ISOLATION, MOLECULAR & PHYSIOLOGICAL CHARACTERIZATION OF SULFATE-REDUCING, HETEROTROPHIC, DIAZOTROPHS

Annaliese Katrin Jones

University of Rhode Island, akjones2188@gmail.com

Follow this and additional works at: https://digitalcommons.uri.edu/theses

Recommended Citation

Jones, Annaliese Katrin, "ISOLATION, MOLECULAR & PHYSIOLOGICAL CHARACTERIZATION OF SULFATE-REDUCING, HETEROTROPHIC, DIAZOTROPHS" (2015). Open Access Master's Theses. Paper 501. https://digitalcommons.uri.edu/theses/501
ISOLATION, MOLECULAR & PHYSIOLOGICAL
CHARACTERIZATION OF SULFATE-REDUCING,
HETEROTROPHIC, DIAZOTROPHS

BY
ANALIESE KATRIN JONES

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER’S OF SCIENCE
IN
INTEGRATIVE AND EVOLUTIONARY BIOLOGY

UNIVERSITY OF RHODE ISLAND
2015
MASTER OF SCIENCE
OF
ANNALIESE KATRIN JONES

APPROVED:

Thesis Committee:

Major Professor Bethany D. Jenkins

Serena Moseman-Valtierra

Daniel Udwary

Nasser H. Zawia
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND
2015
ABSTRACT

Nitrogen (N?2) fixation is the process by which N?2 gas is converted to biologically reactive ammonia, and is a cellular capability widely distributed amongst prokaryotes. This process is essential for the input of new, reactive N in a variety of environments. Heterotrophic bacterial N fixers residing in estuarine sediments have only recently been acknowledged as important contributors to the overall N budget of these ecosystems and many specifics about their role in estuarine N cycling remain unknown, partly due to a lack of knowledge about their autecology and a lack of cultivated representatives. Nitrogenase reductase (nifH) gene composition and prevalence in Narragansett Bay sediments has revealed that two distinct phylogenetic groups dominate N-fixation. Analysis of nifH transcripts has revealed one active group to be the Desulfovibrionaceae, belonging to the Deltaproteobacteria. We see nifH expression from this group across sampling sites and times, despite the fact that Narragansett Bay sediments are replete with combined N, which is thought to inhibit N fixation in the environment. Here were present the genomic and physiological data in relation to N2 fixation by two heterotrophic members of the Desulfovibrionaceae, isolated from sediments of the Narragansett Bay estuary in 2010 and 2011, respectively: Desulfovibrio sp. NAR1 and Desulfovibrio sp. NAR2. To elucidate how nitrogenase activity in these organisms responds to the presence of different sources of combined N, and to link observed physiology with genomic potential (i.e. gene content), we performed a two-part study that coupled high-throughput genome sequencing and analysis with physiological investigations of growth on different N sources with N fixation rate measurements. The genomes of the two diazotrophic
Narragansett Bay Desulfovibrio isolates (NAR1 and NAR2) were sequenced using a high throughput platform, subsequently assembled, annotated, and investigated for genes related to N fixation and overall N metabolism which were then compared across 34 additional Desulfovibrio genomes which were publicly available. To link findings at the molecular level with observations at the physiological level, N fixation rates were measured using the acetylene reduction assay (ARA) under conditions free of reactive N, and under the following combined N conditions: 12 mM urea, 12 mM NO₃⁻, and 12 mM NH₃. Both isolates can sustain growth by N₂ fixation in the absence of biologically available N and our data indicate that nitrogenase activity is completely inhibited by the presence of ammonia, yet uninhibited by nitrate and urea, which are other forms of combined N found in Narragansett Bay. This agrees with observations made at the genome level, as neither our isolates nor the other Desulfovibrio examined in comparison appeared to have the genetic capability to use NO₃⁻ or urea catabolically to meet cellular N demands. This study indicates that the Desulfovibrionaceae are restricted in terms of the N sources they are capable of using, and that this may be a factor contributing to the observed N fixation by this group in sediments that are not limited for sources of combined N. Genome sequencing also reveals both isolates to be metabolically versatile and unique. The NAR1 isolate possesses genes involved in bacterial mercury methylation, and displays near obligate biofilm formation. Genes were also found in the NAR1 isolate which suggest the involvement of c-di-GMP in cell-to-cell communication and biofilm formation. This is particularly interesting since biofilm formation and quorum sensing is not well characterized among the Desulfovibrio, despite biofilm formation being displayed by many members of this
genus. While investigating the role of these organisms as important contributors of fixed N in Narragansett Bay, it was critical that we examine these additional aspects of their metabolism in order to gain a better understanding of controls on growth that may also impact biomass and the ability of these organisms to achieve significant rates of N fixation in the environment.
ACKNOWLEDGMENTS

First and foremost, I would like to express my gratitude to my advisor, Bethany Jenkins, for being an amazing mentor. Thank you for all of your support and words of encouragement, they have kept me motivated even during the most difficult of times. I will always appreciate your advice and expertise. I greatly appreciate all of the opportunities to travel to meetings and collaborate with a variety of scientists over the years. You have truly shaped me as a scientist, speaker, and individual.

Thank you to my committee members, Serena Moseman-Valtierra, Dan Udwary, and Rebecca Robinson for your assistance and time towards this thesis. I am grateful for your support, expertise, and guidance. I would also like to thank the members of MetaGroup. My presentations have benefited greatly from their suggestions and positive critiques.

I would like to thank the many wonderful individuals that I have had the privilege of working with over the years, including past and present members of the Jenkins and Lane labs. I would like to give a special thank-you to Dreux, Joselynn, and Eric for all of the computational help and support. Thank-you to “Team Mud”: Shelley, Andraya and Rodrigue. I would like to give a special thanks to Rodrigue. I will always appreciate how you mentored me as an undergraduate and graduate student, and taught me (almost) everything I know about isolating and cultivating environmental anaerobes. Thank you for all of your support, guidance, and friendship. I could not have asked for a better mentor.

Thank you to my friends and family, for their support and understanding as I finished my thesis. Thank you to Ashley for your kindness and words of
encouragement. Finally, thank you to my parents, Ingrid and Kevin, for your love, encouragement and support throughout my schooling. You have both always believed in me, and I could never have done this without you.
PREFACE

This thesis was formatted in accordance with the manuscript format guidelines established by the Graduate School of the University of Rhode Island.
TABLE OF CONTENTS

ABSTRACT ........................................................................................................................................... ii

ACKNOWLEDGEMENTS ................................................................................................................ v

PREFACE ............................................................................................................................................... vii

TABLE OF CONTENTS ...................................................................................................................... viii

LIST OF TABLES ................................................................................................................................. x

LIST OF FIGURES .............................................................................................................................. xi

INTRODUCTION ................................................................................................................................. 1

The Role of Heterotrophic Diazotrophs in Estuarine Sediments ......................... 1

Thesis Motivation and Outline ................................................................................................. 5

References ........................................................................................................................................... 7

Manuscript 1: Genome comparison and molecular characterization of two novel
diazotrophs from Narragansett Bay, RI ...................................................................................... 11

Abstract ........................................................................................................................................... 12

Introduction ........................................................................................................................................ 13

Materials and Methods ........................................................................................................... 17

Growth and Glycerol Stock Preparation of Desulfovibrio sp. NAR1 ................. 18

Acetylene Reduction Assay ............................................................................................................. 19

Genome sequencing, assembly, annotation, and gene comparison .............. 21

Phylogenetic Analysis and Evolutionary Gene Networks ............................ 24

Whole Genome Alignments ................................................................................................. 25

Results and Discussion ........................................................................................................... 26
Growth and Acetylene Reduction of Desulfovibrio sp. NAR1 under N limited conditions ................................................................. 26
Growth and Acetylene Reduction of Desulfovibrio sp. NAR1 under differing combined N conditions ........................................... 29
Genome sequencing outputs, assembly and annotation .................... 30
Comparison of isolate genomes using IMG/ER and RAST ............... 32
Whole genome comparison and alignments .................................... 36
Gene network analysis ..................................................................... 38
Nitrogen fixation ............................................................................. 41
Urea metabolism ............................................................................. 42
Ammonia metabolism ..................................................................... 45
Nitrate metabolism ......................................................................... 48
Carbon metabolism ......................................................................... 49
Biofilm formation ........................................................................... 49
Mercury methylation ....................................................................... 51
Conclusions ..................................................................................... 52
References ....................................................................................... 54

CONCLUSION ...................................................................................... 89

References ....................................................................................... 92
Supplemental Protocol 1 ................................................................. 80
Supplemental Protocol 2 ................................................................. 83
Appendix A. List of proteins contained in D. sp. NAR1 filtered network ....... 96
| TABLE                      | PAGE |
|----------------------------|------|
| Table 1. Assembly statistics | 75   |
| Table 2. Comparative genome statistics | 76   |
| Table 3. Comparison of gene presence involved in N metabolism | 77   |
| Supplemental Table 1. Contig statistics NAR1 | 85   |
| Supplemental Table 2. Contig statistics NAR2 | 86   |
| Supplemental Table 3. *Desulfovibrio* genomes used in comparison | 87   |
| Supplemental Table 4. Genes used to construct multigene phylogeny | 88   |
## LIST OF FIGURES

| FIGURE | PAGE |
|--------|------|
| Figure 1. Multigene Bayesian phylogenetic tree | 66 |
| Figure 2. Whole genome comparisons | 67 |
| Figure 3. Evolutionary gene networks | 68 |
| Figure 4. N-fixation gene clusters | 69 |
| Figure 5. Proteins required for Hg methylation | 70 |
| Figure 6. Growth of NAR1 under N limitation | 71 |
| Figure 7. Nitrogenase activity of NAR1 under N limitation | 72 |
| Figure 8. Cell enumeration of NAR1 grown under different N treatments | 73 |
| Figure 9. Nitrogenase activity of NAR1 under different combined N conditions | 74 |
| Supplemental Figure 1. Calculations used for biomass estimates | 83 |
INTRODUCTION

The Role of Heterotrophic Diazotrophs in Estuarine Sediments

Estuarine sediments, such as those in Narragansett Bay, have historically been shown to be major regions of nitrogen (N) removal via the activity of denitrifying bacteria (Nixon et al. 1996). Denitrification results in the release of nitrogen (N$_2$) gas and the loss of biologically reactive N. The opposing process of N fixation converts N$_2$ gas to biologically reactive ammonia, and has been demonstrated to occur in Narragansett Bay sediments by anaerobic diazotrophs (Fulweiler et al. 2013; Brown & Jenkins 2014). Historically, N fixation activities in marine systems have primarily been attributed to cyanobacterial species residing in euphotic regions of the ocean (Karl et al. 2002) and contributions to this process from anaerobic bacteria were not previously documented. Additionally, due to the energetically costly nature of biological N fixation (Stam et al. 1987) and the typically abundant concentrations of combined N in estuarine systems including Narragansett Bay (Rabalais 2002; Pryor et al. 2007; DiMilia et al. 2011), it was previously thought that any heterotrophic N fixation occurring in the sediments was likely an inconsequential process (Howarth et al. 1988). However, a growing body of evidence involving direct observations of N$_2$ flux across the sediment water interface in a variety of marine, salt marsh, and sea grass systems (Fulweiler et al. 2007; Gardner et al. 2006; Ferguson et al. 2007; Welsh et al. 1996a; McGlathery et al. 2004) have revealed that heterotrophic N-fixation is an important source of biologically reactive N in these systems and has begun to alter the previous conclusions regarding the role of heterotrophic N fixing bacteria.
In Narragansett Bay the most active of these heterotrophic diazotrophs, based on the number of \textit{nifH} (dinitrogen reductase) transcripts, belong to two distinct phylogenetic groups; members of the \textit{Desulfovibrionaceae} and the \textit{Geobacteraceae} (Fulweiler et al. 2013). In order to gain a better understanding of the ecology of these active N fixers, we attempted to isolate representatives of both phylogenetic groups from Narragansett Bay sediments. We were initially successful in isolating two organisms belonging to the \textit{Desulfovibrionaceae} (Spinette and Jenkins, unpublished data), which are the topic of this thesis. Members of the \textit{Desulfovibrionaceae} are heterotrophic sulfate (SO\textsubscript{4}\textsuperscript{2-}) reducers and have been shown to fix N in culture (Sisler \& ZoBell, 1951) as well as under a variety of environmental conditions (Welsh et al. 1996a; Bertics et al. 2012; Brown \& Jenkins 2014; Fulweiler et al. 2013; Burns et al. 2002). Members of this group hold particular interest within the scientific community for their ability to perform a wide variety of metabolic functions, including; the reduction of sulfate and sulfur to sulfide species, the ability to use recalcitrant carbon sources, the ability to fix N, and the ability to transform certain heavy metal species (Geets et al. 2006; Annachhatre \& Suktrakoolvait, 2001; Sisler \& ZoBell, 1951; Parks et al. 2013). Additionally, the \textit{Desulfovibrio} impact a variety of industries and applications. In the oil industry they are considered to be a nuisance, contributing to the souring of oil via sulfide production and causing the corrosion of pipelines and wells (Muyzer \& Stams, 2008). In wastewater treatment plants, sulfate reducers are frequently employed as a means of removing sulfates and converting hydrogen sulfide species into precipitated heavy metals (Hulshoff et al. 1998). Similarly, the \textit{Desulfovibrio} play an important role in bioremediation by reducing and immobilizing
heavy metals (Martins et al. 2009). Lastly, these organisms hold potential for use as an alternate energy source in biological fuel cells (Gutierrez-Sanchez et al. 2011).

In addition to the properties described above, these organisms are of particular interest regarding their response in environments that experience seasonal hypoxic periods (Ehrlich, 2014). Respiration rates of aerobic bacteria in the water column are high during hypoxia, which depletes the water column of overall O₂ and reduces the amount of O₂ that can penetrate the sediments. Oxygen depleted sediments are an expanded niche for anaerobic bacteria such as sulfate-reducing heterotrophic N-fixing microbes, such as the Desulfovibrio. At the same time, hypoxic conditions prevent the primary process of N removal in these sediments, coupled nitrification-denitrification, from occurring. Sulfide, which results from the respiration of sulfate reducers, directly inhibits nitrification (Joye & Hollibaugh 1995). The inhibition of nitrification means that the subsequent process of denitrification cannot take place, and this hinders removal of reactive N from the sediments. Additional N removal reactions such as anaerobic ammonium oxidation are also inhibited (Thamdrup & Dalsgaard, 2002).

The increased N-fixation activity by Desulfovibrio spp. and impairment of N removal processes during hypoxic conditions could lead to a greater N inventory, increasing eutrophication and exacerbating hypoxic conditions in estuarine systems like Narragansett Bay (Ehrlich, 2014).

However, despite the multiple applications for the Desulfovibrio and previous studies documenting their ecological importance and ability to fix N, there is little previous research investigating their N fixation capabilities and overall N metabolism in a physiological or genomic context. This is likely due to the fact that members of
this genus, and other anaerobic heterotrophic diazotrophs, have only recently begun to be appreciated as significant contributors to N fixation in estuarine systems. As a result investigations have not yet moved beyond those regarding their ecological significance and roles in the environment, which have primarily focused on group abundance and dynamics in situ, to examinations of Desulfovibrio genome content at the individual level in the context of N fixation capabilities and overall N metabolism. In other words, shifting from questions regarding “who, what, and where” to questions regarding “why and how” these organisms fix N in their respective habitats.

