Variation in genome content and predatory phenotypes between *Bdellovibrio* sp. NC01 isolated from soil and *B. bacteriovorus* type strain HD100

Laura E. Williams*, Nicole Cullen*, Joseph A. DeGiorgis*,b, Karla J. Martinez*, Justina Mellone*, Molly Oser*

*aDepartment of Biology, Providence College, Providence, RI, USA

*bCellular Dynamics Program, Marine Biological Laboratory, Woods Hole, MA, USA

Running title: Genome and phenotype of *Bdellovibrio* from soil

*Corresponding author: Laura E. Williams, lwillia7@providence.edu
Abstract

The range of naturally occurring variation in the ability of *Bdellovibrio* strains to attack and kill Gram-negative bacteria is not well understood. Defining phenotypic and associated genotypic variation among *Bdellovibrio* will clarify how divergent lineages within this genus impact microbial communities and will inform development of predatory bacteria as biocontrol agents to combat bacterial infections. We isolated *Bdellovibrio* sp. NC01 from soil and compared its genome and predatory phenotypes to *B. bacteriovorus* type strain HD100. Based on analysis of 16S rRNA gene sequences and average amino acid identity, NC01 belongs to a different species than HD100. Genome-wide comparisons and individual gene analyses indicated that eight NC01 genome regions were likely acquired by horizontal gene transfer (HGT), further supporting an important role for HGT in *Bdellovibrio* genome evolution. Within these regions, multiple protein-coding sequences were assigned predicted functions related to transcriptional regulation and transport; however, most were annotated as hypothetical proteins. Compared to HD100, NC01 has a limited prey range and kills *E. coli* ML35 less efficiently. Whereas HD100 drastically reduces the ML35 population and then maintains low prey population density, NC01 causes a smaller reduction in ML35, after which the prey population recovers, accompanied by a decrease in NC01. In addition, NC01 forms turbid plaques on lawns of ML35, in contrast to clear plaques formed by HD100. Characterizing variation in interactions between *Bdellovibrio* and Gram-negative bacteria, such as observed with NC01 and HD100, is important for understanding the ecological significance of predatory bacteria and evaluating their effectiveness in clinical applications.
Importance

*Bdellovibrio* attack and kill Gram-negative bacteria; however, not all *Bdellovibrio* strains are equally efficient at killing the same Gram-negative bacteria. Defining how *Bdellovibrio* vary in predatory phenotypes and how this phenotypic variation relates to differences in genotype is important for understanding the ecological significance of predatory bacteria and evaluating their effectiveness in biocontrol of bacterial infections. We determined variation in genome content and predatory phenotypes, including prey range and predation efficiency, between *Bdellovibrio* sp. NC01 isolated from soil and *B. bacteriovorus* type strain HD100. NC01 is phylogenetically divergent from HD100, with eight regions of unique gene content likely acquired by horizontal gene transfer. Compared to HD100, the prey range of NC01 is limited, and it was less efficient at killing a strain of *E. coli*. These differences may have important implications for how each strain impacts microbial communities in different environments and for the effectiveness of each in clinical applications.

Introduction

Since their discovery in the early 1960s, predatory bacteria have been isolated from a wide range of terrestrial and aquatic environments, including some with extreme conditions such as high temperature or salinity (1). The impact of predatory bacteria in these environments is not fully known, but initial studies indicate they may play an important role in shaping microbial communities (2–4). In addition, their ability to attack and kill Gram-negative bacterial pathogens has generated interest in developing predatory bacteria as biocontrol agents to combat drug-resistant infections (5–7). Given their potential ecological significance and possible clinical applications, it is important to define phenotypic and genotypic variation among predatory
bacteria. In particular, variation in the outcomes of interactions between different strains of predatory bacteria and prey will determine the impact of bacterial predation in microbial communities, and variation in genome content among predatory bacteria strains may directly affect those outcomes. To broaden our understanding of variation in predatory bacteria, we isolated, assayed and sequenced a *Bdellovibrio* strain from soil and compared its predatory phenotypes and genome content to other sequenced *Bdellovibrio*.

Among the different genera of predatory bacteria, *Bdellovibrio* is the most well-studied. Currently, this genus includes two species that differ in their predation strategy. *B. bacteriovorus* invades the periplasm of Gram-negative prey cells (8), whereas *B. exovorus* attaches to the outside of prey cells but does not enter the cell itself (9). Here, we focus on intraperiplasmic *Bdellovibrio*. This lifecycle includes two main phases: the attack phase and the growth phase (10). During the attack phase, small, highly motile *Bdellovibrio* cells seek out and attach to Gram-negative prey bacteria. *Bdellovibrio* then gains access to the prey periplasm by using lytic enzymes to open a small hole in the prey outer membrane and cell wall (11). After *Bdellovibrio* moves into the prey periplasm, it uses peptidoglycan-modifying enzymes to reseal and reshape the prey cell wall, resulting in formation of a rounded structure called a bdelloplast (12). Within the bdelloplast, *Bdellovibrio* initiates the growth phase, during which it secretes lytic enzymes into the prey cytoplasm. Using digested prey cell contents as nutrients, *Bdellovibrio* grows into a single long filament within the prey periplasm, eventually dividing into multiple progeny, which lyse the bdelloplast and emerge as new attack phase cells.

