r	RIT-PI-r	KPH00125	31	48	9.00E-17	"PixA"
127	Pseudomonas	batumici	UCM B-321	KIH81277	31	49	1.00E-16	AidA
128	Bradyrhizobium	jicamae	PAC68	KRQ92913	31	46	1.00E-	"PixA"

							16	
129	Hyalangium	minutum	minutum	KFE61977	31	49	2.00E-16	AidA
130	Pseudomonas	GM18	GM18	EJM19633	29	47	2.00E-16	"PixA" CD2 (G to A)
131	Burkholderia	stagnalis	MSMB1136	KVC68700	34	51	2.00E-16	"PixA" CD1 (G to N)
132	Burkholderia	stagnalis	MSMB1135	KVN25299	34	51	2.00E-16	"PixA" CD1 (G to N)
133	Burkholderia	stagnalis	MSMB1968	KVN67593	34	51	2.00E-16	"PixA" CD1 (G to N)
134	Burkholderia	stagnalis	MSMB1604WGS	KVO63080	34	51	2.00E-16	"PixA" CD1 (G to N)
135	Burkholderia	stagnalis	MSMB1638WGS	KVP04578	34	51	2.00E-16	"PixA" CD1 (G to N)
136	Burkholderia	stagnalis	MSMB1487WGS	KVW97944	34	51	2.00E-16	"PixA" CD1 (G to N)
137	Burkholderia	stagnalis	MSMB1521WGS	KVZ13594	34	51	2.00E-16	"PixA" CD1 (G to N)
138	Burkholderia	stagnalis	MSMB1960WGS	KWA52221	34	51	2.00E-16	"PixA" CD1 (G to N)
139	Burkholderia	stagnalis	MSMB1943WGS	KWA57423	34	51	2.00E-16	"PixA" CD1 (G to N)
140	Burkholderia	stagnalis	MSMB1956WGS	KWA62384	34	51	2.00E-16	"PixA" CD1 (G to N)
141	Burkholderia	stagnalis	MSMB2169WGS	KWC95061	34	51	2.00E-16	"PixA" CD1 (G to N)
142	Burkholderia	stagnalis	MSMB2168WGS	KWC99373	34	51	2.00E-16	"PixA" CD1 (G to N)
143	Burkholderia	stagnalis	MSMB656WGS	KWH67998	34	51	2.00E-16	"PixA" CD1 (G to N)
144	Burkholderia	stagnalis	MSMB654WGS	KWH76211	34	51	2.00E-16	"PixA" CD1 (G to N)
145	Burkholderia	stagnalis	MSMB653WGS	KWH79032	34	51	2.00E-16	"PixA" CD1 (G to N)
146	Burkholderia	stagnalis	MSMB739WGS	KWI63016	34	51	2.00E-16	"PixA" CD1 (G to N)
147	Burkholderia	stagnalis	MSMB756WGS	KWK33413	34	51	2.00E-16	"PixA" CD1 (G to N)
148	Burkholderia	stagnalis	MSMB757WGS	KWK36657	34	51	2.00E-16	"PixA" CD1 (G to N)
149	Burkholderia	stagnalis	MSMB780WGS	KWK65650	34	51	2.00E-16	"PixA" CD1 (G to N)
150	Burkholderia	stagnalis	MSMB800WGS	KWN07621	34	51	2.00E-16	"PixA" CD1 (G to N)
151	Burkholderia	stagnalis	MSMB802WGS	KWN29939	34	51	2.00E-16	"PixA" CD1 (G to N)
152	Burkholderia	stagnalis	MSMB808WG	KWN43313	34	51	2.00E-	"PixA" CD1 (G to

			S				16	N)
153	Burkholderia	stagnalis	MSMB810WGS	KWN45198	34	51	2.00E-16	"PixA" CD1 (G to N)
154	Burkholderia	stagnalis	MSMB809WGS	KWN46952	34	51	2.00E-16	"PixA" CD1 (G to N)
155	Burkholderia	stagnalis	MSMB835WGS	KWN83274	34	51	2.00E-16	"PixA" CD1 (G to N)
156	Burkholderia	stagnalis	MSMB836WGS	KWN94208	34	51	2.00E-16	"PixA" CD1 (G to N)
157	Burkholderia	stagnalis	MSMB837WGS	KWN98893	34	51	2.00E-16	"PixA" CD1 (G to N)
158	Burkholderia	stagnalis	MSMB1511WGS	KWO76478	34	51	2.00E-16	"PixA" CD1 (G to N)
159	Burkholderia	stagnalis	MSMB2051	KVL95010	32	51	2.00E-16	"PixA" CD1 (G to N)
160	Burkholderia	stagnalis	MSMB2052	KVL98253	32	51	2.00E-16	"PixA" CD1 (G to N)
161	Burkholderia	stagnalis	MSMB2053	KVM03803	32	51	2.00E-16	"PixA" CD1 (G to N)
162	Burkholderia	stagnalis	MSMB1512WGS	KVX61949	32	51	2.00E-16	"PixA" CD1 (G to N)
163	Burkholderia	stagnalis	MSMB719WGS	KWI25677	32	51	2.00E-16	"PixA" CD1 (G to N)
164	Burkholderia	stagnalis	MSMB729WGS	KWI73754	32	51	2.00E-16	"PixA" CD1 (G to N)
165	Burkholderia	stagnalis	MSMB1147	KVM86152	32	51	2.00E-16	"PixA" CD1 (G to N)
166	Burkholderia	stagnalis	MSMB1966	KVN23241	32	51	2.00E-16	"PixA" CD1 (G to N)
167	Burkholderia	stagnalis	MSMB1967	KVN37950	32	51	2.00E-16	"PixA" CD1 (G to N)
168	Burkholderia	stagnalis	MSMB1570WGS	KVN68889	32	51	2.00E-16	"PixA" CD1 (G to N)
169	Burkholderia	stagnalis	MSMB1602WGS	KVO52286	32	51	2.00E-16	"PixA" CD1 (G to N)
170	Burkholderia	stagnalis	MSMB1606WGS	KVO66723	32	51	2.00E-16	"PixA" CD1 (G to N)
171	Burkholderia	stagnalis	MSMB1470WGS	KVW67943	32	51	2.00E-16	"PixA" CD1 (G to N)
172	Burkholderia	stagnalis	MSMB1477WGS	KVW73317	32	51	2.00E-16	"PixA" CD1 (G to N)
173	Burkholderia	stagnalis	MSMB1514WGS	KVX81808	32	51	2.00E-16	"PixA" CD1 (G to N)
174	Burkholderia	stagnalis	MSMB618WGS	KWH29302	32	51	2.00E-16	"PixA" CD1 (G to N)
175	Burkholderia	stagnalis	MSMB642WGS	KWH60258	32	51	2.00E-16	"PixA" CD1 (G to N)
176	Burkholderia	stagnalis	MSMB658WG	KWH88026	32	51	2.00E-	"PixA" CD1 (G to