Anaerobic heterotrophic diazotrophs, specifically members of the Desulfovibrionaceae, present a unique challenge because currently there is no model or set of supporting physiological data that exist which explain why N-fixation activity by these organisms is observed in environments which are not limited for combined N. The paradigm for aerobic N fixing bacteria is that combined N represses N fixation (Mackerras & Smith 1986; Howarth et al. 1988), and this is something that still remains to be investigated in anaerobic diazotrophs. Similarly, there has been very little investigation at the genomic level into the N metabolism of these organisms, their potential ability as a group to assimilate different combined sources of N, or a specific description of controls on their overall N fixation. We hypothesize that members of this genus may not be capable of using certain forms of N in an assimilatory manner and thus would need to continue to fix N under otherwise N replete conditions. This may be an important contributing factor to what has been seen regarding their activity in the environment (Fulweiler et al 2013; Brown & Jenkins 2014; Ehrlich 2014). Because of their N metabolism and with hypoxia expected to
increase due to anthropogenic activities and climate change (Diaz & Rosenberg, 2008; Middleburg and Levin, 2009), we could see an expanded niche and potentially stimulated N fixation by anaerobic heterotrophs.

**Thesis Motivation and Outline**

Members of the *Desulfovibrio* were found to be one of the primary bacterial groups responsible for driving N-fixation activity in Narragansett Bay (Brown, 2013; Fulweiler et al. 2013). Previously, sulfate-reducing bacteria (SRB) as a group have been found to fix N in culture (Sisler & ZoBell, 1951) and drive N fixation in a variety of sediments that were not found to be limited for sources of combined N, such as those colonized by *Zostera* spp. (McGlathery et al. 1998; Welsh et al. 1996b), hypoxic sediments from Eckernförde Bay, Germany (Bertics et al. 2012), and anoxic sediments from Chesapeake Bay, USA (Burns et al. 2002).

Heterotrophic N fixers, such as members of the *Desulfovibrio* spp., have only recently been appreciated as important contributors to the overall N budget of estuarine ecosystems. Presently, the majority of these environmentally active microbes have not been isolated in culture and lack complete genome sequence information. Although there are cultured representatives of the *Desulfovibrionaceae*, there remain relatively few members of the *Desulfovibrio* genus with completely sequenced and annotated genomes; there are currently 2 deep sea species, 1 groundwater species, 1 terrestrial species, a handful of estuarine species, and a similarly few extremophilic representatives with completed genomes (Ji et al. 2013; Pedersen et al., 2014; Hauser et al. 2011; Brown et al. 2011; Brown et al. 2011; Nakazawa et al. 2009; Heidelberg et al. 2004; Pradel et al. 2013). Of those representatives with completed genomes there is
still much yet to be discerned regarding their metabolism, nutrient cycling properties, and responses or adaptations to different environmental stressors, particularly from a genomic perspective and more specifically as those factors pertain to N fixation.

There is a similar lack of information regarding N fixation rate data among the *Desulfovibrio*. Although there are previous studies examining N fixation rates for heterotrophic diazotrophs in sediments using the acetylene reduction assay (McGlathery *et al.* 1998; Welsh *et al.* 1996b), these studies have used environmental samples and thus are examining the overall N fixation ability of a mixed microbial consortia and are not able to establish cell-specific rates of N₂ fixation, or a connection between those rates and the responsible cells. This lack of evidence is most likely due to the fact many of these organisms remain uncultured, and where there are cultivated representatives the potential contributions to heterotrophic N fixation by those specific organisms remain largely disregarded.

The lack of cultivated *Desulfovibrio* representatives and corresponding analysis of their genomic capabilities in regards to N fixation and metabolism, combined with a lack of organism specific physiological data regarding N fixation rates and rate responses to environmental N conditions, have provided motivation for the current study. The genomes of two *Desulfovibrio* spp. isolates, *Desulfovibrio* sp. NAR1 and *Desulfovibrio* sp. NAR2 (Spinette & Jenkins, unpublished data), were sequenced, annotated, and analyzed, both independently and comparatively, in reference to other members of this genus, with particular focus on genomic potential involving N fixation and overall N metabolism. Physiological experiments targeting N fixation rates and rate responses to different combined N sources were also performed for the
Desulfovibrio isolate NAR1 using the acetylene reduction assay (Capone, 1993). This study seeks to begin closing the gap between observed environmental activities and a comprehensive genomic view of environmentally relevant members of the Desulfovibrionaceae.

References:

Annachhatre, A. P., & Suktrakoolvait, S. (2001). Biological sulfate reduction using molasses as a carbon source. Water Environment Research: A Research Publication of the Water Environment Federation, 73(1), 118–126.

Bertics, V. J., Löschler, C. R., Salonen, I., Dale, A. W., Schmitz, R. A., Treude, T. (2012). Occurrence of benthic microbial nitrogen fixation coupled to sulfate reduction in the seasonally hypoxic Eckernförde Bay, Baltic Sea. Biogeosciences Discussions, 9(6), 6489–6533.

Brown, S. D., Gilmour, C. C., Kucken, A. M., Wall, J. D., Elias, D. a, Brandt, C. C., … Palumbo, A. V. (2011). Genome sequence of the mercury-methylating strain Desulfovibrio desulfuricans ND132. Journal of Bacteriology, 193(8), 2078–2079.

Brown, S. D., Wall, J. D., Kucken, A. M., Gilmour, C. C., Podar, M., Brandt, C. C., … Elias, D. a. (2011). Genome sequence of the mercury-methylating and pleomorphic Desulfovibrio africanus Strain Walvis Bay. Journal of Bacteriology, 193(15), 4037–4038.

Brown, S. M. (2013). Using molecular tools to elucidate controls on microbes driving the nitrogen cycle in marine sediments. PhD dissertation, University of Rhode Island.

Brown, S. M., & Jenkins, B. D. (2014). Profiling gene expression to distinguish the likely active diazotrophs from a sea of genetic potential in marine sediments. Environmental Microbiology. e-pub ahead of print 25 Mar 2014, doi: 10.1111/1462-2920.12403.

Burns, J. A., Zehr, J. P., & Capone, D. G. (2002). Nitrogen-fixing phylotypes of Chesapeake Bay and Neuse River estuary sediments. Microbial Ecology, 44(4), 336–343.

Capone, D. (1993). Determination of Nitrogenase Activity in Aquatic Samples Using the Acetylene Reduction Procedure. In P. Kemp, J. Cole, B. Sherr, & E. Sherr (Eds.), Handbook of methods in aquatic microbial ecology. Lewis Publ: Boca Raton, FL.
Diaz, R. J., & Rosenberg, R. (2008). Spreading dead zones and consequences for marine ecosystems. *Science (New York, N.Y.),* 321(5891), 926–929.

DiMilia, P., Nixon, S., Oczkowski, A., Altabet, M., McKinney, R. (2011). Some challenges of an "upside-down" nitrogen budget-- science and management in Greenwich Bay, RI (USA). *Marine Pollution Bulletin,* 62(4), 672-680.

Ehrlich, A. (2014). *Elucidating the Molecular Response of Microbial Nitrogen Fixation in Estuarine Sediments to Hypoxia.* M.S. thesis, University of Rhode Island.

Ferguson, A., Eyre, B., & Gay, J. (2007). Benthic metabolism and nitrogen cycling in a sub-tropical coastal embayment: spatial and seasonal variation and controlling factors. *Aquatic Microbial Ecology,* 48, 175–195.

Fulweiler, R. W., Brown, S. M., Nixon, S. W., & Jenkins, B. D. (2013). Evidence and a conceptual model for the co-occurrence of nitrogen fixation and denitrification in heterotrophic marine sediments. *Marine Ecology Progress Series,* 482, 57–68.

Fulweiler, R. W., Nixon, S. W., Buckley, B. A., & Granger, S. L. (2007). Reversal of the net dinitrogen gas flux in coastal marine sediments. *Nature,* 448(7150), 180–182.

Gardner, W., McCarthy, M., & An, S. (2006). Nitrogen fixation and dissimilatory nitrate reduction to ammonium (DNRA) support nitrogen dynamics in Texas estuaries. *Limnology and Oceanography,* 51(1), 558–568.

Geets, J., Borremans, B., Diels, L., Springael, D., Vangronsveld, J., van der Lelie, D., & Vanbroekhoven, K. (2006). DsrB gene-based DGGE for community and diversity surveys of sulfate-reducing bacteria. *Journal of Microbiological Methods,* 66(2), 194–205.

Gutiérrez-Sánchez, C., Olea, D., Marques, M., Fernández, V. M., Pereira, I. A. C., Vélez, M., & De Lacey, A. L. (2011). Oriented Immobilization of a Membrane-Bound Hydrogenase onto an Electrode for Direct Electron Transfer. *Langmuir,* 27(10), 6449–6457.

Heidelberg, J. F., Seshadri, R., Haveman, S. a, Hemme, C. L., Paulsen, I. T., Kolonay, J. F., … Fraser, C. M. (2004). The genome sequence of the anaerobic, sulfate-reducing bacterium Desulfovibrio vulgaris Hildenborough. *Nature Biotechnology,* 22(5), 554–559.

Howarth, R. W., Marino, R., & Lane, J. (1988). Nitrogen fixation in freshwater, estuarine, and marine ecosystems. 1. Rates and importance. *Limnology and Oceanography,* 33(4), 669–687.
Howarth, R. W., Marino, R., & Cole, J. (1988). Nitrogen fixation in freshwater, estuarine, and marine ecosystems. 2. Biogeochemical controls. *Limnology and Oceanography, 33*(4), 688-701.

Hulshoff Pol, L., Lens, P. L., Stams, A. M., & Lettinga, G. (1998). Anaerobic treatment of sulphate-rich wastewaters. *Biodegradation, 9*(3-4), 213–224.

Ji, B., Gimenez, G., Barbe, V., Vacherie, B., Rouy, Z., Amrani, A., … Pradel, N. (2013). Complete Genome Sequence of the Piezophilic, Mesophilic, Sulfate-Reducing Bacterium Desulfovibrio hydrothermalis AM13(T.). *Genome Announcements, 1*(1), 2–3.

Joye, S., Hollibaugh, J. (1995). Influence of Sulfide Inhibition of Nitrification on Nitrogen Regeneration in Sediments. *Science, 270*(5236), 623-625.

Karl, D., Michaels, A., Bergman, B., Capone, D., Carpenter, E., Letelier, R., … Stal, L. (2002). Dinitrogen fixation in the world's oceans. *Biogeochemistry, 57*(58), 47-98.

Mackerras, A., Smith, G. (1986). Evidence for direct repression of nitrogenase by ammonia in the cyanobacterium *Anabaena cylindrica*. *Biochemical and Biophysical Research Communications, 134*(2), 835-844.

Martins, M., Faleiro, M. L., Barros, R. J., Veríssimo, A. R., Barreiros, M. A., & Costa, M. C. (2009). Characterization and activity studies of highly heavy metal resistant sulphate-reducing bacteria to be used in acid mine drainage decontamination. *Journal of Hazardous Materials, 166*(2-3), 706–713.

McGlathery, K. (1998). Temporal and spatial variation in nitrogen fixation activity in the eelgrass *Zostera marina* rhizosphere. *Marine Ecology Progress Series, 168*, 245–258.

McGlathery, K., Sundbäck, K., & Anderson, I. (2004). The Importance of Primary Producers for Benthic Nitrogen and Phosphorus Cycling. In S. L. Nielsen, G. M. Banta, & M. F. Pedersen (Eds.), *Estuarine Nutrient Cycling: The Influence of Primary Producers* (pp. 231–261). Kluwer Academic Publishers.

Middelburg, J. J., & Levin, L. (2009). Coastal hypoxia and sediment biogeochemistry. *Biogeosciences Discussions, 6*, 3655–3706.

Muyzer, G., & Stams, A. J. M. (2008). The ecology and biotechnology of sulphate-reducing bacteria. *Nature Reviews. Microbiology, 6*(6), 441–454.

Nixon, S. W., Ammerman, J. W., Atkinson, L. P., Berounsky, V. M., Billen, G., Boicourt, W. C., … Seitzinger, S. P. (1996). The fate of nitrogen and phosphorus
at the land-sea margin of the North Atlantic Ocean. *Biogeochemistry, 35*(1), 141–180.

Parks, J. M., Johs, A., Podar, M., Bridou, R., Hurt, R. a, Smith, S. D., … Liang, L. (2013). The genetic basis for bacterial mercury methylation. *Science (New York, N.Y.), 339*(6125), 1332–1335.

Pedersen, K., Bengtsson, A., Edlund, J., Rabe, L., Hazen, T., Chakraborty, R., … Shapiro, N. (2014). Complete Genome Sequence of the Subsurface, Mesophilic Sulfate-Reducing Bacterium Desulfovibrio aespoeensis Aspo-2. *Genome Announcements, 2*(3), 3–4.

Pradel, N., Ji, B., Gimenez, G., Talla, E., Lenoble, P., Garel, M., … Dolla, A. (2013). The first genomic and proteomic characterization of a deep-sea sulfate reducer: insights into the piezophilic lifestyle of Desulfovibrio piezophilus. *PloS One, 8*(1), e55130. doi:10.1371/journal.pone.0055130

Pryor, D., Saarman, E., Murray, D., Prell, W. (2007). Nitrogen Loading from Wastewater Treatment Plants to Upper Narragansett Bay. *Narragansett Bay Estuary Program Report NBEP-2007-126*. Brown University, Providence RI. (http://www.nbep.org/publications/NBEP-126-FNL-Pryor.pdf)

Rabalais, N. (2002). Nitrogen in Aquatic Ecosystems. *AMBIO: A Journal of the Human Environment, 31*(2), 102-112.

Sisler, F. D., & ZoBell, C. E. (1951). Nitrogen Fixation by Sulfate-reducing Bacteria Indicated by Nitrogen/Argon Ratios. *Science, 113*(2940), 511–512.

Stam, H., Stouthamer, A., van Verseveld, H. (1987). Hydrogen metabolism and energy costs of nitrogen fixation. *FEMS Microbiology Letters, 46*(1), 73-92.

Thamdrup, B., & Dalsgaard, T. (2002). Production of N(2) through anaerobic ammonium oxidation coupled to nitrate reduction in marine sediments. *Applied and Environmental Microbiology, 68*(3), 1312–1318.

Welsh, D. T., Bourguès, S., de Wit, R., & Herb, R. A. (1996). Seasonal variations in nitrogen-fixation (acetylene reduction) and sulphate-reduction rates in the rhizosphere of. *Marine Biology, 125*(4), 619–628.
Genome comparison and molecular characterization of two novel diazotrophs from Narragansett Bay, RI

A.K. Jones\(^1\), R. Spinette\(^1\) and B.D. Jenkins\(^{1,2,*}\)
\(^1\)Department of Biological Sciences, University of Rhode Island, Kingston, RI, 02881
\(^2\)Graduate School of Oceanography, University of Rhode Island, Narragansett, RI, 02881

\*corresponding author:
Dr. Bethany Jenkins
University of Rhode Island
120 Flagg Road,
Kingston, RI 02881, USA
bjenkins@uri.edu;
Tel (1) 401-874-7551; Fax (1) 401-874-2202
Abstract:

The ability to reduce atmospheric nitrogen (N₂) to ammonia, known as N₂ fixation or diazotrophy, is a metabolic capability widely distributed amongst prokaryotes and it is essential for the input of new, reactive N to a variety of environments. Recently, heterotrophic bacterial N fixers residing in estuarine sediments have been acknowledged as important contributors to the overall N budget of these ecosystems. However, many specifics about their role in estuarine N cycling remain unknown, partly due to a lack of understanding regarding their autecology and a lack of cultivated representatives. Nitrogenase reductase (nifH) gene composition and prevalence in Narragansett Bay has revealed that two distinct phylogenetic groups dominate N-fixation in these sediments. Analysis of nifH transcripts has revealed one of these groups to be members of the Desulfovibrionaceae, belonging to the Deltaproteobacteria. To examine how nitrogenase activity in these organisms responds to the presence of different sources of combined N, and to link observed physiology with predicted capability based on gene content, we performed a two-part study that coupled high-throughput genome sequencing and comparative analysis to N fixation rate measurements. Here we present the genomic and physiological data in relation to N₂ fixation by two members of the Desulfovibrionaceae, isolated from sediments of the Narragansett Bay estuary in 2010 and 2011, respectively: Desulfovibrio sp. NAR1 and Desulfovibrio sp. NAR2. Genome sequencing revealed both isolates to be metabolically versatile and unique, and both isolates are capable of sustaining growth by N₂ fixation in the absence of reactive N. Our N fixation rate data indicate that nitrogenase activity is completely inhibited by the presence of ammonia, yet
uninhibited by nitrate and urea. This in turn agrees with observations made at the genome level, as both isolates and the additional *Desulfovibrio* genomes examined in comparison appear to lack the genetic capability to use nitrate or urea catabolically in order to meet cellular N requirements. This study indicates that the *Desulfovibrionaceae* may be limited in terms of the N sources they are capable of using catabolically, which could be a factor contributing to the observed N fixation by this group in sediments that are replete for sources of combined N, such as those of the Narragansett Bay estuary.