By invading and lysing Gram-negative bacteria, some strains of *Bdellovibrio* can reduce prey populations by up to 8 log units over 48 hours *in vitro* (6, 13). However, not all *Bdellovibrio* strains are equally effective at killing the same Gram-negative species. Researchers have
observed variation among *Bdellovibrio* strains in which Gram-negative bacteria are susceptible to predation (14) and in how quickly and to what extent a population of Gram-negative prey is reduced (13, 15). These predatory phenotypes, referred to as prey range and predation efficiency, are clearly important in determining the outcomes of interactions between *Bdellovibrio* strains and Gram-negative bacteria; however, the extent of naturally occurring variation in these phenotypes and the underlying mechanisms governing them are not well understood. Defining how predatory phenotypes differ among *Bdellovibrio* will contribute to our understanding of the impact of predatory bacteria on microbial communities in different environments and inform their application in biocontrol of bacterial pathogens.

Here, we describe isolation of *Bdellovibrio* sp. NC01 from soil and compare its genome and predatory phenotypes to other *Bdellovibrio*, focusing primarily on the *B. bacteriovorus* type strain HD100. The intraperiplasmic predator HD100 was the first *Bdellovibrio* strain sequenced (16), and multiple studies have dissected some of the molecular mechanisms governing predation in this strain (17). To investigate variation in genome content, we compared the genome of NC01 to that of HD100 and *B. bacteriovorus* W, which is phylogenetically divergent from HD100 (18, 19). To define variation in predatory phenotypes between NC01 and HD100, we determined prey range by testing each strain against a panel of Gram-negative bacteria, and we measured predation efficiency by quantifying viable *E. coli* ML35 during co-culture with each strain.

**Results**

**Bioswale soil harbors predatory bacteria belonging to *Bdellovibrio***

Using Gram-negative soil bacteria isolated from a residential area as prey, we obtained a pure isolate of predatory bacteria from a sample of soil collected at a bioswale on the Providence College campus. The bioswale is a manmade landscape feature designed to collect and filter...
stormwater runoff from nearby buildings. We designated the bioswale soil isolate as NC01.

Using 1000x phase-contrast microscopy of lysates combining NC01 and Gram-negative prey, we observed small, fast-moving attack phase NC01 cells that attached to prey cells (Figure 1a). We also observed rounded structures characteristic of bdelloplasts, suggesting that NC01 uses a cell invasion strategy similar to *Bdellovibrio bacteriovorus*. Negative stain electron microscopy showed that NC01 attack phase cells have a comma-shaped morphology with a single polar flagellum (Figure 1b).

To classify NC01, we constructed a phylogeny of 16S rRNA gene sequences from this isolate and all *Bacteriovorax*, *Bdellovibrio* and *Halobacteriovorax* strains with complete genomes available in GenBank (Figure 2). In this phylogeny, NC01 clearly clusters with other *Bdellovibrio*. Pairwise alignment of 16S rRNA gene sequences from NC01 and *Bdellovibrio bacteriovorus* HD100, which is the type strain of the species, showed 97% similarity across 1,516 nucleotides. This value is above the 95% similarity cutoff typically used to define bacterial genera (20), which supports classifying NC01 as *Bdellovibrio*. However, it is below the 98.7% similarity cutoff typically used to define species (21), which suggests that NC01 should not be classified as *bacteriovorus*.

To explore this further, we compared the average amino acid identity (AAI) of *Bdellovibrio* strains NC01, W, HD100 and 109J (Table 1). HD100 and 109J, which are both classified as *B. bacteriovorus*, cluster in a well-supported clade in the 16S rRNA phylogeny, and their protein sequences are highly similar (average AAI 99.2%). However, NC01 protein sequences are <70% similar on average to HD100 and 109J protein sequences, supporting the conclusion that NC01 belongs to a different species than these two strains. NC01 protein sequences are even less similar on average to those of strain W, suggesting that NC01 does not
belong to the same species as this strain. Based on these analyses, we chose not to designate a
species for NC01, and we refer to this isolate as *Bdellovibrio* sp. NC01.

**Comparative genomics identifies eight regions of unique gene content possibly acquired by**

*Bdellovibrio* sp. NC01 via horizontal gene transfer

Based on the 16S rRNA gene phylogeny and average AAI comparisons, we chose *B. bacteriovorus* HD100 and *B. bacteriovorus* W to represent phylogenetic diversity and genome
variation among sequenced *Bdellovibrio* for comparative genomics. The genome of *Bdellovibrio*
sp. NC01 is larger than that of both HD100 and W (Table 2). The GC content of the NC01
genome is lower than that of HD100, but slightly higher than that of W. Using the BLAST Ring
Image Generator (BRIG), we conducted pairwise comparisons of NC01 to HD100 and W based
on nucleotide sequence alignments with BLASTN. This analysis identified eight NC01 genome
regions of at least 15,000 nucleotides in which no alignments of >200 nucleotides and e-value
<0.1 to HD100 or W were reported. These eight regions, which may represent unique NC01 gene
content, are marked on Figure 3. The plot of GC content shows that four of the eight regions
(regions 1, 4, 5, and 7) encompass substantial decreases in GC content compared to the typical
variation along the genome. Such deviation in GC content may be an indicator of horizontal gene
transfer (HGT) (22).

To assess whether the genes encoded in these eight NC01 genome regions are present but
divergent in HD100 and W, we used TBLASTN to compare NC01 amino acid sequences
annotated in each region by RAST to the HD100 and W genomes translated into all six reading
frames. RAST annotated 11 to 50 protein-coding sequences within each of the eight NC01
genome regions. By searching for these sequences along the entire HD100 and W genomes, this
analysis detects similar protein-coding sequences even if the genes have undergone
rearrangement. For the majority of the protein-coding sequences in each of the eight NC01 regions, TBLASTN did not produce any alignments to HD100 or W with >70% query coverage. For those protein-coding sequences that did align to HD100 and/or W with >70% query coverage, identity was low (<40%) for most of the sequences. Less than five protein-coding sequences in each region aligned to HD100 or W at >70% coverage and >40% identity. This suggests that the protein-coding sequences in the eight NC01 genome regions are missing from HD100 and W, instead of divergent.