			S				16	N)
177	Burkholderia	stagnalis	MSMB657WGS	KWH89861	32	51	2.00E-16	"PixA" CD1 (G to N)
178	Burkholderia	stagnalis	MSMB659WGS	KWI02369	32	51	2.00E-16	"PixA" CD1 (G to N)
179	Burkholderia	stagnalis	MSMB660WGS	KWI06725	32	51	2.00E-16	"PixA" CD1 (G to N)
180	Burkholderia	stagnalis	MSMB749WGS	KWK05003	32	51	2.00E-16	"PixA" CD1 (G to N)
181	Burkholderia	stagnalis	MSMB754WGS	KWK31753	32	51	2.00E-16	"PixA" CD1 (G to N)
182	Burkholderia	stagnalis	MSMB774WGS	KWK51815	32	51	2.00E-16	"PixA" CD1 (G to N)
183	Burkholderia	stagnalis	MSMB777WGS	KWK60692	32	51	2.00E-16	"PixA" CD1 (G to N)
184	Burkholderia	stagnalis	MSMB811WGS	KWN59881	32	51	2.00E-16	"PixA" CD1 (G to N)
185	Burkholderia	stagnalis	MSMB826WGS	KWN74895	32	51	2.00E-16	"PixA" CD1 (G to N)
186	Burkholderia	stagnalis	MSMB845WGS	KWO16740	32	51	2.00E-16	"PixA" CD1 (G to N)
187	Burkholderia	stagnalis	MSMB847WGS	KWO25883	32	51	2.00E-16	"PixA" CD1 (G to N)
188	Burkholderia	stagnalis	MSMB846WGS	KWO38452	32	51	2.00E-16	"PixA" CD1 (G to N)
189	Xenorhabdus	cabanillasii	JM26	CDL86366	30	47	3.00E-16	True PixA, partial conservation of CD1 (W to M, R to S); and CD2 (W to A, D to E)
190	Pseudomonas	GM21	GM21	EJM14711	30	48	4.00E-16	"PixA"
191	Bradyrhizobium	yuanmingense	yuanmingense	WP_036032112	30	46	5.00E-16	IBP
192	Streptosporangium	roseum	DSM 43021	ACZ83107	34	53	7.00E-16	"PixA" CD1 (G to N) Gram +
193	Burkholderia	stagnalis	MSMB1543	KVN05792	32	50	7.00E-16	"PixA" CD1 (G to N)
194	Burkholderia	stagnalis	MSMB2170WGS	KWD97092	32	50	7.00E-16	"PixA" CD1 (G to N)
195	Burkholderia	stagnalis	MSMB2171WGS	KWE12205	32	50	7.00E-16	"PixA" CD1 (G to N)
196	Burkholderia	stagnalis	MSMB1545WGS	KWO75664	32	50	7.00E-16	"PixA" CD1 (G to N)
197	Bradyrhizobium	elkanii	elkanii	WP_028337824	30	46	1.00E-15	IBP
198	Bradyrhizobium	pachyrhizi	pachyrhizi	WP_050385987	30	46	1.00E-15	"PixA"

199	Burkholderia	ubonensis	MSMB1190	KVD59004	35	49	2.00E-15	"PixA"
200	Burkholderia	ubonensis	MSMB1181WGS	KVU24550	35	49	2.00E-15	"PixA"
201	Burkholderia	ubonensis	MSMB2125WGS	KWC54878	35	49	2.00E-15	"PixA"
202	Burkholderia	ubonensis	MSMB834WGS	KWO17313	35	49	2.00E-15	"PixA"
203	Burkholderia	ubonensis	MSMB1635WGS	KWO89976	35	49	2.00E-15	"PixA"
204	Burkholderia	ubonensis	MSMB1157	KWZ53997	35	49	2.00E-15	"PixA"
205	Burkholderia	ubonensis	MSMB1592WGS	KVO01386	35	49	3.00E-15	"PixA"
206	Burkholderia	ubonensis	MSMB1601WGS	KVO38485	35	49	3.00E-15	"PixA"
207	Burkholderia	ubonensis	MSMB1791WGS	KVP30485	35	49	3.00E-15	"PixA"
208	Shewanella	baltica	OS155	ABN60197	33	50	4.00E-15	"PixA"
209	Shewanella	baltica	OS117	AEH12593	33	50	4.00E-15	"PixA"
210	Shewanella	baltica	BA175	AEG12786	33	50	6.00E-15	"PixA"
211	Shewanella	baltica	OS183	EHQ13675	33	50	6.00E-15	"PixA"
212	Burkholderia	stagnalis	ABCPW-33	KVD85279	32	50	7.00E-15	"PixA" CD1 (G to N)
213	Burkholderia	ambifaria	IOP40-10	EDT06278	34	48	1.00E-14	"PixA" CD1 (G to N)
214	Burkholderia	ubonensis	MSMB1164	KVC98407	31	47	1.00E-14	"PixA" CD1 (G to N)
215	Burkholderia	ubonensis	MSMB1199	KVD61962	31	47	1.00E-14	"PixA" CD1 (G to N)
216	Burkholderia	ubonensis	MSMB0011	KVG32430	31	47	1.00E-14	"PixA" CD1 (G to N)
217	Burkholderia	ubonensis	MSMB1605WGS	KVO57739	31	47	1.00E-14	"PixA" CD1 (G to N)
218	Burkholderia	ubonensis	MSMB1637WGS	KVP02099	31	47	1.00E-14	"PixA" CD1 (G to N)
219	Burkholderia	ubonensis	MSMB1802WGS	KVP53988	31	47	1.00E-14	"PixA" CD1 (G to N)
220	Burkholderia	ubonensis	MSMB1804WGS	KVP65443	31	47	1.00E-14	"PixA" CD1 (G to N)
221	Burkholderia	ubonensis	MSMB1806WGS	KVP74861	31	47	1.00E-14	"PixA" CD1 (G to N)
222	Burkholderia	ubonensis	MSMB1810WGS	KVP88491	31	47	1.00E-14	"PixA" CD1 (G to N)