**Introduction:**

Estuarine sediments typically exhibit a nitrogen (N) cycle that is dominated by processes of N removal, such as coupled nitrification and denitrification (1). The opposing process of N₂ fixation by prokaryotic organisms, known as diazotrophs, is the primary source of reactive N to the world’s oceans and consequently acts as a control on both the N budget and primary production in many marine ecosystems (2-3).

Historically, N fixation has been a process primarily attributed to cyanobacterial species residing in the water column, although genetic potential to fix N is widely distributed amongst prokaryotes, including members of bacteria and archaea (4). Due to the energetically unfavorable nature of biological N fixation and typically abundant concentrations of combined N found in estuarine systems, heterotrophic N fixation was previously thought to be an inconsequential process in these environments (5). However, recent research involving direct observations of N₂ flux across the sediment-water interface in a variety of marine, salt marsh, and sea
grass systems support that heterotrophic N fixation is an important source of reactive N in these systems, and has begun to alter the historically held conclusions regarding the role of heterotrophic diazotrophs (6–10). Additionally, significant N fixation has been documented in waters where cyanobacteria are not believed to be present or active (11–13).

Previous investigation of heterotrophic N fixation in Narragansett Bay, RI, has revealed that most active N fixing organisms belong to two distinct phylogenetic groups, one of which consists of members of the Desulfovibrionaceae (14). Members of the Desulfovibrio spp. (Desulfovibrionaceae) are heterotrophic sulfate (SO$_4^{2-}$) reducers and have been shown to fix N in culture (15). Members of this genus are also noted for their ability to perform a wide variety of metabolic functions, including the reduction of sulfate to sulfur and sulfide species, the ability to utilize recalcitrant carbon sources, the ability to fix N, and the ability to transform certain metal species (15–18). Due to their wide variety of metabolic functions the Desulfovibrio are important in a number of industries and applications. They are considered a nuisance to the oil industry and can contribute to the souring of oil via the production of sulfides and corrosion of pipelines and wells (19), they are used in wastewater treatment plants as a way to remove sulfates and convert hydrogen sulfide byproducts into precipitated heavy metals (20), they play an important role in bioremediation by reducing and immobilizing heavy metals (21), and they hold potential for use as an alternate energy source in biological fuel cells (22).

Additionally, it has been known for some time that many sulfate reducing bacteria (SRB), particularly members of the Desulfovibrio, have the genetic potential
to fix N (23). These organisms have been shown to fix N in a laboratory setting (15, 24) and in a variety of habitats including coral reefs, photosynthetic microbial mats, mangrove sediments, sea grass rhizospheres (25–27), shallow estuarine sediments (6, 14, 28), bioturbated sediments (29), and salt marshes (9). There is further evidence supporting that these organisms play a critical role in supplying fixed N to their environment, particularly in anaerobic or anoxic sediments (14, 30) and benthic sediments which are N deficient (31–33). This mounting evidence has led to the ecological significance of the SRB, in terms of N fixation, being more widely accepted and acknowledged.

However, the majority of investigations involving these diazotrophs have focused on their ecological significance and role in the environment, primarily community composition, and group behavior and dynamics in situ. There have been no examinations to date of *Desulfovibrio* genome content in the context of nitrogen fixation capabilities and overall nitrogen metabolism, despite the fact that multiple complete genomes for diazotrophic *Desulfovibrio* exist (34–39). There could be several reasons for this; one potential reason is that as bacterial genomes have become easier and more cost effective to sequence, genome data is now being produced faster than it can be thoroughly analyzed. If this is the cause then it is likely that as analysis catches up with high-throughput sequencing technologies, thorough investigations of *Desulfovibrio* genomes in the context of their N fixation capabilities will begin to emerge. Another possible explanation is that because these diazotrophic SRB have only recently begun to be more widely accepted among the ecological community as important contributors of fixed N to their environments, investigations have not yet
shifted from answering broader questions regarding the ecological significance and role of these diazotrophs in the environment, to more targeted examinations of genome content at the individual level in the context of N fixation capabilities and overall N metabolism.

Of the cultivated *Desulfovibrio* species there remain relatively few environmental representatives with complete (closed and annotated), well-studied genomes. There are currently 2 deep-sea species, 2 freshwater species, a handful of estuarine species, and similarly few extremophilic representatives with complete genomes (34–42). Of the representatives with completed genomes there is still much yet to be discovered regarding their metabolism, nutrient cycling properties, and responses or adaptations to different environmental conditions and stressors, particularly from a molecular perspective as it pertains to the role of these organisms in N cycling.

There is a similar lack of information regarding N fixation rate data among the *Desulfovibrio*. Although there are previous studies examining N fixation rates for heterotrophic diazotrophs in sediments using the acetylene reduction assay (McGlathery *et al.* 1998; Welsh *et al.* 1996b), these studies have used environmental samples and thus are examining the overall N fixation ability of a mixed microbial consortia and are not able to establish cell-specific rates of N\(_2\) fixation, or a connection between those rates and the responsible cells. This lack of evidence is most likely due to the fact many of these organisms remain uncultured, and where there are cultivated representatives the potential contributions to heterotrophic N fixation by those specific organisms remain largely disregarded.
Whatever the underlying cause, there is an observable disconnect between the acknowledged role of the SRB, specifically the *Desulfovibrio*, as N fixers in the environment and our understanding of them at the individual genomic level as it pertains to their documented role as diazotrophs. The data presented here seeks to address this disconnect and begin closing the gap between observed environmental activities and a comprehensive molecular view of ecologically relevant members of the *Desulfovibrionaceae*.

In this study, we report the nitrogenase activity (NA) of two anaerobic, heterotrophic, diazotrophs (isolated from Narragansett Bay, RI sediments) under different combined N sources (ammonia, nitrate, and urea). We then present a comparative analysis of genomic traits related to observed NA and combined N utilization in these organisms in the context of other members of the *Desulfovibrio* genus.

**Materials and Methods:**

The environmental *Desulfovibrio sp.* examined here were isolated from sulfate-reducing, diazotrophic enrichment cultures inoculated with estuarine sediment gathered at a site off the northern tip of Jamestown, RI (41° 35.292’N 71° 22.748’W) (Spinette and Jenkins, unpublished data). Enrichment cultures were established using anaerobic culture methods described by Hungate (43), and individual species of sulfate reducing bacteria were isolated on roll tubes. NAR1 was isolated from sediment slurry gathered in June 2010 spanning a depth of 0 to 2 cm and NAR2 was isolated from sediment slurry gathered in October of 2010 spanning a depth of 0 to 2 cm.
Growth and Glycerol Stock Preparation of *Desulfovibrio* sp. NAR1

*Desulfovibrio* sp. NAR1 was grown in several media, all of which were variations on the NBSO media (.2 um filtered Narragansett Bay seawater, 60 mM lactate, 1.5 mM sodium thioglycolate, .45 mM sodium sulfide, 1 mM KH$_2$PO$_4$/K$_2$HPO$_4$, 1 mg/L resazurin as a redox indicator, and 300 mM NaHCO$_3$) in gas tight Hungate tubes with N$_2$:CO$_2$ gasses (80:20%) in the headspace. All cultures were grown at 21°C. Regular NBSO was used as the primary media for maintaining the isolate in culture, with the richer NBSOY (NBSO + .05% yeast extract) medium used occasionally. Media formulations used in the acetylene reduction assay were NBSO, NBSO + 12 mM NH$_3$, NBSO + 12 mM NO$_3^-$, and NBSO + 12 mM urea.

Culture biomass was estimated by calculating the number of genomes present in each culture from the total DNA extracted (for calculations used see Supplemental Figure S1). DNA was extracted from cultures grown on glass beads using the Qiagen DNeasy Plant Mini kit (Qiagen, Gaithersburg, MD, USA) with ~.30 g of glass beads as starting material, and a modified protocol (see Supplemental Information for complete extraction protocol) to lyse cells, the resulting cell lysate was then used in a phenol/chloroform extraction followed by an ethanol precipitation. DNA that was extracted for purposes other than genome sequencing was quantified using a Qubit fluorometer with the Quant-iT high sensitivity assay kit (Life Technologies, Grand Island, NY). Glycerol stocks of *Desulfovibrio* sp. NAR1 were made following a general protocol (see Supplemental Protocol 2), and cultures of NAR1 were successfully recovered from those stocks.
Acetylene Reduction Assay

The NA of *Desulfovibrio* sp. NAR1 was measured under various combined N conditions. 10 mL cultures of NAR1 containing 2 g of 1 mm diameter glass beads in carbonate buffered NBSO, NBSO + 12 mM NH$_3$, NBSO + 12 mM NO$_3^-$, NBSO + 12 mM urea were prepared in either duplicate or triplicate and NA was assessed using the acetylene reduction assay (ARA) and methods described previously by Capone (44). Acetylene was generated in house by reacting calcium carbide with water and was collected in a 1 L Supel-Inert film gas sampling bag. Acetylene was added to the culture tubes to a final concentration of 10 to 20% of the total headspace. A Shimadzu GC8 gas chromatograph (Shimadzu Corporation, Kyoto, Japan), with a 2.5 m long stainless steel column containing Haysep T packing 80/100 mesh, was used to measure ethylene production in all ARAs. The injector and column were set to 130°C and 100°C, respectively. Gas samples of 100 uL were taken from the tube’s headspace with a gas tight syringe and immediately injected into the gas chromatograph. Ethylene production was measured over the course of 24 hours. Samples were usually measured 1, 3, 6 and occasionally 24 hours after acetylene was added. Cultures that were part of the same assay were inoculated at the same time using the same parent culture. A set of multiple potential parent cultures was established for each set of experimental tubes; NA was measured in the parent tubes to ensure that the inoculum for the experimental cultures was actively fixing N at the time of inoculation, once NA was established in one member of the parent culture series the next un-tampered parent culture was used to inoculate the experimental culture series. A series of potential parent cultures was necessary, rather than measuring the NA of one culture
and using that same culture as a parent inoculum, to mitigate the negative effects of long term acetylene exposure on the growth of NAR1. Cultures were also no longer sterile after being used in the acetylene reduction assay (ARA), and so could not themselves be used as parents. Additionally, measurement of NA in the parent culture was a critical means of determining when the culture was ready to be used as inoculum, since cell enumeration of the NAR1 isolate remains challenging, and as of yet no hard correlation between culture age and NA has been established.

After establishing a baseline NA using N limited cultures (Fig. 7), ethylene production for cultures of NAR1 grown under different combined N treatments was measured (Fig. 9). NA was measured on separate, triplicate sets of 10mL carbonate-buffered NBSO cultures containing 2g of 1mm diameter glass beads with the addition of one of the following: 12 mM NO₃⁻ (nitrate set), 12 mM NH₃ (ammonia set), 12 mM urea (urea set). A control set of tubes was grown under previously discussed N limited conditions. The measurements taken on Day 8 and Day 10 utilized triplicate cultures for all combined N treatments and the N limited control, an additional control only time-point with culture samples in duplicate was measured on Day 14 to ensure that NA was decreasing and that peak NA had been observed at the time combined N treatments were being examined. Biomass measurements for all samples used were obtained by extracting DNA from each tube, and then calculating the total number of genomes present based on the amount of DNA extracted. This data is presented in Figure 8. Biomass estimates were then used to normalize ethylene production rate to cell abundance.
**Genome sequencing, assembly, annotation, and gene comparison**

*Desulfovibrio sp.* NAR1 was grown under anaerobic conditions in carbonate buffered NBSOY media (filtered Narragansett Bay seawater, 60 mM lactate, 1.5 mM sodium thioglycolate, 450 μM sodium sulfide, 1 mM KH₂PO₄/K₂HPO₄, 0.05% wt/vol yeast extract, 1 mg/L resazurin as a redox indicator, and 300 mM NaHCO₃) with a 20:80% CO₂:N₂ gas mixture in the headspace, until an optical density of 0.2-0.3 was reached. NAR1 will only exhibit planktonic growth when grown in NBSOY media, and the absence of biofilm was necessary for successful nucleic acid extractions. Once the desired OD had been reached, culture samples were harvested by filtering 2 mL of culture onto 0.2 um polyester filter membranes (Millipore, Darmstadt, Germany). Filters were then either placed into sterile 2 mL screw-cap microcentrifuge tubes (Fisher Scientific, Pittsburgh, PA) tubes and flash frozen using liquid nitrogen for later analysis, or immediately used in DNA extractions. NAR2 cultures were grown under anaerobic conditions in carbonate buffered NBSO media (filtered Narragansett Bay seawater, 60 mM lactate, 1.5 mM sodium thioglycolate, 450 μM sodium sulfide, 1 mM KH₂PO₄/K₂HPO₄, 1 mg/L resazurin as a redox indicator, and 300 mM NaHCO₃) + 12 mM NH₃ with a 20:80% CO₂:N₂ gas mixture in the headspace, until an optical density of 0.2-0.3 was reached. Cells were harvested in the same manner as NAR1.