In addition to comparisons against HD100 and W, we also used BLASTP to align amino acid sequences annotated in each region by RAST against the non-redundant GenBank database. For each sequence, we examined the top 250 alignments for any sequences classified as *Bdellovibrio*. This analysis expands the search space beyond complete *Bdellovibrio* genomes to any deposited sequences that were denoted by the submitters as *Bdellovibrio*. In two of the regions (regions 1 and 4), only one-third of amino acid sequences aligned to *Bdellovibrio* sequences. In every other region except region 8, 52-64% of the amino acid sequences aligned to *Bdellovibrio* sequences. In region 8, three-fourths of the amino acid sequences aligned to *Bdellovibrio* sequences. These results suggest that most of the protein-coding sequences in regions 1 and 4 may be unique to *Bdellovibrio* sp. NC01, because they do not align to other *Bdellovibrio* sequences in the database. In the other regions, the larger proportion of shared gene content between NC01 and *Bdellovibrio* in the database may be the result of HGT events that occurred before the divergence of NC01 and these strains.

The hypothesis that the eight NC01 genome regions were acquired by horizontal gene transfer is also supported by predicted functions of the protein-coding sequences encoded within the regions. RAST annotation identified functions related to mobile elements in five of the eight
regions. These include phage tail fibers (regions 1, 5 and 6), site-specific recombinases or integrases (regions 1, 4, and 7), a phage-related DNA primase (region 4), and a plasmid conjugative transfer protein (region 7). In addition, Prophinder analysis of the complete NC01 genome identified a portion of region 5 (1,632,099 – 1,632,833) as a possible prophage. Three regions (regions 2, 3 and 8) did not encode any obvious mobile element functions.

Apart from mobile element-related genes, most of the protein-coding sequences in the eight NC01 genome regions are annotated as hypothetical proteins, therefore it is unclear what functional capacities may have been gained by *Bdellovibrio* sp. NC01 via HGT and whether those functions impact predation. Some of the genes were annotated with predicted functions, which enabled us to assess gene content in the eight regions in terms of broad functional categories. Four of the eight regions (regions 1, 3, 6 and 8) encode predicted functions related to transcriptional regulation, and five regions (regions 1, 3, 5, 6, and 7) encode predicted functions related to transport. Regarding enzymatic activity, region 6 encodes a predicted cell wall endopeptidase, which is interesting given *Bdellovibrio*'s interactions with the Gram-negative prey cell wall during its life cycle.

*Bdellovibrio* sp. NC01 has a limited prey range compared to HD100 and is less efficient at killing *E. coli* ML35

To assess variation in predatory phenotypes between *Bdellovibrio* sp. NC01 and *B. bacteriovorus* HD100, we assayed prey range and predation efficiency. To determine prey range, we tested NC01 and HD100 against a panel of eight Gram-negative bacteria isolated from different environments, including a freshwater stream and soil from a residential area (Table 3). Plaque formation on a bacterial lawn indicates that the predatory isolate is able to attack and lyse that particular Gram-negative prey strain. NC01 formed plaques on lawns of five of the eight
prey strains tested, whereas HD100 formed plaques on lawns of all eight prey strains. NC01 was not able to attack and lyse two prey strains belonging to *Acinetobacter* and *Raoultella* that were isolated from a freshwater stream. In addition, although NC01 formed plaques on *E. coli* ML35, it did not attack and lyse *E. coli* 0057.

To measure predation efficiency, we quantified viable *E. coli* ML35 at different time points over 72 hours of co-culture with NC01 or HD100 (Figure 4). In the absence of predatory bacteria, we did not observe any reduction in viable *E. coli* ML35, demonstrating that co-culture conditions do not adversely affect the ML35 population. After 6 hours of co-culture, NC01 did not appreciably reduce viable *E. coli* ML35 compared to the control, whereas HD100 reduced viable ML35 by approximately 1 log unit. After 24 hours of co-culture, there was a substantial difference in the impact of NC01 and HD100 on the *E. coli* population. NC01 reduced viable ML35 by approximately 3 log units compared to the control, whereas HD100 reduced viable ML35 by 6 log units. Over the next 48 hours, HD100 continued to restrict the viable ML35 population, but the population recovered in co-cultures with NC01, increasing by 2 log units.

For two of the assay replicates, we quantified NC01 and HD100 after 72 hours of co-culture to determine if recovery of the ML35 population is associated with changes in predatory bacteria population density. Based on direct cell counts used to establish co-cultures, both NC01 and HD100 populations were \( \sim 10^8 \) cells/ml at the beginning of the assay (NC01: \( 3.8 \times 10^8 \) cells/ml and \( 7.8 \times 10^8 \) cells/ml, HD100: \( 4.1 \times 10^8 \) cells/ml and \( 9.8 \times 10^8 \) cells/ml). After 72 hours of co-culture with *E. coli* ML35, plaque counts showed a small decrease in the HD100 population (\( 1.5 \times 10^8 \) pfu/ml and \( 4.5 \times 10^8 \) pfu/ml), whereas the NC01 population decreased substantially (\( 2.3 \times 10^5 \) pfu/ml and \( 7.5 \times 10^6 \) pfu/ml). When compared in validation tests, direct cell counts and plaque counts yielded comparable results, therefore the observed decrease in
NC01 population is not due to a difference in methods for determining predator population density. These data demonstrate that for both NC01 and HD100, active predatory bacteria are present after 72 hours of co-culture with *E. coli* ML35, but, in contrast to HD100, the NC01 population density decreases by at least two orders of magnitude over the course of the predation efficiency assay.