223	Burkholderia	ubonensis	MSMB2026WGS	KVR78744	31	47	1.00E-14	"PixA" CD1 (G to N)
224	Burkholderia	ubonensis	MSMB1162WGS	KVT82516	31	47	1.00E-14	"PixA" CD1 (G to N)
225	Burkholderia	ubonensis	MSMB1225WGS	KVU95636	31	47	1.00E-14	"PixA" CD1 (G to N)
226	Burkholderia	ubonensis	MSMB1472WGS	KVW74317	31	47	1.00E-14	"PixA" CD1 (G to N)
227	Burkholderia	ubonensis	MSMB1488WGS	KVX22302	31	47	1.00E-14	"PixA" CD1 (G to N)
228	Burkholderia	ubonensis	MSMB1517WGS	KVX91833	31	47	1.00E-14	"PixA" CD1 (G to N)
229	Burkholderia	ubonensis	MSMB1530WGS	KVZ24952	31	47	1.00E-14	"PixA" CD1 (G to N)
230	Burkholderia	ubonensis	MSMB1559WGS	KVZ50254	31	47	1.00E-14	"PixA" CD1 (G to N)
231	Burkholderia	ubonensis	MSMB2087WGS	KWA86129	31	47	1.00E-14	"PixA" CD1 (G to N)
232	Burkholderia	ubonensis	MSMB2098WGS	KWB36914	31	47	1.00E-14	"PixA" CD1 (G to N)
233	Burkholderia	ubonensis	MSMB2118WGS	KWC20593	31	47	1.00E-14	"PixA" CD1 (G to N)
234	Burkholderia	ubonensis	MSMB2121WGS	KWC22279	31	47	1.00E-14	"PixA" CD1 (G to N)
235	Burkholderia	ubonensis	MSMB2152WGS	KWF05104	31	47	1.00E-14	"PixA" CD1 (G to N)
236	Burkholderia	ubonensis	MSMB720WGS	KWI57165	31	47	1.00E-14	"PixA" CD1 (G to N)
237	Burkholderia	ubonensis	MSMB728WGS	KWI72907	31	47	1.00E-14	"PixA" CD1 (G to N)
238	Burkholderia	ubonensis	MSMB792WGS	KWK94563	31	47	1.00E-14	"PixA" CD1 (G to N)
239	Burkholderia	ubonensis	MSMB791WGS	KWN00095	31	47	1.00E-14	"PixA" CD1 (G to N)
240	Burkholderia	ubonensis	MSMB803WGS"	KWN40082	31	47	1.00E-14	"PixA" CD1 (G to N)
241	Ralstonia	solanacearum	AW1	AAC45946	35	46	1.00E-14	AidA
242	Ralstonia	solanacearum	K60	CCF97746	35	46	1.00E-14	AidA
243	Burkholderia	pyrrocinia	DSM 10685	AKM05012	29	47	1.00E-14	"PixA"
244	Chromobacterium	LK11	LK11	KMN77888	31	50	4.00E-14	"PixA" CD1 (G to D)
245	Caulobacter	henricii	henricii	WP_035070873	31	48	4.00E-14	"PixA"
246	Streptomyces	ipomoeae	91-03	EKX60943	30	51	6.00E-14	"PixA", High G+C content, Gram +

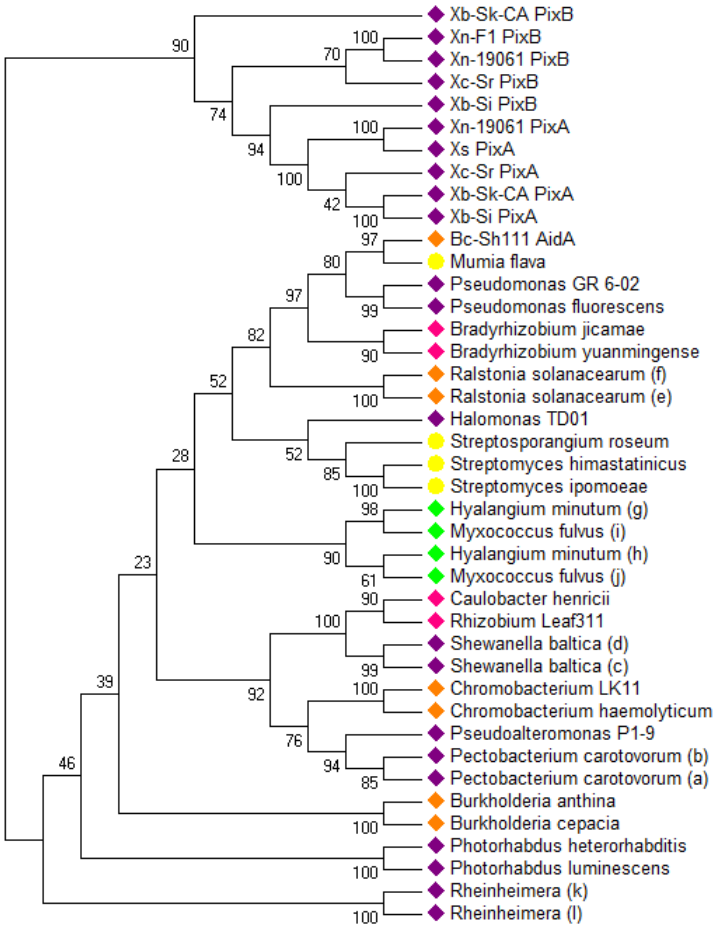
247	Ralstonia	solanacearum	CFBP2957	CBJ41502	34	45	2.00E-13	AidA
248	Ralstonia	solanacearum	23-10BR	KFX30785	34	45	2.00E-13	AidA
249	Ralstonia	solanacearum	CIP120	OAI65516	34	45	2.00E-13	AidA
250	Ralstonia	solanacearum	P597	OAI79107	34	45	2.00E-13	AidA
251	Chromobacterium	haemolyticum	haemolyticum	WP_043641426	31	50	2.00E-13	"PixA" CD1 (G to D)
252	Burkholderia	cepacia	MSMB591WGS	KWF85674	30	47	2.00E-13	"PixA"
253	Ralstonia	solanacearum	UW551	EAP73005	33	45	3.00E-13	AidA
254	Ralstonia	solanacearum	NCPPB 909	KEI33095	33	45	3.00E-13	AidA
255	Ralstonia	solanacearum	NCPPB 282	KFX80041	33	45	3.00E-13	AidA
256	Ralstonia	solanacearum	POPS2	KFX82070	33	45	3.00E-13	AidA
257	Ralstonia	solanacearum	CFIA906	KFZ92724	33	45	3.00E-13	AidA
258	Ralstonia	solanacearum	IPO1609	CEJ20796	33	45	3.00E-13	AidA
259	Ralstonia	solanacearum	UY031	ALF86606	33	45	3.00E-13	AidA
260	Burkholderia	cepacia	RF2-non_BP3	KUY54567	29	45	3.00E-13	"PixA"
261	Burkholderia	cepacia	RF4-BP95	KUY72452	29	45	3.00E-13	"PixA"
262	Burkholderia	cepacia	RF7-non_BP1	KUY99800	29	45	3.00E-13	"PixA"
263	Burkholderia	cepacia	RF7-non_BP4	KUZ03935	29	45	3.00E-13	"PixA"
264	Burkholderia	cepacia	MSMB0856	KVH38414	29	45	3.00E-13	"PixA"
265	Streptomyces	himastatinicus	ATCC 53653	EFL20450	35	51	5.00E-13	"PixA" CD1 (G to R) High G+C content, Gram +
266	Ralstonia	solanacearum	KACC 10722	AMP36154	32	45	6.00E-13	IBP
267	Burkholderia	territorii	MSMB1301WGS	KVV45831	29	46	7.00E-13	"PixA"
268	Burkholderia	territorii	MSMB1502WGS	KVX38723	29	46	7.00E-13	"PixA"
269	Ralstonia	MD27	MD27	KMW46428	28	42	8.00E-13	"PixA"
270	Burkholderia	ambifaria	AMMD	ABI90542	31	45	8.00E-13	"PixA" CD1 (G to D) CD2 (P to A)