Genomic DNA was extracted from polyethylene filters using the Qiagen DNeasy Plant Mini kit (Qiagen, Gaithersburg, MD, USA) with a modified protocol (see supplemental data for complete extraction protocol) to lyse cells, and the resulting cell lysate was then used in a phenol/chloroform extraction followed by an ethanol precipitation. Multiple DNA extracts were pooled together during the ethanol
precipitation step to increase the DNA concentration in the final extract. DNA was re-suspended in 10mM Tris-HCl. DNA quality and quantity was assessed using the Thermo Scientific Nanodrop 2000c, and a Qubit fluorometer was used with the Quant-iT broad range dsDNA assay kit (Life Technologies, Grand Island, NY) as an additional means of quantifying DNA and affirming the concentration reported by the Nanodrop. To verify the purity of the DNA extracts the 16s rDNA and nifH genes were amplified using previously described PCR primers and protocols (14, 45). Sanger sequencing of cloned PCR products was performed at the University of Rhode Island Genome and Sequencing Center (GSC). Both genes were analyzed using at least 10 clones, and all clones for each gene were required to perfectly match and align before the isolate culture was considered pure. Genomic DNA from NAR1 and NAR2 was sheared using a Covaris S220 high performance ultra sonicator (Covaris, Woburn, MA, USA) and a paired end (PE) sequencing library (~700 bp inserts) was constructed using the Apollo 324 robot (WaferGen Biosystems, Fremont, CA, USA) and the Illumina Nextera kit chemistry (Epicentre Technologies, Madison, WI, USA). The sequence libraries for both isolates were then multiplexed and run on the Illumina MiSeq instrument according to manufacturers specifications. Preparation of libraries and sequencing of genomes for both isolates was done by the staff of the URI GSC. Resulting read data for both isolates was trimmed using Trimmomatic (46) to remove adapters and low quality bases. Trimmomatic was used in paired-end (PE) mode using the following parameters and settings: -phred33 (use phred 33 scores), ILLUMINACLIP (cut adapters and other Illumina specific sequences from the read) with a seed mismatch of 3, palindrome clip threshold of 25, simple clip threshold of 9,
minimum adapter length of 4, and the keep both reads parameter set to “true”, LEADING (remove low quality bases from start of read) with a minimum quality score threshold of 15, TRAILING (remove low quality bases from end of read) with a minimum quality score threshold of 15, SLIDINGWINDOW (use a region of bases to assess quality) with a window length of 100 bp and a minimum quality score threshold of 15, MAXINFO (perform adaptive quality trim) with a target length of 100 bp and strictness setting of .8, CROP (removes bases from end of read, user specifies number of bases to keep) set to 240 bp, HEADCROP (removes bases from start of read, user specifies number of bases to remove) set to 10 bp, and MINLEN (removes reads that fall below specified minimum length) set to 100 bp. Genome sequences were assembled using SPAdes de-novo assembler version 3.0.0 (47) with the pre-selected k-mer size range for MiSeq 250bp reads (k= 21, 33, 55, 77, 99, 127) and the --careful (read correcting) parameter turned on. The quality of genome assemblies for both isolates was assessed using QUAST (48), the complementary assessment software to the SPAdes assembler. Genome assemblies were further assessed for quality using the specialized assembly evaluation software REAPR (49) as a final step before annotation. Both genomes were annotated using the Integrated Microbial Genomes Expert Review (IMG/ER) (50) pipeline. Genomes were also annotated using Rapid Annotation using Subsystem Technology (RAST) as a point of comparison to IMG/ER, however, IMG/ER annotations were ultimately accepted for the final draft and used in comparative analysis. It is important to note that annotations generated by both annotators agreed across most major metabolic pathways. A separate set of open reading frames (ORFs) was also called for both isolates using command-line Prodigal
(51) as a means to check the ORF calling of both automatic annotators. Some genes, particularly those of importance to this study (e.g. *nif*), were annotated manually using Geneious version 6.1.6. The gene content of both environmental *Desulfovibrio* isolates in comparison to 34 additional *Desulfovibrio* genomes was examined using the Function Analysis feature in the IMG/ER (50) tool set. Results of the analysis in IMG/ER were confirmed visually by examining the sequence and annotating the gene or genes in question using Geneious version 6.1.6, and by TBLASTN.

**Phylogenetic Analysis and Evolutionary Gene Networks**

20 different vertically transferred marker genes from both of our *Desulfovibrio* isolates and 34 additional *Desulfovibrio* representatives were selected and used as input into a multigene phylogenetic tree. A list of genes can be found in Supplemental Table S4 and a specific list of additional genomes can be found in Supplemental Table S3. The genomes selected for use in this comparison are all complete or well-annotated draft *Desulfovibrio* genomes available in GenBank or IMG databases at the time of this study. The 20 marker genes were extracted from the complete genome sequences or assemblies of all 36 representatives using Geneious v. 6.16. The marker genes were then all aligned separately using MAFFT version 7 with the G-INS-i setting (52), and resulting alignments were concatenated using SequenceMatrix (53). jModelTest (54) was used to assess the best-fit model of nucleotide substitution for each gene alignment. A multi-gene phylogenetic tree (Fig. 1) was constructed from the concatenated gene alignment using Bayesian phylogenetic inference software MrBayes version 3.2.1 (55), with the best-fit model predictions from jModelTest used
as additional input to inform the tree-building software. The program was run for 1000000 generations, which was sufficient for convergence to be reached.

*Desulfovibrio* sp. NAR1 and *Desulfovibrio* sp. NAR2, their three closest relatives (*D. piezophilus*, *D. aespoeensis* str. Aspo-2, *D. desulfuricans* str. ND132 for NAR1; and *D. acrylicus*, *D. desulfuricans* subsp. aestuarii, *D. alaskensis* str. G20 for NAR2), and *D. vulgaris* str. Hildenborough were chosen for inclusion in evolutionary gene network analysis. The base evolutionary gene network was constructed using EGN (56), with amino acid multi-FASTA files of all proteins for all included organisms as input. The local alignment software used was for the creation of input files was BLAST, with a minimum e-value setting of 1e-05. The “quicker” algorithm was selected for creating edge files, with an e-value threshold of 1e-05, a hit identity threshold of 65%, a minimum hit length requirement of at least 70% of the smallest homolog, a minimum hit length of 75 amino acids, the best reciprocal threshold of the best e-value was set to 5%, and the minimum match coverage threshold of the sequence length was set to 65%. The resulting base evolutionary gene network was then filtered using a python script (written by Jim Grandpre and Scott Grandpre), to reveal networks that contained only NAR1 and NAR2 (Fig. 3.3), only NAR2 (Fig. 3.4), and only NAR1 (Fig. 3.2).

**Whole Genome Alignment**

Mauve version 2.3.1 (57) was used to compare FASTA sequences of *Desulfovibrio* sp. NAR1 to the complete genome sequence of *Desulfovibrio piezophilus*, and *Desulfovibrio* sp. NAR2 to the complete genome sequence of *Desulfovibrio alaskensis* str. G20. The contigs of each environmental isolate were
reordered relative to the genome of their respective reference, the final alignment (Fig. 2) shows the reordered draft genomes aligned to their references. Aligning of reordered genomes aided in viewing the regions (local collinear blocks, or LCBs) of draft genomes that were presumably homologous and internally free from genomic rearrangements relative to the parts of the complete genome of their respective references. The reordered draft genomes are also important for any future work involving genome gap closure or additional genome comparisons.

**Results and discussion:**

**Growth and acetylene reduction of *D. sp. NAR1* under N limited conditions**

Cultures of NAR1 were maintained under anaerobic conditions on N limited sulfate reducing media NBSO, either with or without glass beads as a biofilm substrate. Under N limiting conditions NAR1 exhibited biofilm growth, making cell enumeration very challenging. Initial attempts to assess growth of the NAR1 isolate using spectrophotometry failed because the isolate grows as a biofilm under almost all conditions; the only condition under which it will not form a biofilm is when it is grown with NBSOY and there is no physical surface present that the isolate can easily adhere to (such as sediment, glass beads, diatomaceous earth, etc.). Cultures of NAR1 grown with NBSOY and no additional physical surface were planktonic and could be measured using spectrophotometry. However, since our assessment of NA in this isolate involved media formulations other than NBSOY that also contained glass beads, spectrophotometry was not a viable option for assessing growth in those cultures. Other methods, such as flow cytometry and qPCR were attempted as a means to measure the growth of NAR1. Flow cytometry was not a practical option as it was
only possible with planktonic cells (grown in rich media), and even non-biofilm cells produced enough exopolymeric substances (EPS) to cause cells to stick to the inside of the injection line and interfere with the functioning of the instrument. Quantitative PCR attempts also failed, most likely due to the EPS produced by this organism. DNA extracts from older cultures did not amplify as well as DNA extracted from younger cultures, suggesting an inhibition of the DNA polymerase enzyme. As DNA concentration was not very high in older cultures (≤10 ng/μL), and DNA of comparatively lower concentrations from younger cultures amplified, it is unlikely that DNA concentration is the cause of the failed amplification in DNA samples taken from older cultures. However, it is likely that older cultures have more EPS present, supported by the obvious biofilm formation seen in cultures that are ≥2 weeks old. Protein contamination in DNA extracts, most likely due to EPS and biofilm material, was also an issue that was encountered in gathering DNA for genome sequencing, and was resolved by ethanol precipitation and pooling of multiple samples to purify and concentrate the DNA. However, because the goal of DNA extractions from experimental tubes used in the ARA was to enumerate cells in a specific culture, pooling of multiple replicates would defeat the purpose, and the amount of DNA extracted from a single culture was low enough that additional ethanol precipitation to purify DNA and remove EPS would likely cause most of the sample to be lost. For these reasons, qPCR was not a practical method for quantifying biomass in this study. Biomass estimates were ultimately made by extracting DNA from pure cultures grown under N limited conditions, and subsequently using the DNA yield to calculate the total number of genomes present in the culture volume, which was used as
representation of cell abundance. Figure 6 shows the log growth of NAR1 under N limited conditions. This organism exhibits very slow growth and very little change in overall biomass when grown under N limited conditions. However, slow growth is not unusual for environmental bacteria such as members of the *Desulfovibrio* (58), and it is likely that there is some loss of DNA during the extraction process, which would result in a lower reported cell count.

Ethylene production by NAR1 (Fig. 7) was first measured on N limited cultures over a period of ~14 days prior to examining the effects of combined nitrogen on NA in the NAR1 isolate. Previous pilot experiments had shown that no NA occurred before the third day after inoculation, so these time points were not included in the study. Figure 7 shows NA for the NAR1 isolate for the same time period in which growth was measured (Fig. 6). These measurements indicate that peak NA in N limited cultures of NAR1 occurs early on in the growth of the organism, late lag or early log phase, which is a trend that had been observed in previous assays (data not shown). Peak NA is within range of that described for other bacterial isolates (59).

The observed high activity on Day 5, when the culture is young and the cell count is lower, may be due to the N demands of the organisms as they prepare to enter log phase. It is unclear at this time whether biofilm formation plays a role in increased NA or overall N demand. Since *Desulfovibrio* biofilms are known to be composed primarily of protein (60, 61), this remains a possibility and should be considered in future investigations. Further studies would be needed in order to elucidate what specific factors contribute to timing of peak NA in this isolate.
Growth and acetylene reduction of NAR1 under differing combined N conditions

After establishing a baseline NA using N limited cultures (see previous section), ethylene production for cultures of NAR1 grown under different combined N treatments was measured using the ARA (Fig. 9). NA was not observed with the ammonia treatment at either time-point and this agrees with what is known regarding nitrogenase gene regulation in response to ammonium availability for aerobic N fixers (62), although the inhibition of NA by ammonia has not previously been established for sulfate-reducing diazotrophs, or other anaerobic N fixers. No negative effect on NA was observed with either the nitrate treatment or the urea treatment, which is not in agreement with the previously held idea that NA is inhibited by other sources of combined nitrogen (63). The fact that neither nitrate nor urea have an inhibitory effect on NA also agrees with our genomic analysis of this isolate, since there is nothing in the genome of NAR1 to suggest that this organism can use either nitrate or ammonia as a source of N. Rates of ethylene production for the nitrate treatment are very similar to those of the N limited control, while the rates for the urea treatment are slightly higher for both time-points. This increase in NA with the presence of urea has been observed in previous assays with this isolate, and the cause is currently unknown. As the organism lacks any evidence of a urease or a urea amidolyase, it is unlikely that urea is being broken down into any intermediate that could be affecting N fixation, or utilized as a carbon source, which could affect growth. The lack of effect on growth is further demonstrated by examining cell abundance for the urea treatment at both time-points (Fig. 8), which is very similar to that of the N limited control.
The total cell abundance data corresponding to the ARA experiment does not follow what would be expected for the typical logarithmic growth of bacteria. There could be several reasons for this, the most likely cause having to do with how the replicates are set up and inoculated. Since NAR1 grows in a biofilm, it is very difficult to control how many cells are contained in the inoculum, even when the same volume of culture is used for every inoculation. Because it is impossible to ensure a homogenous suspension of cells with a biofilm-forming organism, it is likely that one tube may be inoculated with a large segment of dislodged biofilm while another tube may receive fewer cells in comparison. Currently, our methods do not allow us to break up the biofilm into segments that can be accurately quantified and then used as inoculum, or to force NAR1 to grow planktonically under N limited conditions. Keeping in mind that replicates in this experiment are separate cultures and that Fig. 8 shows the total number of cells present in these cultures measured at the respective time-point, instead of the same culture with change in cell abundance quantified over time, it is plausible that the lower cell count on day 10 for all treatments except nitrate is due to fewer cells being present in the inoculum used for those cultures compared to the day 8 cultures. Possibly inflated ethylene production rates due to discrepancies in culture density are corrected for by normalizing ethylene production rates to cell abundance.

**Genome sequencing outputs, assembly and annotation**

Genome sequencing using the Illumina MiSeq platform was carried out for *Desulfovibrio* sp. NAR1 and *Desulfovibrio* sp. NAR2. The number of paired-end, 250bp reads obtained for each isolate was more than 16 million. Phred per-base
quality scores of ≥ 10 were reported for NAR1 and NAR2 raw reads, which represent inferred base call accuracy of at least 90% (64, 65), for all reads. Ambiguous N bases were not detected in either raw data set. Although these values indicate that the probability of an incorrect base call was minimal for the raw sequences, both data sets were trimmed prior to being used in de novo assemblies. The number of paired-end reads remaining for each isolate after trimming was more than 13 million, which represents a range between 700x and 900x expected genome coverage for each isolate. Phred per-base quality scores of ≥ 30 were reported for NAR1 and NAR2 trimmed reads, which in turn represent inferred base call accuracy of at least 99.9% for all reads. Ambiguous N bases were not detected in either trimmed data set. The average read length after trimming was 200bp and 210bp for NAR1 and NAR2, respectively. These resulting quality values represent an improvement over the raw data sets and indicate that the probability of an incorrect base call is minimal for these sequences, at this point both trimmed data sets were considered to be appropriate for generating de novo assemblies.

Statistics for the de novo assemblies of D. sp. NAR1 and D. sp. NAR2 are shown in Table 1 (Table 1). The NAR1 genome data was assembled first, and used to establish a pipeline for working with the NAR2 data. Both isolate data sets were trimmed and assembled using identical parameters, using all reads in their respective sets. Individual contig coverage values were averaged for both assemblies, and are reported in Table 1. Fold coverage for individual contigs for NAR1 and NAR2 ranged from 339x-909x, and 705x-1,876x, respectively (Supplemental Tables S1 and S2). Comparisons of assembly statistics such as \( N_{50} \), \( N_{75} \), number of contigs, and longest
contigs for these isolates compared to other published draft genomes indicate that good-quality assemblies have been achieved for both isolates (66, 67). Results of assessments with assembly evaluation software REAPR (49) also supported that high-quality draft genomes had been generated, finding 95.33% of assembled genome bases to be error free for NAR1 and 99.36% for NAR2. The total number of nucleotides for NAR1 was near to the completely sequenced genomes of *Desulfovibrio desulfuricans* subsp. *aestuarii* (draft) and *Desulfovibrio piezophilus* (Table 2) (38, 68), which are both ~3.6Mb. The total number of nucleotides for NAR2 was near to the completely sequenced genome of *Desulfovibrio vulgaris* str. Miyazaki F (Table 2) (68), which is ~4.0Mb. The G+C contents of both isolates are near to those of other *Desulfovibrio*, which are between 45 and 67% (34–42, 50). Contigs from both isolate assemblies were ultimately reordered using Mauve alignments (57), as shown in Figure 2.

Genomes were annotated with IMG/ER, RAST (50, 69) and manually using BLAST tools (70) into protein coding genes (PCGs), RNA coding genes (RCGs) and hypothetical PCGs (Table 2). The majority of PCGs in both isolates, ~80%, have function assignments.

**Comparison of isolate genomes using IMG/ER and RAST**

A combination of IMG/ER, RAST, and manual annotation using BLAST tools (50, 69, 70) was used to annotate both isolate draft genomes. IMG/ER was used as the primary annotator and comparison tool, RAST was used sparingly, and manual annotation was used primarily where automatic annotations seemed questionable, were missing, or when genes were of special interest to this study.
Total coding bases were split into PCGs, RCGs, and hypothetical PCGs. This data, along with the breakdown of RCGs into those accounting for rRNA, tRNA, and other RNA genes, is shown in Table 2 along with the same data for two closely related Desulfovibrio species for each isolate; *D. desulfuricans* str. ND132 and *D. piezophilus* for NAR1, and *D. desulfuricans* subsp. *aestuarii* and *D. alaskensis* str. G20 for NAR2. The more phylogenetically distant representative, *D. vulgaris* str. Miyazaki F, was also included as an additional point of comparison. Tabulated attributes are similar overall between isolates and the previously sequenced genomes listed. However, the NAR1 genome is note-worthy with respect to the two annotated clustered regularly interspaced short palindromic repeat regions (CRISPRs), where 3 of the other genomes included in Table 2 have only one CRISPR region and 3 do not have any. CRISPR-Cas systems play a role in adaptive immunity against phages and other invading genetic elements and are present in approximately 40% of sequenced eubacteria genomes and 90% of archaea genomes (71, 72). It is interesting to note that NAR2 lacks a CRISPR-Cas system, particularly since these organisms were isolated from the same site and would presumably have been exposed to similar phage attacks and the same foreign DNA in the environment.