**Plaques formed by *Bdellovibrio* sp. NC01 on *E. coli* ML35 prey lawns are not as clear as those formed by HD100**

In addition to quantitative differences in the ability of *Bdellovibrio* sp. NC01 and *B. bacteriovorus* HD100 to kill *E. coli* ML35, we also observed qualitative differences in the appearance of plaques formed by each strain on lawns of ML35 (Figure 5). In contrast to HD100, which forms completely clear, regular plaques, NC01 forms plaques of a similar size, but with incomplete clearance of the prey lawn within the plaques. Some NC01 plaques remain completely turbid over eight days of incubation, whereas others retain a turbid center but begin to clear on the edges of the plaque.

**Discussion**

We isolated *Bdellovibrio* from soil collected at a manmade landscape feature designed for stormwater management on a college campus. By comparing 16S rRNA gene sequences and analyzing average amino acid identity (AAI) among multiple sequenced *Bdellovibrio*, we determined that the *Bdellovibrio* isolate was sufficiently divergent from *B. bacteriovorus* type strain HD100 and other *B. bacteriovorus* strains that it should be classified as a different species. The only other species currently described within *Bdellovibrio* is *exovorus*; however, this species uses an epibiotic predation strategy that differs from the cell invasion strategy observed by microscopy for the soil *Bdellovibrio* isolate described here. Pending further characterization, we
designated the isolate as *Bdellovibrio* sp. NC01. Our analysis of 16S rRNA sequences and average AAI also suggests that *Bdellovibrio* strain W is divergent enough from both HD100 and NC01 to warrant its own species designation. This is consistent with the reported ability of W to form a resting stage called a “bdellocyst”, which has not been observed for any other *Bdellovibrio* so far (18, 23).

Because *Bdellovibrio* sp. NC01 is phylogenetically divergent from HD100 and other *Bdellovibrio* with sequenced genomes, this isolate provides a valuable opportunity to assess variation in genome content and predatory phenotypes among *Bdellovibrio*. Genome-wide comparisons of NC01 to HD100 and W using BLASTN identified eight regions of unique gene content along the NC01 genome, and TBLASTN confirmed that the protein-coding sequences within these regions are missing from the genomes of HD100 and W. Multiple lines of evidence support the hypothesis that these eight regions were acquired by *Bdellovibrio* sp. NC01 via horizontal gene transfer. Four of the regions have substantially lower GC content than the genome average, which may indicate horizontal acquisition from a bacterial genome with a lower average GC content than that of NC01 (22). Five of the regions encode genes with predicted functions related to mobile elements such as phage and plasmids, and one of these regions was identified as a possible prophage by phage finding sequence analysis software.

Alignment by BLASTP of amino acid sequences encoded within two of the eight regions against the GenBank database suggested that these two regions may have been acquired recently by *Bdellovibrio* sp. NC01, because most of the sequences did not align to *Bdellovibrio* sequences (complete or partial) in the database. This indicates that, given the current representation of *Bdellovibrio* in the database, gene content in these two regions is mostly unique to NC01. For the other six regions, the majority of amino acid sequences encoded in these regions aligned to
Bdellovibrio sequences in the database. This suggests that these Bdellovibrio strains may be closely related to NC01, and the HGT events corresponding to acquisition of these six regions occurred in the ancestral lineage before NC01 diverged from these strains. Addition of more complete Bdellovibrio genomes to the database will improve our ability to reconstruct evolutionary events such as horizontal gene transfer.

The analysis of Bdellovibrio sp. NC01 gene content reported here further illustrates the role of horizontal gene transfer as an important force in predatory bacteria genome evolution. Comparative genomics has yielded evidence of both ancient and recent HGT in Bdellovibrio and in the saltwater-adapted predatory genus Halobacteriovorax (24–27). However, lack of functional annotation of many Bdellovibrio protein-coding sequences prevents a thorough assessment of how HGT events impact predatory phenotypes. Within the eight NC01 genome regions likely acquired via HGT, many of the protein-coding sequences are annotated as hypothetical proteins, which provides no insight into their function and their potential roles in the predatory life cycle. To fully understand how Bdellovibrio genome variation relates to variation in predatory phenotypes, it is essential to assign functions to protein-coding genes.

In addition to genome comparisons, we also compared predatory phenotypes between Bdellovibrio sp. NC01 and B. bacteriovorus type strain HD100. Based on plaque formation on lawns of eight different Gram-negative bacteria, NC01 has a more restricted prey range than HD100. In particular, HD100 was able to attack and kill both tested strains of E. coli, whereas NC01 was only able to attack and kill E. coli ML35. Other researchers have also observed that a particular Bdellovibrio strain may vary in its ability to kill different members of the same species (14, 28); however, the mechanisms governing these outcomes are not known. By demonstrating that strain-level variation within a Gram-negative species determines susceptibility to predation,
these data have potentially important implications for developing predatory bacteria as biocontrol agents in the treatment of bacterial infections. If a *Bdellovibrio* isolate attacks and kills one representative of a Gram-negative species, it does not guarantee that the isolate will attack and kill other strains within the species. Further work to define the molecular mechanisms underlying variation in prey range may enable prediction of whether a Gram-negative strain will be susceptible to predation by a particular *Bdellovibrio* strain.