271	Hyalangium	minutum	DSM 14724	KFE61976	30	49	9.00E-13	AidA CD1 (G to N)
272	Pseudomonas	protegens	Pf-5	AAY91957	29	45	1.00E-12	"PixA" CD2 (G to A)
273	Pseudomonas	protegens	CHA0	AGL84519	29	45	1.00E-12	"PixA" CD2 (G to A)
274	Ralstonia	solanacearum	PSI07	CBJ49535	33	45	1.00E-12	"PixA"
275	Ralstonia	solanacearum	FQY_4	AGH85845	33	46	1.00E-12	AidA
276	Ralstonia	solanacearum	SD54	ESS51648	33	46	1.00E-12	AidA
277	Ralstonia	solanacearum	Phyl III-seqv23	CUV46738	33	46	1.00E-12	AidA
278	Ralstonia	solanacearum	eST002_mST005	CUV54360	33	46	1.00E-12	AidA
279	Ralstonia	solanacearum	P781	OAI65750	33	46	1.00E-12	AidA
280	Ralstonia	AU12-08	AU12-08	EPX95752	27	42	2.00E-12	"PixA"
281	Ralstonia	NT80	NT80	GAQ28071	27	42	2.00E-12	"PixA"
282	Pseudomonas	fluorescens	fluorescens	WP_019093798	29	45	2.00E-12	"PixA" CD2 (G to A)
283	Burkholderia	cepacia	DWS 37UF10B-2	KGC06457	29	45	3.00E-12	"PixA"
284	Burkholderia	cepacia	NRF60-BP8	KVA10947	29	45	3.00E-12	"PixA"
285	Burkholderia	cepacia	MSMB1826	KVL10081	29	45	3.00E-12	"PixA"
286	Burkholderia	cepacia	MSMB1835	KVL41003	29	45	3.00E-12	"PixA"
287	Burkholderia	cepacia	MSMB2157WGS	KWE64093	29	45	3.00E-12	"PixA"
288	Ralstonia	solanacearum	Po82	AEG67471	32	45	3.00E-12	AidA
289	Ralstonia	solanacearum	P673	EUJ16320	32	45	3.00E-12	AidA
290	Ralstonia	solanacearum	UW163	AMP68876	32	45	3.00E-12	AidA
291	Ralstonia	solanacearum	IBSBF1503	AMP74218	32	45	3.00E-12	AidA
292	Ralstonia	solanacearum	CFBP6783	OAI75034	32	45	3.00E-12	AidA
293	Ralstonia	syzygii	R24	CCA85668	31	45	3.00E-12	"PixA"
294	Pseudomonas	protegens	Cab57	BAO62075	29	45	4.00E-12	"PixA" CD2 (G to A)

295	Ralstonia	R229	R229	CCA80947	31	45	4.00E-12	"PixA"
296	Photorhabdus	heterorhabditis	heterorhabditis	WP_054479578	30	50	4.00E-12	"PixA" CD2 (W to H, D to P)
297	Burkholderia	ambifaria	MC40-6	ACB65610	30	44	6.00E-12	"PixA" CD1 (G to D); CD2 (P to A)
298	Pseudomonas	entomophila	L48	CAK17669	31	44	1.00E-11	"PixA"
299	Burkholderia	ABCPW 111	ABCPW 111	KGS08272	30	45	2.00E-11	"PixA"
300	Burkholderia	ABCPW 111	MSMB0265	KVG48864	30	45	2.00E-11	"PixA"
301	Burkholderia	ABCPW 111	MSMB0266	KVG49456	30	45	2.00E-11	"PixA"
302	Burkholderia	ABCPW 111	MSMB2040	KVG86324	30	45	2.00E-11	"PixA"
303	Burkholderia	ABCPW 111	MSMB2041	KVG90602	30	45	2.00E-11	"PixA"
304	Burkholderia	ABCPW 111	MSMB2042	KVH00256	30	45	2.00E-11	"PixA"
305	Burkholderia	ABCPW 111	MSMB0852	KVH27067	30	45	2.00E-11	"PixA"
306	Burkholderia	ABCPW 111	MSMB617WGS	KWH37574	30	45	2.00E-11	"PixA"
307	Burkholderia	ABCPW 111	BDU18	KWZ37749	30	45	2.00E-11	"PixA"
308	Pseudomonas	P482	P482	KDN99481	32	45	2.00E-11	"PixA"
309	Pseudomonas	donghuensis	donghuensis	WP_010221516	32	45	3.00E-11	"PixA"
310	Burkholderia	MSMB1498	MSMB1498	KVK82119	30	45	4.00E-11	"PixA"
311	Burkholderia	pseudomallei	BDU19	KWZ48220	30	45	5.00E-11	IBP CD1 (S/T to N)
312	Rhizobium/Agrobacterium group	Rhizobium	Leaf311	KQQ58627	32	47	6.00E-11	"PixA" CD1 (G to E, R1 to H)
313	Burkholderia	pyrrocinia	Lyc2	KFL52734	30	44	8.00E-11	"PixA" CD1 (G to D, R1 to Y); CD2 (P to A)
314	Myxococcus	fulvus	124B02	AKF85240	30	53	2.00E-10	"PixA" CD1 (G to S)
315	Pectobacterium	carotovorum	YC D46	KHS91609	33	44	2.00E-10	"PixA" CD1 (R2 to T)
316	Burkholderia	ubonensis	MSMB2006WGS	KVQ79384	43	60	1.00E-09	"PixA" NO CD2, PixB fragment
317	Rheinheimera	A13L	A13L	EGM79634	28	46	2.00E-08	"PixA" CD1 (R2 to T)
318	Rheinheimera	KL1	KL1	KOO59823	28	45	3.00E-08	"PixA" CD1 (R2 to T)