Isolate PCGs with KEGG, COG, Pfam and TIGRfam (73–76) annotations in IMG/ER were compared across subcategories in terms of number of genes contained in each subcategory and the percentage of total PCGs with pathway association that the gene number represents. The percentage is not particularly meaningful on its own because it is taken from the total number of genes with pathway association, which can vary depending on which database is being considered, and it is not a
representation of total PCGs. Even so, the percentage of genes represented in each of
the various subcategories remained relatively consistent between the two isolates. So
from a broad, overall standpoint, the two isolates look similar in terms of gene content
with function assignments. However, not all possible metabolic subcategories are
represented, and details regarding gene content rather than overall gene number cannot
be assessed in this manner.

The RAST database uses SEED (69) as the core means of annotation and
subsystem assignment, which IMG/ER does not, so RAST was used as an additional
point of comparison for isolate annotations made by IMG/ER. Comparison of
subsystem gene content for both isolates in RAST revealed the NAR1 isolate to have a
relatively large number of genes involved in cell wall and capsule formation; 134
genes compared to NAR2, which has 81 genes with that subsystem assignment. Of the
136 genes involved in cell wall and capsule formation in NAR1, 57 are involved in
capsular and extracellular polysaccharide production and an additional 47 genes are
involved in cell wall and capsule formation with no subcategory. The NAR2 isolate,
comparatively, has 35 genes involved in extra cellular polysaccharide production and
only 15 genes are involved in cell wall and capsule formation with no subcategory. This
difference is in agreement with the observed exudate and biofilm production in
NAR1, which is not observed at all in the NAR2 isolate. The number of genes
assigned to motility and chemotaxis subsystems is also of interest; the NAR1 isolate
has 136 genes assigned to this subcategory, with 100 associated with flagellar motility
and 36 being assigned to bacterial chemotaxis. NAR2, comparatively, has 94 genes
assigned to motility and chemotaxis, with all of these genes being associated with
flagellar motility. This difference between the two isolates in terms of genes involved in chemotaxis could also be related to the observed biofilm formation in NAR1, and could be part of the pathway that signals biofilm formation. As this pathway is not yet described among the Desulfovibrio, these genes should be considered as targets for future studies involving biofilm formation in this genus.

Both isolates are notable in regards to the number of genes assigned to virulence, disease, and defense subcategories. The NAR1 isolate has 61 genes assigned to this subsystem, and the NAR2 isolate has a comparative 70 genes. Within this subsystem, both isolates have genes involved in copper homeostasis, mercury, cobalt, zinc, and cadmium resistance. The NAR1 isolate has a gene encoding MliC, a membrane bound inhibitor of c-type lysozyme. This may partially explain why the NAR1 isolate was difficult to lyse enzymatically. The NAR2 isolate does not have this gene, which agrees with the fact that no difficulties with enzymatic lyses were noted for this isolate. The NAR2 isolate has genes encoding proteins involved in multidrug-resistance (MDR) efflux pumps, and is predicted to possess a functional multidrug and toxic-compound extrusion (MATE) family MDR efflux pump (77). The NAR1 isolate has genes encoding proteins involved in arsenic resistance (ars, arr, and ACR3), NAR2 does not have any of these genes present. The Desulfovibrio genus is often regarded for its resistance to heavy metals and for the ability to transform certain toxic heavy metal species (18, 21), so the presence of genes involved in this sort of metabolism in both isolates can be expected. And while it may not be surprising that both isolates possess these sorts of resistances, it does uniquely equip them for survival in Narragansett Bay sediments, as relatively high levels of Cr, Cu, Ni, Pb, and
Zn have been found in Bay sediments (78). Additionally, Narragansett Bay sediments have been shown to be contaminated with Hg, MeHg (79), and arsenic (80).

Also within the virulence, disease, and defense subcategory, both isolates have predicted beta-lactamases. Antibiotic resistant bacteria have been found widely in aquatic environments (81, 82), including urban river sediments (83), and Desulfovibrio have been shown to contain beta-lactamases, with certain species having implications in human Desulfovibrio infections (84). So although the presence of predicted beta-lactamases in these isolates may not be surprising in regards to environment or genus assignment, it should still be noted and considered as a possible target for future physiological studies, especially since the potential human pathogenicity of these isolates remains unknown. The NAR1 isolate in particular should be considered, as it contains additional putative genes for colicin V production and fosfomycin resistance. The presence of these genes has been observed and annotated in other Desulfovibrio, but there has yet to be physiological confirmation of colicin V production in this genus, while fosfomycin resistance has already been observed in some Desulfovibrio species with clinical relevance (85).

**Whole genome comparison and alignments**

The draft genomes of isolates Desulfovibrio sp. NAR1 and Desulfovibrio sp. NAR2 were compared to all complete or well-annotated draft Desulfovibrio genomes available in GenBank or IMG databases (50, 86) at the time of this study (for a list of genomes see Supplemental Table S3), which was a total of 34 additional genomes. A multi-gene phylogenetic tree (Figure 1) for all 36 Desulfovibrio representatives was constructed using 20 different vertically transferred genes (Table S4), to firmly
establish the phylogenetic relationship among the genomes being examined and to provided a more comprehensive assessment and point of reference in regards to which *Desulfovibrio* representatives were most closely related to environmental isolates NAR1 and NAR2. This kind of assessment had been made prior solely on the basis of 16s, *nifH*, and *dsrAB* genes. The two environmental isolates do not cluster close to one another within this group; the NAR1 isolate clusters most closely with *D. piezophilus*, *D. desulfuricans* str. ND132, and *D. aespoensis* str. Aspo-2, while NAR2 clusters most closely with *D. acrylicus*, *D. desulfuricans* subsp. aestuarii, and *D. alaskensis* strains. These results agree for the most part with what had been seen previously with the three-gene assessment, however *D. alaskensis* strains lack *nif*, so a relationship based on N fixing genes could not be examined.

From the *Desulfovibrio* genome phylogeny shown in Fig. 1, the closest relative to each isolate with a complete, gapless genome was selected as a reference for whole genome comparison, alignment, and contig re-ordering using Mauve version 2.3.1 (57). Results of this alignment can be seen in Figure 2. Each local collinear block (LCB) contains a colored similarity profile of the local sequence, with the height of the colored profile corresponding to the average degree of sequence conservation in that section. Areas that are white, which are seen most frequently in the NAR2 alignment, suggest that this genome contains a significant amount of sequence variability compared with *D. alaskensis* str. G20, which is not surprising as G20 is only 89% similar to NAR2 at the 16s level. The more closely related organisms to NAR2 did not have closed genomes and so could not be used as references in this assessment. The NAR1 alignment contains less white space than the NAR2 alignment,
which agrees with the higher level of sequence similarity, 95%, between NAR1 and \textit{D. piezophilus} at the 16s level. In both alignments note the “X” pattern formed by connected LCBs, this typically occurs at the origin of replication in aligned genomes. The number, arrangement, and heights of similarity profiles in the whole genome alignments of both environmental isolates are indicative of organisms related at the genus level, but not at the species or strain level.

**Gene network analysis**

Evolutionary gene networks (87) were used to compare the genomes of the two environmental isolates, three of their closest relatives, and \textit{D. vulgaris} Hildenborough, whose genome has been well studied and serves as an additional point of comparison. The genomes of \textit{D. piezophilus}, \textit{D. desulfuricans} str. ND132, and \textit{D. aespoeensis} str. Aspo-2 were included as part of the NAR1 cluster, and the genomes of \textit{D. acrylicus}, \textit{D. desulfuricans} susbsp. aestuarii and \textit{D. alaskensis} str. G20 were included as part of the NAR2 cluster. The initial gene network was run using parameters discussed previously in the Methods section, and results were subsequently filtered to select networks that consisted of only NAR1 and NAR2 (Fig 3.2), only NAR2 (Fig 3.4), and only NAR1 (Fig. 3.3). Selected connected components involving N fixation proteins from the un-filtered network are shown in Fig 3.1.

The un-filtered network consisted of approximately 6,000 connected components, primarily housekeeping proteins, proteins involved in basal metabolic activities, and numerous hypothetical proteins. Connected components that were of particular interest in the unfiltered network were those involving the N fixation (\textit{nif}) genes, as shown in Fig 3.4. The NifH network (Fig. 3.1 A) consists of all
representatives included in this analysis that are capable of fixing N, the shape of the
network indicates how similar the NifH sequences are between these representatives at
the amino-acid level, with an edge (black line) connecting every node (circle) to every
other node in the network. The NifB network (Fig. 3.1 B) consists of all included
representatives capable of fixing N except *D. vulgaris*. This network is particularly
interesting in that we can see sequences from each of the environmental isolates
clustering distinctly with sequences from their closest relatives, the warmer colored
circles being NAR1 and associated relatives, and the cooler colored circles being
NAR2 and associated relatives. The NifK protein network (Fig. 3.1 C) shows similar
relationships to those seen with the NifH network, except that *D. vulgaris* is once
again absent. The NifD network was identical to the NifK network, and not shown.

The network filtered for connected components containing only NAR1 and
NAR2 consisted of a total of 9 connected components. These were proteins involved
in choline/glycine/betaine transport (Fig 3.2 A), metal dependent hydrolases of the
beta-lactamase superfamily (Fig. 3.2 B), glyoxalase family proteins-glyoxalase I (Fig.
3.2 C), xanthine dehydrogenases catalases/peroxidases, permeases of drug metabolite
transporters (DMT), periplasmic aromatic aldehyde oxidoreductases, and hypothetical
membrane proteins. Because these two isolates do not share a close phylogenetic
relationship within the *Desulfovibrio* considered here; as indicated by separate
clustering in Fig. 1, and only 87% 16s sequence similarity between the two isolates,
these networks could indicate proteins that confer specific benefits for survival in
Narragansett Bay sediments. The fact that some of these genes (e.g. beta-lactamases,
DMTs) are involved in bacterial defense, bacterial detoxification (glyoxalase family
proteins) or as of yet have an un-described function but are exclusively shared between these two isolates, support this hypothesis. Because of this potential connection these proteins should be considered as targets for future investigations, especially those involving transcriptomic or gene expression analysis.

The filtered network containing only NAR1 had approximately 650 connected components (see Appendix A for a list of corresponding proteins), the majority of these connected components consisted of proteins involved in signal transduction and amino acid transport, with a smaller portion of the overall networks consisting of proteins involved in bacterial defense. Components of particular interest are shown in Fig 3.3, which include a putative sensory box/GGDEF family protein network (Fig 3.3 A), a putative diguanylate cyclase and receptor proteins network (Fig 3.3 B), and a periplasmic binding and signal transduction proteins network (Fig 3.3 C), as they have implications in biofilm formation in NAR1. These proteins and their potential role in this isolate’s biofilm production will be discussed further in a later section. Additional proteins found to be unique to NAR1 primarily had to do with signal transduction and amino acid transport, which is suggestive of involvement in biofilm formation and exudate production. The fact that these proteins network exclusively from any predicted proteins in NAR2 agrees with what we have observed at the physiological level, with biofilm and exudate production being restricted to NAR1 and not observed at all in NAR2.

There were only 2 connected components that consisted of just NAR2 (Fig 3.4), non-specific predicted membrane proteins (Fig. 3.4 A), and bacteriophage head to tail connecting proteins (Fig. 3.4 B). The two non-specific predicted membrane
proteins (Dn2DRAFT 02896, Dn2DRAFT 02903) share the highest degree of BLAST homology using BLASTX with a hypothetical protein in *Thiocapsa marina*, a purple sulfur bacterium. Dn2DRAFT 02896 shares 59% amino acid identity across 97% of the query, and Dn2DRAFT 02903 shares 60% amino acid identity across 98% of the query with a hypothetical protein (Seq ID: ref|WP_007193013.1) in *T. marina*. All BLAST results with a high enough degree of amino acid similarity to be of interest (≥ 60% sequence similarity) were to other hypothetical proteins, primarily from beta-proteobacteria, and so did not reveal any insight as to the possible function of this protein in NAR2.

**Nitrogen fixation**

Of the 34 representatives used in comparison to the environmental isolates, only 7 representatives lack the *nif* operon (Table 3). Only one representative, *Desulfovibrio* sp. U5L has genes for an alternate Fe-Fe nitrogenase. Both environmental isolates NAR1 and NAR2 have a full *nif* operon (Fig. 4), with the arrangement of their N fixation gene cluster being similar to those of their closest N-fixing relatives. There is also supporting physiological evidence that both isolates fix N. The presence of an iron-molybdenum nitrogenase appears to be a shared characteristic for this representative group of *Desulfovibrio*.

In addition to the examination of N fixation, an analysis of metabolism of other N substrates (ammonia, nitrate, and urea) was performed and discussed in the following sections. The analysis of additional aspects of N metabolism in these isolates is critical to improving our understanding of why *D*. sp. NAR1 and *D*. sp. NAR2 exhibit the N fixation behavior we have observed in Narragansett Bay, as an
inability to use other sources of N that are commonly found in the environment could account for a continued need to fix N. To do that it is important for us to be able to take the physiology discussed in the previous sections, and connect it with related functional gene content, which is discussed below.

**Urea metabolism**

Currently, the most attention regarding bacterial urea metabolism is given to organisms that make up mammalian gut consortia and intestinal human pathogens, and little focus has been placed on the urea metabolism of environmental representatives like the *Desulfovibrio*, or on sulfate-reducing bacteria in general. However, there have been some examinations of the uptake and metabolism of urea by environmental bacteria and phytoplankton (88, 89). These studies have shown that rates of bacterial urea uptake in the environment are highly variable, genes for urea transport and catabolism are not wide-spread amongst bacteria, and that other forms of N are generally preferred to urea, which could be due to the fact that urea catabolism is an energetically expensive process. Although some members of the *Desulfovibrio* are known to have urea transporters and ureases, little is known about the fate of urea once it enters a *Desulfovibrio* cell.

There exist at least four families of transporters that facilitate selective permeation of urea: an ATP-dependent ABC type urea transporter (90), an ion motive force-dependent urea transporter (91), an acid-activated urea channel that belongs to the urea/amide channel family (92), and the urea transporter (UT) family, this last type being the most widely distributed family. UT members are found in bacteria, fungi, insects and vertebrates (91, 93–96). In many bacteria and eukaryotes, urea in the cell
can be broken down to ammonia and CO$_2$ by a urease. Some bacteria and eukaryotes also use urea amidolyases (UALase) to decompose urea (88). A crystalline structure for a UT family urea transporter from *Desulfovibrio vulgaris* Hildenborough is available and its activity and mechanism of action have been proven in vivo (97). However, subsequent urea metabolism after transport has not been thoroughly investigated.

In this study, an examination using IMG/ER, with visual sequence confirmation in Geneious, and further confirmation using tblastn revealed that of the 34 additional *Desulfovibrio* genomes considered here, only 2 representatives have a complete set of genes necessary for urea transport and catabolism (*D. desulfuricans* subsp. desulfuricans and *D. fructosivorans*), the remaining organisms lack either a transporter, a urease, or all genes involved in urea transport and catabolism (Table 3).