Although both NC01 and HD100 can attack and kill *E. coli* ML35, predation by HD100 causes a greater reduction in the ML35 population after 24 hours *in vitro* compared to predation by NC01. HD100 then maintains a low ML35 population density, although it does not completely eliminate viable ML35. By contrast, in co-cultures of NC01 and ML35, the ML35 population recovers after 24 hours, whereas the NC01 population decreases. This is likely not the result of general predator-prey dynamics, such as an increase in search time due to reduced ML35 numbers, because we did not observe ML35 population recovery in co-cultures with HD100. Other researchers have observed similar prey population recovery (29, 30), and Shemesh and Jurkevitch demonstrated that recovery of *Erwinia carotovora* in co-cultures with *Bdellovibrio* isolated from soil was due to development of phenotypic resistance to predation. It is possible that *E. coli* ML35 uses a similar mechanism to develop resistance to predation by *Bdellovibrio* sp. NC01, but this mechanism is either not expressed or not successful in co-cultures with *B. bacteriovorus* HD100. Further work will test this hypothesis and define the molecular mechanisms governing ML35 population recovery.

Differences in predation efficiency between NC01 and HD100 on *E. coli* ML35 may also be related to variation in plaque phenotype. Plaques formed by NC01 on lawns of ML35 are turbid, whereas plaques formed by HD100 are completely clear. Turbid plaques have been
reported as a characteristic of prey-independent mutants of *Bdellovibrio*; however, these plaques are typically smaller than those formed by the wild-type strain and sometimes have a small colony of host-independent *Bdellovibrio* at their center (31, 32). Neither of these features were observed when comparing NC01 and HD100 plaques, suggesting that the NC01 plaque phenotype is not the result of prey-independent growth. Turbid plaques formed by NC01 may arise from incomplete lysis of *E. coli* ML35 cells. If prey cells are successfully invaded but not lysed, *Bdellovibrio* progeny will be trapped, which would explain the incomplete clearance of the prey lawn and the decrease in the NC01 population observed during predation efficiency assays. Further work will test this hypothesis.

Overall, comparisons between *Bdellovibrio* sp. NC01 and *B. bacteriovorus* HD100 revealed variation in predation phenotypes that may have important implications for the ecological impact of each of these strains and their effectiveness in clinical applications. The molecular mechanisms governing phenotypes such as prey range and predation efficiency are not known. Comparative genomics can generate hypotheses linking gene content differences to observed phenotype differences, but the lack of functional annotation for many protein-coding genes complicates this effort. It is also likely that the outcomes of interactions between predatory bacteria and prey, such as whether or not two different *Bdellovibrio* can kill the same Gram-negative strain, are dependent upon mechanisms specific to the *Bdellovibrio* strain and mechanisms specific to the prey strain. Defining these mechanisms is key for understanding variation in predatory phenotypes among divergent *Bdellovibrio*.

**Materials and Methods**

**Isolation and classification of environmental bacteria for use as prey**
We isolated bacteria from an urban freshwater stream in Providence, RI (41.835, 71.44299) and from soil in a residential area of Massachusetts (42.0877, -71.2309) for use as prey. We also obtained *E. coli* ML35 from Mark Martin (University of Puget Sound, Puget Sound, WA) and *Escherichia* 0057 from Brett Pellock (Providence College, Providence, RI).

To classify these isolates, we sequenced the 16S rRNA gene following a previously described procedure (24). Briefly, we used primers 63F (33) and 1378R (34) to amplify the 16S rRNA gene. Sanger sequencing of purified PCR products was performed by GeneWiz (South Plainfield, NJ) using the same primers. We trimmed and assembled reads using Phred/Phrap/Consed (35–37), then classified the assembled 16S rRNA gene sequences using BLAST (38), the SILVA Incremental Aligner (39) and the Ribosomal Database Project classifier (40). Supplementary Table 1 shows the full classification results. The RDP classifier and the SILVA Incremental Aligner classify sequences to the genus level, but not the species level; therefore, species names are not provided for these isolates.

**Isolation and classification of *Bdellovibrio* sp. NC01 from bioswale soil**

In September 2015, we collected 50 g of soil from a bioswale on the Providence College campus (41.84277, -71.43944). We combined the soil sample with 500 ml of sterile HM buffer (25 mM HEPES adjusted to pH 7.4 and supplemented with 3 mM calcium chloride dihydrate and 2 mM magnesium chloride hexahydrate). We then filtered this mixture through sterile cheesecloth into a sterile container to remove large particulate matter. We vortexed the filtrate to release bacterial cells from the small particles of soil that passed through the cheesecloth. To pellet the soil particles, we centrifuged the filtrate at 1753 rpm for 10 minutes, after which we retained the supernatant containing suspended bacterial cells.
To enrich for predatory bacteria in the soil sample supernatant, we used *Pseudomonas* sp. NC02 as prey, because it was also isolated from a soil environment (41). We cultured *Pseudomonas* sp. NC02 overnight in 120 ml tryptic soy broth (TSB) at 28°C and 200 rpm, then centrifuged the culture at 6000 rpm for 10 minutes, washed the resulting pellet with 100 ml sterile HM buffer, centrifuged again and resuspended the resulting pellet in 2.5 ml HM buffer. We combined 1 ml of *Pseudomonas* sp. NC02 resuspension with 20 ml of soil sample supernatant and incubated this enrichment at 28°C and 200 rpm.