319	Photorhabdus	luminescens	luminescens	WP_040154142	27	44	3.00E-08	"PixA" CD1 (G to A,); CD2 (W to H)
320	Myxococcus	fulvus	124B02	AKF85239	27	48	5.00E-08	"PixA" CD1 (G to N, D to N); CD2 (G to S)
321	Photorhabdus	luminescens	luminescens	WP_049583339	38	52	5.00E-08	"PixA" NO CD2
322	Pectobacterium	carotovorum	carotovorum	WP_052325947	31	47	1.00E-07	PixC
323	Photorhabdus	luminescens	luminescens	WP_052739541	37	51	3.00E-07	"PixA" NO CD2
324	Photorhabdus	luminescens	BA1	EYU13442	28	44	4.00E-07	"PixA" CD2 (W to H) Part of multiple gene duplication event (X4) flanked by repeat region with varying degrees of conservation
325	Pectobacterium	carotovorum	carotovorum	WP_052238864	30	46	5.00E-07	PixC
326	Pectobacterium	carotovorum		WP_052238948	30	46	6.00E-07	PixC
327	Photorhabdus	luminescens	LN2	KGM28847	28	44	6.00E-07	"PixA" CD1 (G to A); CD2 (W to H)
328	Photorhabdus	luminescens	BA1	EYU13441	36	51	7.00E-07	"PixA" NO CD2 Part of multiple gene duplication event (X4) flanked by repeat region with varying degrees of conservation
329	Burkholderia	ubonensis	MSMB1069WGS	KVS97884	36	48	1.00E-06	"PixA" 115AA Is smallest PixB Identified with perfect conservation of domains
330	Photorhabdus	heterorhabditis	VMG	KOY61631	24	45	2.00E-06	"PixA" CD2 (W to H) Part of multiple gene duplication event (X4) with varying degrees of conservation
331	Photorhabdus	asymbiotica	asymbiotica	WP_036773243	29	44	4.00E-06	"PixA" CD2 (W to H)
332	Yersinia	bercovieri	127/84	CNH75565	33	46	5.00E-06	PixC

333	Pseudoalteromonas	P1-9	P1-9	KPV95834	28	44	8.00E-06	"PixA" CD1 (W to F, R2 to S)
334	Yersinia	kristensenii	IP6945	CNH07449	31	45	8.00E-06	PixC
335	Burkholderia	pyrrocinia	pyrrocinia	WP_017331102	29	43	1.00E-04	"PixA" CD1 (R2 to A)
336	Rahnella	aquatilis	CIP 78.65	AEX52531	36	50	2.00E-04	PixC
337	Rahnella	aquatilis	ATCC 33071	KFD06648	36	50	2.00E-04	PixC
338	Pectobacterium	carotovorum	CFIA1033	KFF65531	25	41	8.00E-04	"PixA" CD1 (G to D, R to T)
339	Pectobacterium	carotovorum	carotovorum	KFX18911	24	41	0.003	"PixA" CD1 (G to D, R to T)



Appendix E:

Figure S1: Molecular Phylogenetic analysis by Maximum Likelihood method of PixB homologs. The evolutionary history was inferred by using the Maximum Likelihood method based on the Poisson correction model. The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 4.9310)). The analysis involved 41 amino acid sequences. There was a total of 255 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

Appendix F:

Table S5: Genus level residue analysis of the PixB motif in identified PixB homologs. Linker domain of all identical sequences occurring in each strain were grouped and averaged across species to derive a genus level average for each class. Separate analysis was conducted for both the PixA and PixC variants in a similar manner. Conserved domains were analyzed for residue deviation from consensus sequences

Alphaproteobacteria

	Bradyrhizobium	*Caulobacter*	*Rhizobium/ Agrobacterium group*	Average	% Residues of Avg	
Ala	6	6	9	7	7	CD1 conservation
Arg	6	3	1	3	3	
Asn	7	6	5	6	6	
Asp	3	5	5	4	4	G to (E)
Gln	1	7	4	4	4	D
Glu	5	2	4	3	4	X
Gly	6	8	8	8	8	X
His	2	1	2	2	2	R1 to (H)
Ile	6	7	5	6	6	W
Leu	6	7	9	7	7	R2
Lys	2	1	2	2	2	X
Met	1	2	2	2	2	X
Phe	7	4	3	4	4	S/T
Pro	7	7	7	7	7	CD2 conservation
Ser	11	7	6	8	8	G
Thr	7	14	12	11	11	Y/C
Tyr	5	7	8	7	7	X
Val	6	7	8	7	7	X
Trp	3	2	2	2	2	W
Cys	2	0	0	1	1	D
Residue #	99	100	100	100		P

Betaproteobacteria

	Burkholderia	*Ralstonia*	*Chromobacterium*	Average	% Residues of Avg	
Ala	5	3	3	4	4.2	CD1 conservation
Arg	5	5	3	5	5.1	
Asn	5	4	8	4	4.4	G to (N), (D)
Asp	5	5	6	5	5.2	D
Gln	4	3	9	3	3.5	X
Glu	6	4	3	5	4.9	X
Gly	6	7	5	7	6.5	R1 to (Y)
His	1	1	0	1	1.1	W
Ile	6	6	7	6	6.0	R2 to (A)
Leu	6	4	5	5	5.1	X
Lys	4	5	4	5	4.8	X
Met	0	0	0	0	0.3	S/T to (N)
Phe	5	6	3	5	5.4	CD2 conservation
Pro	7	9	8	8	8.1	
Ser	8	5	10	7	6.5	G
Thr	8	9	8	8	8.4	Y/C
Tyr	7	6	8	6	6.2	X
Val	7	8	13	7	7.5	X
Trp	3	4	4	4	3.7	W
Cys	2	4	1	3	3.1	D
Residue #	100	99	108	100		P to (A)