*D. sp. NAR1* and *D. sp. NAR2* both lacked genes for all proteins discussed above according to an examination in IMG/ER and visual examination of annotations in Geneious. A tblastn search for all genes previously discussed and genes encoding an alternate catabolic enzyme, a urea amidolyase, was performed for both isolates. No homologs for proteins involved in urea transport or catabolism from *D. desulfuricans* subsp. desulfuricans were found in NAR1, however, a homolog of the amino acid/amide transport permease protein UrtC was found in NAR2. The NAR2 amino acid sequence shared 44% homology across 227 amino acid residues to UrtC from *D. desulfuricans* subsp. desulfuricans (locus tag: G449DRAFT_2704). The gene encoding the UrtC homolog in NAR2 is flanked by other genes involved in amino acid transport, specifically branched chain amino acid transport. Since UrtC belongs to a
larger family of amino acid/amide transporters, it is probable that the UrtC homolog in NAR2 has a role in amino acid transport and that the blast similarity to the urea transporter is due to the structural similarity of the substrates that the two enzymes bind. There were no homologs in NAR2 for the other *D. desulfuricans* subsp. desulfuricans proteins involved in urea transport and catabolism, including the urease. No homologs for the UT type urea transporter from *D. vulgaris* Hildenborough or amidolyase from *Rhizobium* sp. was found in either isolate.

Based on these findings, it can be concluded that the *D. sp.* NAR1 draft genome does not contain genes for any of the described enzymatic machinery necessary for urea transport or catabolism. The *D. sp.* NAR2 draft genome does contain a possible urea transporter, but there is no evidence supporting that this isolate has an ability to use urea catabolically. These findings in the genome data agree with what has been observed at the physiological level for both isolates, to the extent that both isolates fix N even in the presence of urea, which supports the conclusion that they cannot catabolize urea. However, the NAR1 isolate has exhibited increased NA in the presence of urea, which would seem to indicate that the isolate has some means of sensing its presence. It would appear that any possible method NAR1 could be employing to sense urea and/or transport it into the cell is not a part of described urea metabolism or transport in bacteria. Since urea metabolism in environmental representatives is currently not very well described, further physiological and molecular investigations are needed to elucidate the mechanism and response seen in NAR1. The fact that neither isolate is predicted to be capable of catabolizing urea does
not make them unique amongst the *Desulfovibrio*, or amongst the eubacteria in general.

**Ammonia metabolism**

Ammonium is known to be the most universally utilized source of biologically available N, and is taken up preferentially by estuarine microbes (89). Accordingly, we would expect to find genes for ammonium uptake and incorporation in all *Desulfovibrio* representatives. An examination of environmental isolates and additional representatives revealed that all *Desulfovibrio*, with the exception of *D. cuneatus*, possess at least one copy of the ammonium transporter (*amt*, TIGR accession: TIGR00836) (Table 3). It is possible that because *D. cuneatus* is a draft, the ammonium transporter was either missed in annotation or is missing from the assembly, and that the organism may in fact have the transporter. The majority of representatives have multiple copies of the ammonium transporter, where both NAR1 and NAR2 have a single copy, making them slightly atypical in this regard. There are, however, 4 additional representatives that also have a single copy of the transporter. A BLAST search did not reveal any additional copies of the ammonium transporter in either isolate, but it is possible that both isolates could have an additional copy/copies of the transporter that are missing from the current assemblies, or that they have a different protein acting as ammonium transporter that has not yet had that function formally assigned to it. The ammonium transporter for NAR1 is located on DESnar1_contig6, at 122,261-123,619bp, in the forward direction. It is immediately followed by a copy of N regulatory protein P-II, two hypothetical proteins, and a copy of glutamate synthase ~3.5kb downstream, an arrangement which makes sense in
terms of nitrogen regulation and activity which has been seen in other bacteria (34, 36, 38). In the NAR2 isolate, the ammonium transporter is located on DESnar2_contig2 at 904,931-905,257bp, in the forward direction. It is immediately preceded by a copy of N regulatory protein P-II, and immediately followed by an isocitrate dehydrogenase and a protein disulfide isomerase. NAR2 does have a copy of glutamate synthase, however it is located on a different contig.

Whether ammonia is used directly from the environment or is derived from other N sources, its assimilation involves metabolites. Some metabolites, such as 2-oxoglutarate, signal N sufficiency or deficiency to the regulatory apparatus (98). To confirm the potential for ammonia uptake and incorporation in both isolates, additional proteins involved in the N assimilatory pathway were examined in both isolates and the additional Desulfovibrio representatives (Table 3). The N regulatory protein P-II is a 2-oxoglutarate (2OG) sensor, which is involved in the adenylation cascade that regulates the activity and concentration of glutamine synthetase (GS), in response to N source availability (99). The majority of the Desulfovibrio genomes examined here, including both environmental isolates, have between 2-4 copies of this protein. NAR1 has 4 copies of the protein, with two copies located side by side in between nifH and nifD on DESnar1_contig13, and a third copy located near the previously mentioned ammonium transporter. The fourth copy is located on DESnar1_contig1 at 248,250-248,591 in the forward direction. It is surrounded on either side by hypothetical proteins, further upstream are proteins involved in cellular respiration and downstream are proteins involved in the shikimate pathway. NAR2 has 3 copies of the N regulatory protein P-II, with two copies located side by side and
preceded immediately by $nifH$ and followed immediately by $nifD$, the same arrangement seen in NAR1. The third copy is located on DESnar2_contig2 and was previously mentioned in relation to the ammonium transporter, which immediately follows this copy of N regulatory protein P-II. The locations of all copies of the P-II protein in both isolates make sense in terms of transcription, regulation, and activity, given the proximity to other genes involved in N metabolism.

Because the majority of prokaryotes possess the glutamine synthetase (GS)/glutamate synthase (GOGAT) pathway of assimilation, both environmental isolates were assessed for components of this pathway. The isolates were also assessed for an alternate pathway involving the NADP-linked glutamate dehydrogenase, which catalyzes the amination of 2-oxoglutarate to form glutamate. NAR1 has two copies of glutamine synthetase, the first copy is annotated as being a type III glutamine synthetase, and the second is annotated as a type I glutamine synthetase, both of which have been found previously in prokaryotes (100, 101). NAR1 also has all subunits for the NADPH type GOGAT and an NAD-specific glutamate dehydrogenase. NAR2 has a typical prokaryotic type I glutamine synthetase, all subunits of the NADPH GOGAT, and a glutamate/leucine dehydrogenase.

Both environmental isolates appear to have a complete GS/GOGAT system of ammonia assimilation, as well as a glutamate dehydrogenase. They both possess the critically important N regulatory protein P-II, with copies of this gene found at genomic locations that make sense in terms of N sensing and regulation. This provides evidence that both isolates have a predicted means of sensing ammonia in the cell, a means of assimilating it, and a means of signaling the regulation of other genes.
involved in N metabolism under differing N conditions. These findings support our physiological observations, in that neither isolate exhibited NA in the presence of ammonia.

**Nitrate metabolism**

No member of the *Desulfovibrio* examined here, including the environmental isolates, has a predicted means of assimilating nitrate (Nas-type nitrate reductases) (Table 3). However, some *Desulfovibrio* species have previously been shown to reduce nitrate or nitrite and use it as a terminal electron acceptor (102, 103), and a sufficient number of the *Desulfovibrio* genomes examined here appear to have the genetic potential for dissimilatory or respiratory nitrate/nitrite reduction. However, nitrate and nitrite reductases share a significant amount of amino acid sequence similarity with sulfate and sulfite reductases, which poses a challenge to automatic gene annotators and increases the likelihood of incorrect annotations, especially when the genomes of sulfate-reducing organisms (which are known to have sulfate and sulfite reductases) are being examined. It is therefore important to approach assertions of nitrate and nitrite reductases in these organisms that are solely based on computational analyses with caution. Additional evidence at the physiological level for organisms predicted to have the ability to reduce nitrate or nitrite would be needed to confirm the computational assertion. Currently the only member of the representatives examined here to have a confirmed, active, respiratory (Nap) type nitrate reductase is *Desulfovibrio desulfuricans* subsp. *desulfuricans* str. ATCC27774 (102, 104). Both environmental isolates have a putative NapC-type respiratory nitrate reductase subunit and NAR2 has a predicted NorZ apoprotein nitric oxide reductase.
The lack of genes involved in nitrate assimilation or assimilatory nitrate reduction appears to be typical regarding the *Desulfovibrio*, and the presence of genes involved in dissimilatory or respiratory nitrate reduction seems to be variable among the group. The lack of genetic potential for utilizing nitrate in an assimilatory manner in both isolates supports what we have seen at the physiological level, as NA in both isolates was unaffected by the presence of nitrate.

**Carbon metabolism**

Both isolates have all proteins necessary for glycolysis and gluconeogenesis. It does appear that both isolates have an incomplete TCA cycle, although this is not unusual amongst the *Desulfovibrio* (105, 106). Additionally, both isolates have genes encoding an L-lactate transporter and an L-lactate dehydrogenase, which supports the observed growth of both isolates on lactate. The presence of a complete glycolytic/gluconeogenic pathway further supports both assemblies as being comprehensive, well annotated drafts.

**Biofilm formation**

The NAR1 isolate has observed biofilm formation, except under conditions where a rich media (NBSO + .05% lysed yeast cell extract) is used and no additional surface for cellular attachment is present (e.g. glass beads, sediment, etc.). This has made it a very challenging organism to work with. Many *Desulfovibrio* are known to form biofilms, including the well-studied *Desulfovibrio vulgaris* Hildenborough (107). However, the genes responsible for biofilm formation in these organisms have not been conclusively identified or well-studied, and cell-to-cell communication and
quorum sensing pathways involved in biofilm formation for these organisms remains unclear.

An examination of NAR1 gene annotations in IMG/ER and Geneious revealed no genes belonging to the well described lux family of quorum sensing genes (108, 109), these results were then confirmed using a tblastn analysis with the assembled NAR1 contigs as the database and Lux proteins from Vibrio fischeri as the queries. Putative genes involved in biofilm formation in NAR1 were ultimately discovered using a filtered evolutionary gene network (87) (Fig. 3.3). The majority of these genes belong to a family of diguanylate cyclases with GGDEF (110) domains, and c-di-GMP receptor domain proteins.

Cyclic di-GMP (c-di-GMP) is a bacterial second messenger that is widely utilized by bacteria, with more than 80% of sequenced bacteria predicted to use this signal (111, 112). C-di-GMP controls a variety of phenotypes, including biofilm formation, motility, and virulence in multiple bacteria (111, 113, 114). The fact that the predicted diguanylate cyclases in NAR1 did not network with proteins from any other closely related Desulfovibrio representatives support the hypothesis that these proteins may serve a unique function in NAR1. Given the documented role of diguanylate cyclases and c-di-GMP in biofilm formation, and that NAR1 is unique in its near obligate biofilm lifestyle when compared with NAR2 and other representatives of the Desulfovibrio, it is possible that the unique function served by these proteins is coordination of biofilm formation in NAR1. These genes and the proteins they code for should be further investigated to confirm any role in biofilm formation in this isolate. A separate examination of c-di-GMP levels in NAR1 and other biofilm-
forming *Desulfovibrio* during biofilm growth and during planktonic growth should be considered in order to elucidate the role of c-di-GMP as a signaling molecule for biofilm formation in these organisms.

**Mercury methylation**

Mercury (Hg) is a pervasive global pollutant known to found in Narragansett Bay sediments (79); in its methylated form (CH$_3$Hg$^+$), it bioaccumulates and is highly toxic to humans and other organisms (115). Unlike inorganic forms of Hg, which originate from atmospheric deposition and point discharge, CH$_3$Hg$^+$ is generated in the environment by microorganisms. Hg methylation is largely restricted to the proteobacteria and primarily to anaerobic organisms (116). Sulfate-reducing bacteria, such as the *Desulfovibrio*, are the main producers of CH$_3$Hg$^+$ (117, 118), although iron-reducing bacteria and methanogens can also be involved (119, 120).

The genetic basis for bacterial mercury methylation was recently described by Parks, et al. (18). Because some of the closest relatives to NAR1 are confirmed Hg methylators and because Narragansett Bay sediments are known to contain Hg, the draft genomes of both environmental isolates were searched for hgcA and hgcB, the genes required for bacterial Hg methylation (18). The amino acid sequences for both Hg methylation proteins from *D. desulfuricans* str. ND132 were used as queries to search the draft genomes of NAR1 and NAR2 using tblastn. No homologs for these proteins were found in the genome of NAR2, however homologs were found in the genome of NAR1 (Fig. 5). A homolog for HgcA was found on DESnar1_contig14 at 52,665-53,621 base pairs in the forward direction with 63% sequence similarity to HgcA from *D. desulfuricans* str. ND132, and a homolog for HgcB was found on
DESnar1_contig14 at 53,697-53,984 base pairs in the forward direction with 64.6% sequence similarity to HgcB from *D. desulfuricans* str. ND132. The alignment of both proteins found in NAR1 against other confirmed (black arrowhead) and predicted Hg-methylators belonging to the *Desulfovibrio* is shown in Fig. 5. The level of amino acid sequence similarity between NAR1 and the other Hg-methylators combined with the fact that the proteins in NAR1 all have highly conserved domains required for the two previously described Hgc proteins to function in Hg methylation (18), provide evidence for NAR1 having the genetic capability to methylate Hg. Further examination at the physiological level would be needed in order to confirm Hg methylation in this isolate and to establish Hg methylation rates. Hg methylation potential in NAR1 and similar bacteria residing in Narragansett Bay sediments is an important consideration, as CH$_3$Hg$^+$ production has the potential to negatively impact other organisms living in the bay.

**Conclusions**

This study combined high-throughput sequencing, comparative genomics, and physiological N fixation rate approaches to further understand the controls on N fixation and N metabolism in members of the *Desulfovibrionaceae*. Our genomic data and physiological N fixation rate data suggest that environmental isolates *Desulfovibrio* NAR1 and *Desulfovibrio* NAR2 are capable of using ammonia as an N source but are not capable of using nitrate or urea to meet cellular N requirements, although there is a possibility that both isolates may be able to respire nitrate. The majority of additional *Desulfovibrio* genomes examined in comparison to our environmental isolates also showed a lack of genes involved urea and nitrate
catabolism, indicating that an inability to utilize these sources of combined N may be widespread amongst the *Desulfovibrionaceae*. This could in turn provide an explanation for why we see *nifH* expression by this group even in environments that are not limited for combined N, such as sediments of the Narragansett Bay estuary (121, 122). It also suggests that we will continue to see input of fixed N by this group of organisms in Narragansett Bay and other similar estuarine systems, despite the presence of combined N, which could potentially exacerbate eutrophication and linked seasonal hypoxic events in these environments.

The sequenced genomes of both isolates were additionally found to be metabolically versatile and unique compared to one another. Both isolates possessed a significant number of genes involved in bacterial defense and heavy metal resistance, which is not surprising for members of this genus, however, their presence may still provide an advantage for survival in Narragansett Bay sediments. Isolates were also shown to possess different physiological traits that were then reflected at the genome level, such as the near obligate biofilm formation exhibited by NAR1 and the complete lack of biofilm formation seen in NAR2. Results of gene network analysis indicated that cyclic-diguanlylate (c-di-GMP) could be an important cell-to-cell signaling molecule involved in biofilm formation in NAR1. This could have implications for future studies involving this isolate or other biofilm forming *Desulfovibrio*, especially since biofilm formation and related cell signaling is not well characterized for members of this genus. NAR1 was also found to have genes involved in bacterial mercury methylation, an environmentally harmful process that is primarily restricted to anaerobic sulfate reducing bacteria (117, 118). Since mercury is
known to be present in Narragansett Bay sediments (79), and changes in Narragansett Bay seem to be shifting towards conditions that favor these sulfate reducers (123), bacterial mercury methylation activity by this diazotrophic community might be an additional aspect to consider in terms of how the biogeochemical roles these organisms play might be linked and how that affects the overall health of Narragansett Bay.