Once we observed small, fast-moving cells under 1000x phase-contrast microscopy, we filtered the enrichment using a 0.45 µm filter and performed 10-fold serial dilutions of the filtrate in sterile HM buffer. Dilutions were then plated using a double agar overlay method. For each dilution of the filtrate to be plated, we combined 100 µl of the dilution, 250 µl of *Pseudomonas* sp. NC02 at 10⁸ cfu/ml and 4.5 ml of molten HM top agar (0.6% agar). After vortexing, we poured this mixture onto HM agar plates (1.5% agar). We allowed the top agar to solidify at room temperature, then incubated plates at 28°C until we observed plaques in the lawn of *Pseudomonas* sp. NC02. To begin the process of obtaining a pure isolate, we used a sterile Pasteur pipette to extract single, well-defined plaques from the top agar. We combined each plaque with 500 µl of sterile HM buffer, incubated the plaque suspension at room temperature for at least 5 minutes, then vortexed it vigorously.

To culture predatory bacteria from the plaque suspensions, we made lysates combining 300 µl of each plaque suspension, 1.5 ml of an overnight culture of *Serratia* 0043 and 20 ml sterile HM buffer. *Serratia* 0043 was also isolated from a soil environment, and switching to this prey species during the isolation procedure selected for predatory isolates that were capable of attacking multiple different soil bacteria. We incubated lysates at 28°C and 200 rpm. After 48 h,
one lysate showed small, fast-moving cells, and we used this lysate to begin two additional rounds of serial dilution, plating and plaque picking to ensure a pure isolate. We used *Serratia 0043* as prey during both of these rounds. After the final round, we filtered the plaque suspension using a 0.45 µm filter, then combined 500 µl of filtrate with 500 µl of sterile 50% glycerol and stored the stock at -80°C.

To classify NC01, we extracted the 16S rRNA gene sequence from the NC01 genome based on the annotation generated by PGAP (see below for genome sequencing and annotation methods). Using SINA 1.2.11 (39), we aligned this sequence against 16S rRNA gene sequences from all nine delta-proteobacterial obligate predatory bacteria with complete genomes available in GenBank at the time (*Bacteriovorax stolpii* DSM12778 CP025704, *Bdellovibrio bacteriovorus* 109J CP007656, *Bdellovibrio bacteriovorus* HD100 NC_005363, *Bdellovibrio bacteriovorus* SSB218315 CP020946, *Bdellovibrio bacteriovorus* Tiberius NC_019567, *Bdellovibrio bacteriovorus* W NZ_CP002190, *Bdellovibrio exovorus* JSS NC_020813, *Halobacteriovorax marinus* BE01 CP017414, *Halobacteriovorax marinus* SJ NC_016620). We then inferred a phylogeny from the alignment with the online server RAxML Blackbox (42), using 100 bootstrap replicates to estimate the confidence of the topology.

To further investigate the relationships among NC01 and other *Bdellovibrio*, we used an average amino acid identity (AAI) calculator (http://enve-omics.ce.gatech.edu/) (43) to analyze the similarity among protein-coding sequences from NC01, 109J, HD100 and W, as annotated by RAST (see below for genome sequencing and annotation methods).

**Negative stain electron microscopy of *Bdellovibrio* sp. NC01**

To obtain samples of *Bdellovibrio* sp. NC01 for negative stain electron microscopy, we added a small amount of the -80°C freezer stock to HM buffer mixed with 1.5 ml of an overnight
culture of *E. coli* ML35. After 48 h of incubation at 28°C and 200 rpm, we stained samples with uranyl acetate following a previously described procedure (24) and imaged the resulting specimens using a JEOL CX 2000 transmission electron microscope.

**Genome sequencing, annotation and analysis of *Bdellovibrio* sp. NC01**

To obtain genomic DNA, we combined a small amount of the -80°C stock, 1.5 ml of an overnight culture of *E. coli* ML35 and 25 ml of sterile HM buffer. We chose ML35 as prey because the genome sequence is available for screening reads if needed (44). After 72 hours of incubation at 28°C and 200 rpm, we centrifuged the lysate at 9100 rpm for 10 minutes to pellet bacterial cells. Because PacBio sequencing technology requires at least 10 µg of genomic DNA, we did not filter the lysate to prevent loss of predatory bacteria cells. We then used the Wizard Genomic DNA purification kit to extract genomic DNA, following the manufacturer’s instructions beginning with resuspension of the pellet in 600 µl of the kit’s nuclei lysis solution. For the final rehydration step, we rehydrated the DNA pellet overnight at 4°C. Quantification by Qubit estimated a final concentration of 340 ng/µl genomic DNA.

For long-read sequencing, library construction and sequencing were performed at the Institute for Genome Sciences at the University of Maryland Baltimore on a Pacific Biosciences RSII instrument using P6-C4 chemistry. Quality control analysis of the reads showed that <10% of reads aligned by BLAST to *E. coli* or *Shigella*, indicating that representation of the prey genome in the read dataset is low. To analyze and assemble the reads, we launched an instance of SMRT Portal 2.3.0 using Amazon EC2. A single SMRT cell yielded 79,088 postfilter polymerase reads (N50 17,028) and 94,031 subreads (N50 13,572). Within SMRT Portal, we assembled the reads with HGAP3 (45), using 4.5 Mbp for the genome size. This generated 186 contigs. The largest contig was 4,001,137 bp and aligned to *Bdellovibrio* in the database using
The other 185 contigs were each <48,000 bp in length, and all of them aligned to *E. coli* by BLASTN, which is expected due to the inclusion of prey genomic DNA in the sequencing library.