Gammaproteobacteria

	Xenorhabdus	*Pseudomonas*	*Shewanella*	*Halomonas*	*Rheinheimera*	*Photorhabdus*	*Pectobacterium*	*Pseudoalteromonas*	Average	% Residues of Avg	CD1 conservation
Ala	6	5	3	1	7	6	5	3	4	4.3	G to (L), (A), (D)
Arg	3	4	3	4	1	5	3	6	4	3.5	D to (S)
Asn	6	6	6	8	9	5	5	10	7	6.4	X
Asp	7	5	5	8	6	6	7	6	6	6.1	X
Gln	2	3	3	4	4	6	6	7	4	4.2	R1 to
Glu	5	4	4	3	1	3	5	4	4	3.4	W to (F)
Gly	6	5	10	4	10	8	10	9	8	7.3	R2 to (H), (T), (S)
His	2	2	1	1	2	3	2	1	2	1.6	X
Ile	5	6	4	7	4	9	5	3	5	5.0	X
Leu	9	5	10	7	7	9	6	3	7	6.6	S/T
Lys	7	4	1	1	2	5	3	3	3	3.1	CD2 conservation
Met	2	0	4	1	1	2	1	0	1	1.2	
Phe	4	5	4	4	7	3	6	9	5	4.9	G to (A)
Pro	10	7	7	9	5	5	6	4	6	6.1	Y/C
Ser	9	10	12	6	10	10	9	12	10	9.3	X
Thr	8	7	10	12	13	7	7	5	9	8.2	X
Tyr	8	4	8	6	7	8	8	6	7	6.5	W to (M), (H)
Val	12	9	9	8	7	6	6	11	9	8.2	D to (P)
Trp	3	3	2	4	2	2	3	1	3	2.4	P
Cys	0	5	1	1	0	1	0	0	1	1.0	D
Residue #	112	100	107	99	102	109	105	105	105		P to (A)

Deltaproteobacteria

	Hyalangium	*Myxococcus*	Average	% Residues of Avg	
Ala	5	6	6	5	CD1 conservation
Arg	2	3	3	3	
Asn	5	7	6	6	G to (N), (S)
Asp	4	5	5	4	D to (N)
Gln	8	6	7	7	X
Glu	1	1	1	1	X
Gly	10	6	8	8	R1
His	0	1	0	0	W
Ile	5	4	5	5	R2
Leu	6	8	7	7	X
Lys	2	3	2	2	X
Met	2	1	1	1	S/T
Phe	6	4	5	5	CD2 conservation
Pro	6	7	7	7	
Ser	9	7	8	8	G to (S)
Thr	13	11	12	12	Y/C
Tyr	5	5	5	5	X
Val	6	10	8	8	X
Trp	2	2	2	2	W
Cys	0	1	1	1	D
Residue #	105	101	103		P

PixA

	Xenorhabdus	Average	% Residues of Avg	CD1 conservation
Ala	3	3	3	
Arg	4	4	3	
Asn	6	6	5	**G**
Asp	8	8	7	**D**
Gln	3	3	2	**X**
Glu	6	6	5	**X**
Gly	5	5	4	**R1**
His	2	2	2	**W to (F), (L), (M)**
Ile	9	9	7	**R2 to (N), (S)**
Leu	7	7	6	**X**
Lys	9	9	8	**X**
Met	19	19	16	**S/T**
Phe	1	1	1	**CD2 conservation**
Pro	7	7	6	
Ser	6	6	5	**G**
Thr	3	3	3	**Y/C**
Tyr	8	8	7	**X**
Val	12	12	10	**X**
Trp	1	1	1	**W to (I), (V), (A)**
Cys	1	1	1	**D to (E)**
Residue #	119	119	3	**P to (T)**

PixC

	Pectobacterium	*Rahnella*	*Yersinia*	*Salmonella*	Average	% Residues of Avg	CD1 conservation
Ala	3	2	2	3	2	2	
Arg	1	3	3	3	3	2	
Asn	7	6	5	4	5	4	G
Asp	4	4	5	7	5	4	D
Gln	2	3	3	2	2	2	X
Glu	2	3	2	3	3	2	X
Gly	28	22	23	21	24	20	R1
His	1	2	2	3	2	2	W
Ile	7	6	8	7	7	6	R2
Leu	4	5	4	4	4	3	X
Lys	6	4	5	6	5	4	X
Met	1	2	3	2	2	2	S/T
Phe	2	3	3	3	3	2	CD2 conservation
Pro	5	4	5	6	5	4	
Ser	5	4	5	5	5	4	G
Thr	9	9	10	9	9	8	Y/C
Tyr	3	3	5	3	4	3	X
Val	4	9	6	4	6	5	X deletion
Trp	2	3	3	2	2	2	W
Cys	4	2	2	2	2	2	D
Residue #	141	116	119	100	119		

Appendix G:
Table S6: Predicted binding residues of each PixC variant.

Strain	Chlorine	Calcium	Glycerol	BETA-D-GLUCOSE
>Pectobacterium_carotovorum_947_PixC	V88 I89 T97 P98	I105	H48 A49 Q50 D51 R68 K102 F127 E129 C159 G160 S161 N162	
>Pectobacterium_carotovorum_948_PixC	V88 I89 T97 P98	I105	A49 Q50 D51 R68 K102 F127 E129	N189 N190
>Pectobacterium_carotovorum864_PixC	V88 I89 P98 V10 D11 A12 T73 L74	I105 I76	H48 A49 Q50 D51 R68 K102 F127 E129	
>Rahnella_aquatilis_PixC	V88 I89 T97 P98 V10 D11 A12 T73 L74	M76 T105	H48 A49 R50 D51 R68 G102 F127 E129	P13 I16 Y80 V148 G149
>Yersinia_bercovieri_PixC	V88 I89 P98	M76 V105	V49 Q50 D51 R68 G102 F130 E132	
>Yersinia_kristensenii_PixC	V88 I89 T97 P98	M76 T105	V49 Q50 D51 R68 A102 F130 E132	
>Salmonella_enterica539_PixC	V89 I90 T98 P99	M77 G173 T106	I50 Q51 D52 R69 A103 K105 F131 E133	
>Salmonella_enterica465_PixC	V89 I90 P99	M77 T106	W47 I50 Q51 D52 R69 A103 K105 E133	
>Salmonella_enterica_590_PixC	V89 I90 T98 P99	M77 T106	R69 A103 F131 E133 I50 Q51 D52	E14 M17 D83 Y84 Y188 K189

Appendix H:
Table S7: Predicted binding residues of *Xenorhabdus* PixB proteins

Binding Predictions						
Strain	**Chlorine**	**Calcium**	**Glycerol**	**BETA-D-GLUCOSE**	**Heme**	**Copper**
>XNC1_PixB	I11 D12 A13 V73 S74 L75 R89 A90 T100 V101	H108 S97 T100 K143 Q144 K77 D78	V49 K50 Q51 G52 K38			
>XBI1_PixB	F93 E94 S95 K82	D109 M151 L153 P154 D120	V54 Y55 N57 Q30 G167 Y169 F174 Y177		V14 V15 K16 V17 T79 L80	L90 L164 F166
>XNFF1_PixB	A91 P92 M105 D107	I115 D101 E145 N79	V52 S53 I54			Y156 E158 I175
>XCR1_PixB	K95 P96 A107 V11 D12 T80 L81	K83 H114	T41 P50 I56 G58 R75 Q111 I113 E145	Q14 M17 Y87 I168 Y169		
>XBKQ1_PixB	V10 D11 A12 T71 I72 I86 L87 T104	K74 A111	N43 V46 N47 T48 R66 L108 Y134 T136			