References:

1. Nixon SW, Ammerman JW, Atkinson LP, Berounsky VM, Billen G, Boicourt WC, Boynton WR, Church TM, Ditoro DM, Elmgren R, Garber JH, Giblin AE, Jahnke RA, Owens NJP, Pilson MEQ, Seitzinger SP. 1996. The fate of nitrogen and phosphorus at the land-sea margin of the North Atlantic Ocean. Biogeochemistry 35:141–180.

2. Karl D, Letelier R, Tupas L, Dore J, Christian J, Hebel D. 1997. The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. Nature 388:533–538.

3. Karl D, Michaels A, Bergman B, Capone DG, Carpenter E, Letelier R, Lipshultz F, Paerl HW, Stigman D, Stal L. 2002. Dinitrogen fixation in the world’s oceans. Biogeochemistry 57:47–98.

4. Zehr JP, Jenkins BD, Short SM, Steward GF. 2003. Minireview Nitrogenase gene diversity and microbial community structure: a cross-system comparison. Environ. Microbiol. 5:539–554.

5. Howarth RW, Marino R, Lane J. 1988. Nitrogen fixation in freshwater, estuarine, and marine ecosystems. 1. Rates and importance. Limnol. Oceanogr. 33:669–687.

6. Fulweiler RW, Nixon SW, Buckley BA, Granger SL. 2007. Reversal of the net dinitrogen gas flux in coastal marine sediments. Nature 448:180–182.

7. Gardner W, McCarthy M, An S. 2006. Nitrogen fixation and dissimilatory nitrate reduction to ammonium (DNRA) support nitrogen dynamics in Texas estuaries. Limnol. Oceanogr. 51:558–568.
8. **Ferguson A, Eyre B, Gay J.** 2007. Benthic metabolism and nitrogen cycling in a sub-tropical coastal embayment: spatial and seasonal variation and controlling factors. Aquat. Microb. Ecol. 48:175–195.

9. **Welsh DT, Bourguès S, de Wit R, Herb RA.** 1996. Seasonal variations in nitrogen-fixation (acetylene reduction) and sulphate-reduction rates in the rhizosphere of Zostera noltii: nitrogen fixation by sulphate reducing bacteria. Mar. Biol. 125:619–628.

10. **McGlathery K, Sundbäck K, Anderson I.** 2004. The Importance of Primary Producers for Benthic Nitrogen and Phosphorus Cycling, p. 231–261. In Nielsen, SL, Banta, GM, Pedersen, MF (eds.), Estuarine Nutrient Cycling: The Influence of Primary Producers. Kluwer Academic Publishers.

11. **Halm H, Lam P, Ferdelman TG, Lavik G, Dittmar T, LaRoche J, D’Hondt S, Kuypers MMM.** 2012. Heterotrophic organisms dominate nitrogen fixation in the South Pacific Gyre. ISME J. 6:1238–49.

12. **Farnelid H, Bentzon-Tilia M, Andersson AF, Bertilsson S, Jost G, Labrenz M, Jürgens K, Riemann L.** 2013. Active nitrogen-fixing heterotrophic bacteria at and below the chemocline of the central Baltic Sea. ISME J. 7:1413–23.

13. **Fernandez C, Farías L, Ulloa O.** 2011. Nitrogen fixation in denitrified marine waters. PLoS One 6:e20539.

14. **Fulweiler RW, Brown SM, Nixon SW, Jenkins BD.** 2013. Evidence and a conceptual model for the co-occurrence of nitrogen fixation and denitrification in heterotrophic marine sediments. Mar. Ecol. Prog. Ser. 482:57–68.

15. **Sisler FD, ZoBell CE.** 1951. Nitrogen Fixation by Sulfate-reducing Bacteria Indicated by Nitrogen/Argon Ratios. Science. 113:511–512.

16. **Geets J, Borremans B, Diels L, Springael D, Vangronsveld J, van der Lelie D, Vanbroekhoven K.** 2006. DsrB gene-based DGGE for community and diversity surveys of sulfate-reducing bacteria. J. Microbiol. Methods 66:194–205.

17. **Annachhatre AP, Suktrakoolvait S.** 2001. Biological sulfate reduction using molasses as a carbon source. Water Environ. Res. 73:118–26.

18. **Parks JM, Johs A, Podar M, Bridou R, Hurt R a, Smith SD, Tomanicek SJ, Qian Y, Brown SD, Brandt CC, Palumbo A V, Smith JC, Wall JD, Elias DA, Liang L.** 2013. The genetic basis for bacterial mercury methylation. Science 339:1332–5.
19. **Muyzer G, Stams AJM.** 2008. The ecology and biotechnology of sulphate-reducing bacteria. Nat. Rev. Microbiol. **6**:441–454.

20. **Hulshoff Pol L, Lens PL, Stams AM, Lettinga G.** 1998. Anaerobic treatment of sulphate-rich wastewaters. Biodegradation **9**:213–224.

21. **Martins M, Faleiro ML, Barros RJ, Veríssimo AR, Barreiros MA, Costa MC.** 2009. Characterization and activity studies of highly heavy metal resistant sulphate-reducing bacteria to be used in acid mine drainage decontamination. J. Hazard. Mater. **166**:706–13.

22. **Gutiérrez-Sánchez C, Olea D, Marques M, Fernández VM, Pereira IAC, Vélez M, De Lacey AL.** 2011. Oriented Immobilization of a Membrane-Bound Hydrogenase onto an Electrode for Direct Electron Transfer. Langmuir **27**:6449–6457.

23. **Zehr JP, Mellon M, Braun S, Litaker W, Steppe T, Paerl HW.** 1995. Diversity of heterotrophic nitrogen fixation genes in a marine cyanobacterial mat. Appl. Environ. Microbiol. **61**:2527–32.

24. **Postgate JR, Kent HM, Robson RL.** 1988. Nitrogen fixation by Desulfovibrio, p. 457–471. In Cole, JA, Ferguson, SJ (eds.), The nitrogen and sulphur cycles. Society of General Microbiology Symposium 42. Cambridge University Press.

25. **Capone DG.** 1983. Benthic nitrogen fixation, p. 85–123. In Carpenter, EJ, Capone, DG (eds.), Nitrogen in the Marine Environment. John Wiley & Sons Ltd, New York.

26. **Capone DG.** 1988. Benthic Nitrogen Fixation, p. 105–137. In Blackburn, TH, Sorensen, J (eds.), Nitrogen Cycling in Coastal Marine Environments. Springer, New York.

27. **Carpenter EJ, Capone DG.** 2008. Nitrogen fixation in the marine environment, p. 141–198. In Capone, DG, Bronk, DA, Mulholland, MR, Carpenter, EJ (eds.), Nitrogen in the Marine Environment. Academic Press, Elsevier, San Diego.

28. **Burns JA, Zehr JP, Capone DG.** 2002. Nitrogen-fixing phylotypes of Chesapeake Bay and Neuse River estuary sediments. Microb. Ecol. **44**:336–43.

29. **Bertics V, Sohm J, Treude T, Chow C, Capone D, Fuhrman J, Ziebis W.** 2010. Burrowing deeper into benthic nitrogen cycling: the impact of bioturbation on nitrogen fixation coupled to sulfate reduction. Mar. Ecol. Prog. Ser. **409**:1–15.
30. Bertics VJ, Löscher CR, Salonen I, Dale a. W, Schmitz R a., Treude T. 2012. Occurrence of benthic microbial nitrogen fixation coupled to sulfate reduction in the seasonally hypoxic Eckernförde Bay, Baltic Sea. Biogeosciences Discuss. 9:6489–6533.

31. Herbert RA. 1975. Heterotrophic nitrogen fixation in shallow estuarine sediments. J. Exp. Mar. Bio. Ecol. 18:215–225.

32. Nedwell DB, Azni bin Abdul Aziz S. 1980. Heterotrophic nitrogen fixation in an intertidal saltmarsh sediment. Estuar. Coast. Mar. Sci. 10:699–702.

33. McGlathery K. 1998. Temporal and spatial variation in nitrogen fixation activity in the eelgrass Zostera marina rhizosphere. Mar. Ecol. Prog. Ser. 168:245–258.

34. Heidelberg JF, Seshadri R, Haveman SA, Hemme CL, Paulsen IT, Kolonay JF, Eisen JA, Ward N, Methe B, Brinkac LM, Daugherity SC, Deboy RT, Dodson RJ, Durkin AS, Madupu R, Nelson WC, Sullivan SA, Fouts D, Haft DH, Selengut J, Peterson JD, Davidsen TM, Zafar N, Zhou L, Radune D, Dimitrov G, Hance M, Tran K, Khouri H, Gill J, Utterback TR, Feldblyum TV, Wall JD, Voordouw G, Fraser CM. 2004. The genome sequence of the anaerobic, sulfate-reducing bacterium Desulfovibrio vulgaris Hildenborough. Nat. Biotechnol. 22:554–9.

35. Brown SD, Wall JD, Kucken AM, Gilmour CC, Podar M, Brandt CC, Teshima H, Detter JC, Han CS, Land ML, Lucas S, Han J, Pennacchio L, Nolan M, Pitluck S, Woyke T, Goodwin L, Palumbo A V, Elias D a. 2011. Genome sequence of the mercury-methylating and pleomorphic Desulfovibrio africanus Strain Walvis Bay. J. Bacteriol. 193:4037–8.

36. Brown SD, Gilmour CC, Kucken AM, Wall JD, Elias DA, Brandt CC, Podar M, Chertkov O, Held B, Bruce DC, Detter JC, Tapia R, Han CS, Goodwin LA, Cheng JF, Pitluck S, Woyke T, Mikhailova N, Ivanova NN, Han J, Lucas S, Lapidus AL, Land ML, Hauser LJ, Palumbo AV. 2011. Genome sequence of the mercury-methylating strain Desulfovibrio desulfuricans ND132. J. Bacteriol. 193:2078–2079.

37. Morais-Silva FO, Rezende AM, Pimentel C, Santos CI, Clemente C, Varela-Raposo A, Resende DM, da Silva SM, de Oliveira LM, Matos M, Costa DA, Flores O, Ruiz JC, Rodrigues-Pousada C. 2014. Genome sequence of the model sulfate reducer Desulfovibrio gigas: a comparative analysis within the Desulfovibrio genus. Microbiology open 3:513–530.

38. Pradel N, Ji B, Gimenez G, Talla E, Lenoble P, Garel M, Tamburini C, Fourquet P, Lebrun R, Bertin P, Denis Y, Pophillat M, Barbe V, Ollivier B, Dolla A. 2013. The first genomic and proteomic characterization of a deep-sea
sulfate reducer: insights into the piezophilic lifestyle of Desulfovibrio piezophilus. PLoS One 8:e55130.

39. **Pedersen K, Bengtsson A, Edlund J, Rabe L, Hazen T, Chakraborty R, Goodwin L, Shapiro N.** 2014. Complete Genome Sequence of the Subsurface, Mesophilic Sulfate-Reducing Bacterium Desulfovibrio aespoeensis Aspo-2. Genome Announc. 2:3–4.

40. **Ji B, Gimenez G, Barbe V, Vacherie B, Rouy Z, Amrani A, Fardeau M-L, Bertin P, Alazard D, Leroy S, Talla E, Ollivier B, Dolla A, Pradel N.** 2013. Complete Genome Sequence of the Piezophilic, Mesophilic, Sulfate-Reducing Bacterium Desulfovibrio hydrothermalis AM13(T.). Genome Announc. 1:2–3.

41. **Hauser LJ, Land ML, Brown SD, Larimer F, Keller KL, Rapp-Giles BJ, Price MN, Lin M, Bruce DC, Detter JC, Tapia R, Han CS, Goodwin LA, Cheng J-F, Pitluck S, Copeland A, Lucas S, Nolan M, Lapidus AL, Palumbo A V, Wall JD.** 2011. Complete genome sequence and updated annotation of Desulfovibrio alaskensis G20. J. Bacteriol. 193:4268–4269.

42. **Nakazawa H, Arakaki A, Narita-Yamada S, Yashiro I, Jinno K, Aoki N, Tsuruyama A, Okamura Y, Tanikawa S, Fujita N, Takeyama H, Matsunaga T.** 2009. Whole genome sequence of Desulfovibrio magneticus strain RS-1 revealed common gene clusters in magnetotactic bacteria. Genome Res. 19:1801–1808.

43. **Hungate R.** 1969. A roll tube method for the cultivation of strict anaerobes, p. 117–132. In Norris, J, Ribbons, D (eds.), Methods in Microbiology. Academic Press, New York, New York.

44. **Capone D.** 1993. Determination of Nitrogenase Activity in Aquatic Samples Using the Acetylene Reduction Procedure, p. In Kemp, P, Cole, J, Sherr, B, Sherr, E (eds.), Handbook of methods in aquatic microbial ecology. Lewis Publ: Boca Raton, FL.

45. **Lane D.** 1991. 16S/23S rRNA sequencing, p. 115–175. In Stackebrandt, E, Goodfellow, M (eds.), Nucleic acid techniques in bacterial systematics. John Wiley & Sons Ltd, New York.

46. **Bolger AM, Lohse M, Usadel B.** 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

47. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham SON, Prijibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev M, Pevzner PA.** 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19:455–477.
48. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–5.

49. Hunt M, Kikuchi T, Sanders M, Newbold C, Berriman M, Otto TD. 2013. REAPR: a universal tool for genome assembly evaluation. Genome Biol. 14:R47.

50. Markowitz VM, Chen I-MA, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob B, Huang J, Williams P, Huntemann M, Anderson I, Mavromatis K, Ivanova NN, Kyrpides NC. 2012. IMG: the Integrated Microbial Genomes database and comparative analysis system. Nucleic Acids Res. 40:D115–D122.

51. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119.

52. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30:772–80.

53. Vaidya G, Lohman DJ, Meier R. 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27:171–180.

54. Posada D. 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25:1253–6.

55. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61:539–42.

56. Halary S, McInerney JO, Lopez P, Bapteste E. 2013. EGN: a wizard for construction of gene and genome similarity networks. BMC Evol. Biol. 13:146.

57. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5:e11147.

58. Postgate JR, Campbell LL. 1966. Classification of Desulfovibrio species, the nonsporulating sulfate-reducing bacteria. Bacteriol. Rev. 30:732–8.

59. Bentzon-Tilia M. 2014. The Biology of Heterotrophic N2-fixing Bacteria in Marine and Estuarine Waters. PhD dissertation, University of Copenhagen.
60. Beech IB, Sunner JA, Hiraoka K. 2005. Microbe-surface interactions in biofouling and biocorrosion processes. Int. Microbiol. 8:157–68.

61. Clark ME, Edelmann RE, Duley ML, Wall JD, Fields MW. 2007. Biofilm formation in Desulfovibrio vulgaris Hildenborough is dependent upon protein filaments. Environ. Microbiol. 9:2844–54.

62. Mackerras AH, Smith GD. 1986. Evidence for direct repression of nitrogenase by ammonia in the cyanobacterium Anabaena cylindrica. Biochem. Biophys. Res. Commun. 134:835–44.

63. Moseman-Valtierra SM, Armaiz-Nolla K, Levin LA. 2009. Wetland response to sedimentation and nitrogen loading: diversification and inhibition of nitrogen-fixing microbes. Ecol. Appl. 20:1556–1568.

64. Ewing B, Hillier L, Wendl MC, Green P. 1998. Base-Calling of Automated Sequencer Traces Using Phred. I. Accuracy Assessment. Genome Res. 8:175–185.