To close the genome, we used BLASTN to identify overlap in the ends of the contig and trimmed the sequence to generate a closed contig of 3,975,771. We also adjusted the start site so that the first nucleotide of the closed genome sequence corresponds to the first nucleotide in the start codon of *dnaA*. To polish this sequence, we used 250 bp paired-end Illumina MiSeq data generated at the Rhode Island Genomics and Sequencing Center. We filtered and trimmed the raw read dataset to retain only those read pairs for which the quality score of every base was ≥13 and the read length was ≥65. This yielded 1,256,135 read pairs. Using bwa-mem (46), samtools (47) and Pilon (48), we aligned these reads to the closed genome sequence and corrected 1 SNP, 84 small insertions totaling 85 bases and 4 small single-base deletions. We manually examined changes flagged by Pilon using the alignment viewer Tablet (49), generating a final, polished *Bdellovibrio* sp. NC01 genome sequence of 3,975,853 bp. This sequence has been deposited in GenBank as CP030034.

In addition to annotation with NCBI’s Prokaryotic Genome Annotation Pipeline (PGAP) version 4.5, we also annotated the sequence using ClassicRAST (50, 51), Infernal (52), and tRNAScan-SE (53, 54). For genome-wide pairwise comparisons, we used the BLAST Ring Image Generator (BRIG) (55).

**Phenotype assays to test prey range and predation efficiency**

We obtained *Bdellovibrio bacteriovorus* HD100 from Mark Martin (University of Puget Sound, Puget Sound, WA) for predatory phenotype comparisons. To culture *Bdellovibrio* NC01 and HD100, we added a small amount of -80°C stocks to 20 ml HM buffer combined with 1.5 ml
of an overnight culture of *E. coli* ML35. After 72 hours of incubation at 28°C and 200 rpm, we filtered the lysate using a 0.45 µm filter to separate predatory cells from prey cells and cell debris. For each phenotype assay described below, we followed this protocol to obtain attack phase *Bdellovibrio* cells.

To test prey range and plaque appearance, we plated 10-fold serial dilutions of filtered lysate using a double agar overlay method as described above for the isolation procedure. Each tested prey strain was grown in 120 ml TSB overnight at 28°C and 200 rpm, with the exception of *Escherichia* strains, which were grown at 37°C. After centrifugation and washing, we resuspended prey cell pellets in 5 ml of sterile HM buffer. Based on cfu/ml counts, this yielded concentrations of at least $10^8$ cfu/ml, which ensured a dense prey lawn for visualizing plaques. For some prey strains, we added 500 µl of prey resuspension to top agar to improve our ability to detect and observe plaques in the prey lawn. We incubated plates at 28°C for 8 days before scoring plaque formation and performed three replicates for each prey strain. In observations of plaque appearance, we photographed plates at multiple time points over the course of 8 days.

To test predation efficiency, we adapted a protocol described in (15). We centrifuged the filtered lysate at 8500 rpm for 10 minutes. The resulting pellet of predatory bacteria was resuspended in sterile HM buffer to yield ~$10^8$ cells/ml based on direct cell counts using a Petroff-Hausser counting chamber. To ensure that direct cell counts provided accurate estimates of active predatory bacteria concentrations, we validated the method by comparing cells/ml estimated using the Petroff-Hausser counting chamber with plaque-forming units/ml estimated using serial dilution and double agar overlay to count plaques in a prey lawn. Results from two replicates showed that direct cell counts were slightly lower than pfu/ml counts, but within the
same order of magnitude (direct cell counts: $3.94 \times 10^8$ cells/ml and $2.58 \times 10^8$ cells/ml, plaque
counts: $9.2 \times 10^8$ pfu/ml and $4.7 \times 10^8$ cells/ml).

To initiate the predation efficiency assay, we combined 500 µl of *Bdellovibrio* cell
resuspension (at ~$10^8$ cells/ml) with 12.5 ml of an overnight culture of *E. coli* ML35 diluted to
OD$_{600}$ 0.15 (range: 0.135 – 0.165). This yielded ~$10^8$ cfu/ml, therefore the predation efficiency
assay results reported here are based on a predator:prey ratio of approximately 1:1. As a control,
we used 500 µl of sterile HM buffer in place of *Bdellovibrio* cell resuspension. Using cfu/ml
counts, we quantified viable *E. coli* ML35 at the beginning of the assay (time zero) and then after
6, 24, 48 and 72 hours of incubation at 28°C and 200 rpm.

**Acknowledgements**

This research was supported by an Institutional Development Award (IDeA) from the
National Institute of General Medical Sciences of the National Institutes of Health under grant no.
P20GM103430 and by funding from Providence College. This material is based upon work
conducted at a Rhode Island NSF EPSCoR research facility, the Genomics and Sequencing
Center, supported in part by the National Science Foundation EPSCoR Cooperative Agreement
#EPS-1004057. This material is based upon work supported in part by the National Science
Foundation under EPSCoR Cooperative Agreement #OIA-1655221. The funders had no role in
study design, data collection and interpretation, or the decision to submit the work for
publication.

We thank Lisa Sadzewicz and Luke Tallon at the Institute for Genome Sciences at the
University of Maryland Baltimore for PacBio sequencing services. We thank Janet Atoyan at the
Rhode Island Genomics and Sequencing Center for Illumina MiSeq sequencing services. We
thank Sean O’Donnell for isolating freshwater prey bacteria and Mark Martin and Brett Pellock for donating bacterial strains.