	Binding Predictions			
Strain	**Iron**	**N-ACETYL-D--GLUCOSAMINE**	**ISOPROPYL ALCOHOL**	**Sodium**
>XNC1_PixB			I11 A13 Q14 M17 Y39 Y81 A83 V160 I162 Y163 V172	
>XBI1_PixB				N173 F174
>XNFF1_PixB			I10 R11 E13 I16 L32 Y41 Y83 S84 I85 I160 V173 I175	
>XCR1_PixB				
>XBKQ1_PixB	R16 Y78 N156 A157 G158			

Appendix I:
Table S8: Predicted binding residues of *Xenorhabdus* PixA proteins

Binding Predictions					
Strain	**Chlorine**	**Calcium**	**Glycerol**	**BETA-D-GLUCOSE**	**Heme**
>XCR1_PixA	D92 S93 M109 P132	H128 E131 N168 M169 T139 N80	D49 V52 T53 R40 L185 I187 L193		I16 D19 F20 V31 S32 K41 Y42 Y84 S85 V86 I197 R12 D15
>XNC1_PixA	E91 P92 S105 L106 V10 K11 A12 S76 L77	S102 S105 N142 L144 I79 V113	E47 D48 V52 I159 I161 L167 M168 Y170 N40	E13 M16 Y83 Y162 R163	V10 K11 A12 I15 I32 K39 Y41 S84 A85 A160 Y162 I171
>XSR1_PixA	E91 P92 M103	H79 N99 L100 S102 V141 N142 E143 E110	V51 V52 C53 K40		
>XDD1_PixA	S91 P92 S116 M117	M111 Q112 M113 T153 I79 D80 F123 M124	K40 E47 D48 D25 I26 P29 Y166 Y181 V88 K89 M120 W148		
>XBI1_PixA	A91 P92 L108	N79 K104 M105 D107 N144 E145 I115	D48 V52 S53 V37 R71 E112 Y114 Y138 M140		
>XBKQ1_PixA	A91 P92 L108	I115 K104 M105 D107 N144 E145 N79	D48 V52 S53 R71 E112 Y114 Y138 M140 V37		
>XBP1_PixA	A91 P92 D107 L108 R11 T12 L77	I75 T76 N79 D101 K104 N144 E145 N125	D48 V52 S53		D14 I15 I16 D18 Y19 L32 H33 K40 Y41 S84 I85 G162 V173 I175 R11 R43 I75 L77 M8 I10 R11 I15 L32 D39 K40 Y41 I42 V87 E158 I160 G162 I175 P177
>XiS1v1_PixA	D92 P93 P108	M103 T104 N144 N149 M115 M80	E48 D49 L52		
>XBFFL1v2_PixA	A91 P92 M105 D107	D101 E145 N79 I115	V52 S53 I54		

Binding Predictions				
Strain	**Copper**	**Iron**	**N-ACETYL-D--GLUCOSAMINE**	**ISOPROPYL ALCOHOL**
>XCR1_PixA				
>XNC1_PixA				
>XSR1_PixA	L42 L87 V88 R156 T158 I171 V173			
>XDD1_PixA	V10 K11 A12 T76 L77 V10 L32 H33 H39 K40 Y41 V42 T87 F165 E167 I169 I182 M184 P186			
>XBI1_PixA	V87 Y156 E158 I175	I16 Y83 I160 G162 I163 Q170	I10 R11 T12 Y41 R43 T76 L77	
>XBKQ1_PixA	V87 Y156 E158 I175	I16 Y83 I160 G162 I163 Q170	I10 R11 T12 Y41 R43 T76 L77	
>XBP1_PixA	V87 L88 E158 I160			
>XiS1v1_PixA	E158 I160 A175 M89 E158 I160 E158 I160 A175	Y84 V161 G162 I163 L169 Q170		
>XBFFL1v2_PixA	Y156 E158 I175			I10 R11 E13 I16 L32 Y41 Y83 S84 I85 I160 V173 I175

Appendix J:
Table S9: Predicted Binding residues for PixB DDRP homologs

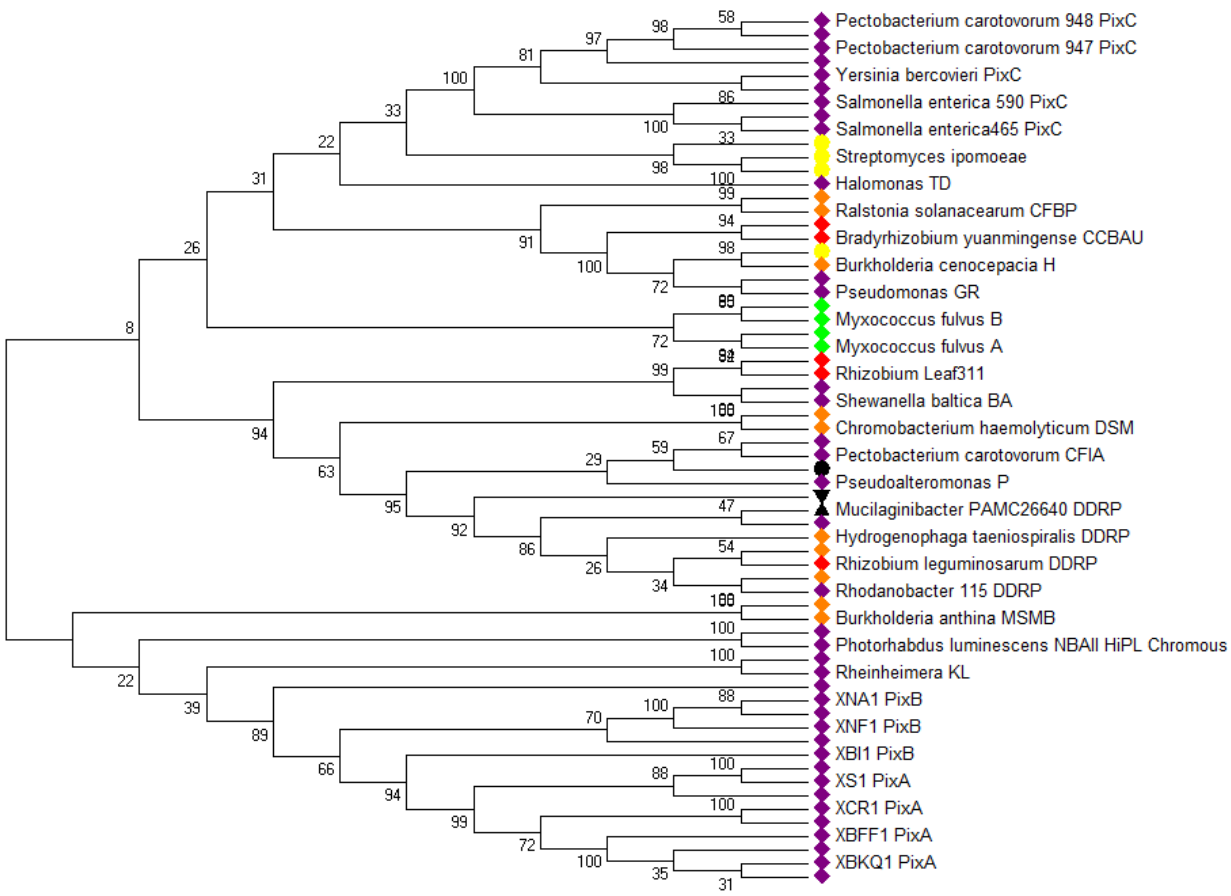
Binding Predictions							
Strain	**Chlorine**	**Calcium**	**Glycerol**	**BETA-D-GLUCOSE**	**N-ACETYL-D-GLUCOSAMINE**	**Sodium**	**Zinc**
>Burkholderia cepacia complex_DDRP	P96 R97 P106	D84 V113	G56 V57 K58 G59 Q76 N110 T141 D143				
>Chryseobacterium_DDRP	Q98 K99 G108 I18 D19 T20 S83 I84	Q86 V115	Q98 K99 G108 S78 Y112 S143 D145	D21 D90 T164 L165			
>Collimonas arenae_DDRP	Q97 Y98 Q107	D85 T114	I57 I58 S59 G60 V19 I20 D21 F48 A77 D111 S140 D142				
>Cytophagales bacterium B6_DDRP	A97 W98 I21 D22 T23 S82 E83	V103 V104 D106 V142 V143 T114	T57 I58 G59 G60 R77 N111 F137 Q139	D24 T27 S89 V158 R159			
>Hydrogenophaga taeniospiralis_DDRP	I18 D19 T20 S80 V81	A83 Q101 V102 N104 V145 G146 T112	T6 S75 N109 S140 D142 I55 I56 S57 G58	K24 D87 T161 L162		N170 L171	Q101 K147
>Mucilaginibacter sp. PAMC 26640_DDRP	I18 D19 T20 S80 I81	K95 N104 Q105 Q83 K112	D51 G55 I56 S57 W75 V109 S138 N140	E21 D87 T159 L160	V17 I18 D19 F46		
>Pseudoalteromonas flavipulchra_DDRP	K94 Q95 I17 D18 T19 S79 I80	Q100 V101 N103 V144 R145 N111	D50 I54 V55 S56 S74 N108 T139 D141	K23 D86 T160 L161			
>Rhizobium leguminosarum_DDRP	Q106 N115 Q116	D94 T123	A66 V67 T68 G69 S86 N120 T149 D151				
>Rhodanobacter sp. 115_DDRP	I20 D21 T22 S82 I83	T114 D85 Q97 N106 Q107	I57 V58 S59 G60 A77 N111 S140 D142	K26 D88 D89 T161 L162			

MSLVACLCVFFFRLLNAFFIRTFFQPDEYYQALEPAHRLVYGYGYLTWEWHEALRSSLHPLLYAFAYKLF
NIFILESWAVVVAPKIMGALIATTSDIYTYWFASRYWQSPSVAKVALGLSLISSWNWYVSTRSFLNNLEMALTAAGLSY
WPWHQYKLMPLLISCLFGFLSCIVRPTNSILWAFLAVCLLAKNIRRFSRLVSLLVSVSLVF
MLTSMLASLADRFFYGRWTFPIYNFLEFNVFRNLLIFYGSAPWHFYLLQGIPLILMGYLPLFCLAIYQYR
RTTLVGLCVFFVSAFSLIAHKEFRFLQPLHPVMITLCAKTIYQRRHINLAKLFVFVVVIFHLFIAYFFTR
VHEAGEIRVVELLRDDPMVESIGFLTPCHSTPWHSMFHRPDLISKSWFLTCEPPLHFESGTAENIQTYRD
QLDLFFDDPVTFLQSLPYEWPSHLVLFQPMQKVAMSELTEYYECKRVF**N**GYFHWDPRRSGDIIVYCRKEILMI

Appendix K:

Figure S2: Possible Eukaryotic PixB Homolog

A putative PixB homolog was identified in the fungal parasite *Metschnikowia bicuspidata*. *M. bicuspidata* is an aquatic pathogen of crustaceans that gains access to the host via the gut and migrates to the hemolymph where it reproduces. The protein is designated a GPI mannosyltransferase 3 and contains a perfect C-terminal CD2 consensus sequence as well as an N-terminal DxxxΩxxx(S/T) CD1 sequence (conserved domains underlined). The LD is 370 residues long and is enriched (33%) with leucine (14%), phenylalanine (11%), and serine (8%). The predicted omega cleavage site is in bold.



Appendix L:

Figure S3: Molecular Phylogenetic analysis by Maximum Likelihood method of PixB homologs including the PixC and DNA directed RNA polymerase (DDRP) homologs. The evolutionary history was inferred by using the Maximum Likelihood method based on the Poisson correction model. The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+*G*, parameter = 5.1383)). The analysis involved 64 amino acid sequences. There was a total of 322 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

Appendix M:
Table S10. Phenotypic characterization of *X. nematophila 19061*, NMI1 and NMI2.

Condition	**19061**	**NMI1**	**NMI2**
Colony color on LBTA (48hrs)	Blue	Blue	Blue
Hemolysis on blood agar	Beta positive	Beta positive	Beta positive
Indole production	Negative	Negative	Negative
Nitrate production	Negative	Negative	Negative
Catalase production	Negative	Negative	Negative
Urease production	Negative	Negative	Negative
Gelatinase	Negative	Negative	Negative
Glucose fermentation	Positive	Positive	Positive
H2S production	Negative	Negative	Negative
Sucrose fermentation	Negative	Negative	Negative
Oxidase	Negative	Negative	Negative
Amylase	Negative	Negative	Negative
Lactose fermentation	Negative	Negative	Negative
DNase	Positive	Positive	Positive
Pigmentation on EMB	Dark purple	Dark purple	Dark purple