References


**Figures and Figure Legends**

**Figure 1.** Microscopy of *Bdellovibrio* sp. NC01 isolated from bioswale soil. (A) Compilation of four images taken using 1000x phase-contrast microscopy to show small, comma-shaped NC01 attack phase cells attached to larger *E. coli* ML35 cells after 48 hours of co-culture. (B) Negative stain electron microscopy of a NC01 attack phase cell. The scale bar is 500 nm.
Figure 2. 16S rRNA gene phylogeny. We aligned 16S rRNA gene sequences using SINA 1.2.11, then constructed a phylogeny using RAxML BlackBox with 100 bootstrap replicates. The phylogeny is rooted on the branch joining *Bdellovibrio* with *Bacteriovorax/Halobacteriovorax*, and bootstrap support for each node is shown.
Figure 3. Genome-wide pairwise comparisons of *Bdellovibrio* nucleotide sequences. The genome of *Bdellovibrio* sp. NC01 is shown as the innermost ring. The next ring is a plot of GC content and the purple and green ring is a plot of GC skew, both of which were generated by the BLAST Ring Image Generator (BRIG) using sliding windows to determine deviation from the average value for the genome. The blue ring and red ring show pairwise nucleotide sequence alignments of NC01 to HD100 and W, respectively. Each individual alignment is shown as a block along the ring, and color gradation of the blocks reflects the percent similarity of
alignments. The outermost ring denotes eight regions lacking alignments, which may indicate unique NC01 gene content.
**Figure 4.** Efficiency of *Bdellovibrio* sp. NC01 and *B. bacteriovorus* HD100 predation on *E. coli* ML35. We quantified viable *E. coli* ML35 over 72 hours on its own (control), combined with *Bdellovibrio* sp. NC01 (NC01) and combined with *B. bacteriovorus* HD100 (HD100). The resulting cfu/ml data were log transformed and plotted using ggplot in RStudio. Each time point shows the mean and standard deviation for at least five replicates.
Figure 5. Plaque phenotypes of *Bdellovibrio* sp. NC01 and *B. bacteriovorus* HD100. Using a double agar overlay method, we assessed plaque formation on lawns of *E. coli* ML35. These images were taken after 6 days of incubation and illustrate reproducible differences in the appearance of plaques formed by HD100 (A) and NC01 (B).
### Tables

#### Table 1. Average amino acid identity among *Bdellovibrio* genomes

<table>
<thead>
<tr>
<th></th>
<th>NC01</th>
<th>W</th>
<th>HD100</th>
<th>109J</th>
</tr>
</thead>
<tbody>
<tr>
<td>NC01</td>
<td>64.4</td>
<td>68.8</td>
<td>68.8</td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>64.4</td>
<td>66.4</td>
<td>66.5</td>
<td></td>
</tr>
<tr>
<td>HD100</td>
<td>68.8</td>
<td>66.4</td>
<td>99.2</td>
<td></td>
</tr>
<tr>
<td>109J</td>
<td>68.8</td>
<td>66.5</td>
<td>99.2</td>
<td></td>
</tr>
</tbody>
</table>

#### Table 2. Genome comparison of *Bdellovibrio* sp. NC01, *Bdellovibrio bacteriovorus* HD100 and *Bdellovibrio bacteriovorus* W

<table>
<thead>
<tr>
<th></th>
<th><em>Bdellovibrio</em> sp. NC01</th>
<th><em>Bdellovibrio bacteriovorus</em> HD100</th>
<th><em>Bdellovibrio bacteriovorus</em> W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genome size (bp)</td>
<td>3,975,853</td>
<td>3,782,950</td>
<td>3,006,992</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>44.5</td>
<td>50.6</td>
<td>43.3</td>
</tr>
<tr>
<td>Genes</td>
<td>3,840</td>
<td>3,828</td>
<td>3,597</td>
</tr>
<tr>
<td>Protein-coding</td>
<td>3,773</td>
<td>3,776</td>
<td>3,548</td>
</tr>
<tr>
<td>Hypothetical proteins</td>
<td>1,712</td>
<td>1,431</td>
<td>1,245</td>
</tr>
<tr>
<td>tRNA</td>
<td>34</td>
<td>34</td>
<td>36</td>
</tr>
<tr>
<td>rRNA</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Other RNA</td>
<td>4</td>
<td>12</td>
<td>4</td>
</tr>
</tbody>
</table>

#### Table 3. Prey range of *B. bacteriovorus* HD100 and *Bdellovibrio* sp. NC01

<table>
<thead>
<tr>
<th>Genus</th>
<th>Strain ID</th>
<th>Environment</th>
<th>Plaque formation(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Acinetobacter</em></td>
<td>0036</td>
<td>Freshwater</td>
<td>Yes</td>
</tr>
<tr>
<td><em>Aeromonas</em></td>
<td>0031</td>
<td>Freshwater</td>
<td>Yes</td>
</tr>
<tr>
<td>Enterobacteriaceae</td>
<td>0032</td>
<td>Freshwater</td>
<td>Yes</td>
</tr>
<tr>
<td><em>Raoultella</em></td>
<td>0037</td>
<td>Freshwater</td>
<td>Yes</td>
</tr>
<tr>
<td><em>Pseudomonas</em></td>
<td>NC02 (0042)(^b)</td>
<td>Soil</td>
<td>Yes</td>
</tr>
<tr>
<td><em>Serratia</em></td>
<td>0043</td>
<td>Soil</td>
<td>Yes</td>
</tr>
<tr>
<td><em>Escherichia</em></td>
<td>0057</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td><em>Escherichia</em></td>
<td>ML35</td>
<td></td>
<td>Yes</td>
</tr>
</tbody>
</table>

\(^a\)Plaque formation data based on three replicates

\(^b\)CC-BY 4.0 International license under a not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available
\textit{Pseudomonas} sp. NC02 is referred to as \textit{0042} in (27)