65. Ewing B, Green P. 1998. Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities. Genome Res. 8:186–194.

66. Magoc T, Pabinger S, Canzar S, Liu X, Su Q, Puiu D, Tallon LJ, Salzberg SL. 2013. GAGE-B: an evaluation of genome assemblers for bacterial organisms. Bioinformatics 29:1718–25.

67. Marcelletti S, Ferrante P, Petriccione M, Firrao G, Scortichini M. 2011. Pseudomonas syringae pv. actinidiae draft genomes comparison reveal strain-specific features involved in adaptation and virulence to Actinidia species. PLoS One 6:e27297.

68. Tatusova T, Ciufo S, Fedorov B, O’Neill K, Tolstoy I. 2014. RefSeq microbial genomes database: new representation and annotation strategy. Nucleic Acids Res. 42:D553–D559.

69. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards R a, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res. 42:D206–14.

70. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–10.

71. Heler R, Marraffini L, Bikard D. 2014. Adapting to new threats: the generation of memory by CRISPR Cas immune systems. Mol. Microbiol. Published.
72. **Grissa I, Vergnaud G, Pourcel C.** 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinformatics **8**:172.

73. **Kanehisa M, Goto S.** 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. **28**:27–30.

74. **Tatusov RL, Galperin MY, Natale DA, Koonin E V.** 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. **28**:33–6.

75. **Bateman A.** 2002. The Pfam Protein Families Database. Nucleic Acids Res. **30**:276–280.

76. **Haft DH.** 2003. The TIGRFAMs database of protein families. Nucleic Acids Res. **31**:371–373.

77. **Piddock LJ V.** 2006. Multidrug-resistance efflux pumps - not just for resistance. Nat. Rev. Microbiol. **4**:629–36.

78. **Rincón CE.** 2006. Surface Sediment Physical and Chemical Gradients of Upper Narragansett Bay, Rhode Island: A Close Look at Modern Metal Distributions. M.S. thesis, Brown University.

79. **Taylor DL, Linehan JC, Murray DW, Prell WL.** 2012. Indicators of sediment and biotic mercury contamination in a southern New England estuary. Mar. Pollut. Bull. **64**:807–19.

80. **Kutcher T.** 2009. Chapter 12-Human Impacts on Narragansett Bay, pg. 147-162. In Raposa, KB, Schwartz, ML (eds.), An Ecological Profile of the Narragansett Bay National Estuarine Reserve (nbnerr.org/profile.htm). Prudence Island, RI.

81. **Baquero F, Martínez J-L, Cantón R.** 2008. Antibiotics and antibiotic resistance in water environments. Curr. Opin. Biotechnol. **19**:260–5.

82. **Xi C, Zhang Y, Marrs CF, Ye W, Simon C, Foxman B, Nriagu J.** 2009. Prevalence of antibiotic resistance in drinking water treatment and distribution systems. Appl. Environ. Microbiol. **75**:5714–8.

83. **Lu S-Y, Zhang Y-L, Geng S-N, Li T-Y, Ye Z-M, Zhang D-S, Zou F, Zhou H-W.** 2010. High diversity of extended-spectrum beta-lactamase-producing bacteria in an urban river sediment habitat. Appl. Environ. Microbiol. **76**:5972–6.
84. Goldstein EJC, Citron DM, Peraino VA, Cross SA. 2003. Desulfovibrio desulfuricans Bacteremia and Review of Human Desulfovibrio Infections. J. Clin. Microbiol. 41: 2752-2754.

85. Lozniewski A, Maurer P, Schuhmacher H, Carlier JP, Mory F. 1999. First isolation of Desulfovibrio species as part of a polymicrobial infection from a brain abscess. Eur. J. Clin. Microbiol. Infect. Dis. 18:602–3.

86. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2013. GenBank. Nucleic Acids Res. 41:D36–42.

87. Halary S, Leigh J, Cheaib B, Lopez P, Bapteste E. 2010. Network analyses structure genetic diversity in independent genetic worlds. Proc. Natl. Acad. Sci. 107:127–132.

88. Solomon C, Collier J, Berg G, Glibert P. 2010. Role of urea in microbial metabolism in aquatic systems: a biochemical and molecular review. Aquat. Microb. Ecol. 59:67–88.

89. Jorgensen NOG. 2006. Uptake of urea by estuarine bacteria. Aquat. Microb. Ecol. 42:227–242.

90. Valladares A, Montesinos ML, Herrero A, Flores E. 2002. An ABC-type, high-affinity urea permease identified in cyanobacteria. Mol. Microbiol. 43:703–15.

91. Kojima S, Bohner A, von Wirén N. 2006. Molecular mechanisms of urea transport in plants. J. Membr. Biol. 212:83–91.

92. Weeks DL. 2000. A H+-Gated Urea Channel: The Link Between Helicobacter pylori Urease and Gastric Colonization. Science. 287:482–485.

93. Sebbane F, Bury-Moné S, Cailliau K, Browaeys-Poly E, De Reuse H, Simonet M. 2002. The Yersinia pseudotuberculosis Yut protein, a new type of urea transporter homologous to eukaryotic channels and functionally interchangeable in vitro with the Helicobacter pylori UreI protein. Mol. Microbiol. 45:1165–74.

94. You G, Smith CP, Kanai Y, Lee WS, Stelzner M, Hediger MA. 1993. Cloning and characterization of the vasopressin-regulated urea transporter. Nature 365:844–7.

95. Raunser S, Mathai JC, Abeyrathne PD, Rice AJ, Zeidel ML, Walz T. 2009. Oligomeric structure and functional characterization of the urea transporter from Actinobacillus pleuropneumoniae. J. Mol. Biol. 387:619–27.
96. **Maciver B, Smith CP, Hill WG, Zeidel ML.** 2008. Functional characterization of mouse urea transporters UT-A2 and UT-A3 expressed in purified Xenopus laevis oocyte plasma membranes. Am. J. Physiol. Renal Physiol. 294:F956–64.

97. **Levin EJ, Quick M, Zhou M.** 2010. Crystal structure of a bacterial homolog of the kidney urea transporter. Nature 462:757–761.

98. **Leigh JA, Dodsworth JA.** 2007. Nitrogen regulation in bacteria and archaee. Annu. Rev. Microbiol. 61:349–377.

99. **Durand A, Merrick M.** 2006. In vitro analysis of the Escherichia coli AmtB-GlnK complex reveals a stoichiometric interaction and sensitivity to ATP and 2-oxoglutarate. J. Biol. Chem. 281:29558–67.

100. **Brown JR, Masuchi Y, Robb FT, Doolittle WF.** 1994. Evolutionary relationships of bacterial and archaeal glutamine synthetase genes. J. Mol. Evol. 38:566–76.

101. **Van Rooyen JM, Abratt VR, Sewell BT.** 2006. Three-dimensional structure of a type III glutamine synthetase by single-particle reconstruction. J. Mol. Biol. 361:796–810.

102. **Dalsgaard T, Bak F.** 1994. Nitrate Reduction in a Sulfate-Reducing Sulfate-Reducing Bacterium, Desulfovibrio desulfuricans, Isolated from Rice Paddy Soil: Sulfide Inhibition, Kinetics, and Regulation Nitrate Reduction in. Appl. Environ. Microbiol. 60:291–297.

103. **Seitz HJ, Cypionka H.** 1986. Chemolithotrophic growth of Desulfovibrio desulfuricans with hydrogen coupled to ammonification of nitrate or nitrite. Arch. Microbiol. 146:63–67.

104. **Marietou A, Richardson D, Cole J.** 2005. Nitrate reduction by Desulfovibrio desulfuricans: A periplasmic nitrate reductase system that lacks NapB, but includes a unique tetraheme c-type cytochrome, NapM. FEMS Microbiol. Lett. 248:217–225.

105. **Tang Y, Pingitore F, Mukhopadhyay A, Phan R, Hazen TC, Keasling JD.** 2007. Pathway confirmation and flux analysis of central metabolic pathways in Desulfovibrio vulgaris hildenborough using gas chromatography-mass spectrometry and Fourier transform-ion cyclotron resonance mass spectrometry. J. Bacteriol. 189:940–9.

106. **Lewis A, Miller J.** 1975. Keto Acid Metabolism in Desulfovibrio. J. Gen. Microbiol. 90:286–292.
107. Clark M, He Z, Redding A. 2012. Transcriptomic and proteomic analyses of Desulfovibrio vulgaris biofilms: carbon and energy flow contribute to the distinct biofilm growth state. BMC Genomics 13:138.

108. Meighen EA. 1993. Bacterial bioluminescence: organization, regulation, and application of the lux genes. FASEB J. 7:1016–1022.

109. Visick KL, Fuqua C. 2005. Decoding microbial chatter: Cell-cell communication in bacteria. J. Bacteriol. 187:5507–5519.

110. Pei J, Grishin N V. 2001. GGDEF domain is homologous to adenylyl cyclase. Proteins Struct. Funct. Genet. 42:210–216.

111. Römling U, Galperin MY, Gomelsky M. 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiol. Mol. Biol. Rev. 77:1–52.

112. Seshasayee ASN, Fraser GM, Luscombe NM. 2010. Comparative genomics of cyclic-di-GMP signalling in bacteria: Post-translational regulation and catalytic activity. Nucleic Acids Res. 38:5970–5981.

113. Simm R, Morr M, Kader A, Nimtz M, Römling U. 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessibility to motility. Mol. Microbiol. 53:1123–1134.

114. Newell PD, Monds RD, O’Toole GA. 2009. LapD is a bis-(3’,5’)-cyclic dimeric GMP-binding protein that regulates surface attachment by Pseudomonas fluorescens Pf0-1. Proc. Natl. Acad. Sci. U. S. A. 106:3461–6.

115. Wood JM. 1974. Biological Cycles for Toxic Elements in the Environment. Science (80-. ). 183:1049–1052.

116. Hintelmann H. 2010. Organomercurials. Their formation and pathways in the environment. Met. Ions Life Sci. 7:365–401.

117. Compeau GC, Bartha R. 1985. Sulfate-reducing bacteria: principal methylators of mercury in anoxic estuarine sediment. Appl. Environ. Microbiol. 50:498–502.

118. Gilmour CC, Henry EA, Mitchell R. 1992. Sulfate stimulation of mercury methylation in freshwater sediments. Environ. Sci. Technol. 26:2281–2287.

119. Fleming EJ, Mack EE, Green PG, Nelson DC. 2006. Mercury methylation from unexpected sources: molybdate-inhibited freshwater sediments and an iron-reducing bacterium. Appl. Environ. Microbiol. 72:457–64.
120. **Wood JM, Kennedy FS, Rosen CG.** 1968. Synthesis of methyl-mercury compounds by extracts of a methanogenic bacterium. Nature **220**:173–4.

121. **Pryor D, Saarman E, Murray D, Prell W.** 2007. Nitrogen Loading from Wastewater Treatment Plants to Upper Narragansett Bay. Narragansett Bay Estuary Program Report NBEP-2007-126. (http://www.nbep.org/publications/NBEP-126-FNL-Pryor.pdf). Brown University, Providence, RI.

122. **DiMilla PA, Nixon SW, Oczkowski AJ, Altabet MA, McKinney RA.** 2011. Some challenges of an “upside down” nitrogen budget--science and management in Greenwich Bay, RI (USA). Mar. Pollut. Bull. **62**:672–80.

123. **Ehrlich A.** 2014. Elucidating the Molecular Response of Microbial Nitrogen Fixation in Estuarine Sediments to Hypoxia. M.S. thesis, University of Rhode Island.

124. **Hunter JL, Severin GB, Koestler BJ, Waters CM.** 2014. The Vibrio cholerae diguanylate cyclase VCA0965 has an AGDEF active site and synthesizes cyclic di-GMP. BMC Microbiol. **14**:22.

125. **Bellows LE, Koestler BJ, Karaba SM, Waters CM, Lathem WW.** 2012. Hfq-dependent, co-ordinate control of cyclic diguanylate synthesis and catabolism in the plague pathogen Yersinia pestis. Mol. Microbiol. **86**:661–674.
Figure 1. Multigene Bayesian phylogenetic tree. Phylogenetic tree inferred from a concatenated alignment of 20 genes using MrBayes v. 3.2.1. Posterior probabilities are shown at each node. Positions of environmental isolates Desulfovibrio sp. NAR1 and Desulfovibrio sp. NAR2 are indicated with black arrowheads.
Figure 2. Whole genome comparisons. Whole genome comparisons of both environmental isolates and their closest relatives using Mauve, version 2.3.1. The pairwise alignment of *Desulfovibrio* sp. NAR1 against *Desulfovibrio piezophilus* (A) and *Desulfovibrio* sp. NAR2 against *Desulfovibrio alaskensis* str. G20 (B) is shown. The similarly colored blocks are presumably homologous and internally free of rearrangements between genomes. White areas within blocks indicate sequences that were not aligned to the other genome and represent non-homologous regions. Long red lines indicate contig boundaries for draft genomes.
Figure 3.1 Representative protein gene networks consisting of environmental isolates and other Desulfovibrio representatives. (A) NifH protein network. (B) NifA protein network. (C) NifK protein network.

Figure 3.2 Representative protein gene networks consisting of exclusively NAR1 and NAR2. (A) Choline/glycine/betaine transporters. (B) Metal dependent hydrolases of the beta-lactamase superfamily. (C) Glyoxalase family proteins-glyoxalase I.

Figure 3.3 Representative protein gene networks consisting of only NAR1. (A) Putative sensory box/GGDEF family proteins-role in signaling and biofilm persistence. (B) Putative diguanylate cyclase and receptor proteins (GGDEF family). (C) Periplasmic binding and signal transduction proteins.

Figure 3.4 All protein gene networks consisting of only NAR2. (A) Non-specific predicted membrane proteins. (B) Bacteriophage head to tail connecting proteins.
Figure 4. N fixation gene clusters. N fixation gene clusters of Narragansett Bay isolates (black arrows) and three closely related *Desulfovibrio* representatives. Color-coded arrows indicate coding sequences (CDS) and their orientation.
Figure 5. Proteins required for Hg methylation. Multiple sequence alignments of 5 Desulfovibrio HgcA and HgcB orthologs, including environmental isolate NAR1. Confirmed mercury methylating organisms are indicated with a black arrowhead. Red boxes indicate highly conserved regions previously described by Parks et al. (Parks et al., 2013), including the putative cap helix (consensus sequence motif, N(V/I)WCA(A/G)GK in HgcA, two strictly conserved CX2CX2CX3C motifs characteristic of [4Fe-4S] clusters, and a conserved vicinal pair of cysteines located at the C terminus of HgcB.
Figure 6. Growth of NAR1 under N limitation. Growth of *Desulfovibrio* sp. NAR1 under N limited conditions. Error bars represent one standard deviation from sample mean. For the first 3 time-points, error bars are too small to be seen.
Nitrogenase activity of NAR1 under N limitation.

Nitrogenase activity of NAR1 grown under N limited conditions. Reported as nmols C₂H₄ produced per cell per day, measured using the acetylene reduction assay. Error bars represent one standard deviation from sample mean.
Figure 8. Cell enumeration of NAR1 cultures grown under different N treatments. Total cell counts as genomes per culture tube for cultures of NAR1 grown with N limited, 12mM ammonia, 12mM nitrate, and 12mM urea treatments. Error bars represent one standard deviation from sample mean.
Figure 9. Nitrogenase activity of NAR1 under different combined N conditions. Nitrogenase activity of NAR1 grown under different combined N conditions: N limited, 12mM ammonia, 12mM nitrate, and 12mM urea. Nitrogenase activity is reported as nmols C\textsubscript{2}H\textsubscript{4} produced per cell per day, measured using the acetylene reduction assay. Error bars represent one standard deviation from sample mean.
### Table 1. Assembly statistics.

Assembly statistics for the draft genomes of both environmental isolates. “% Reads mapped” refers to the number of reads that mapped back to the final genome assembly.

| Assembly            | Bases    | Contigs | Mean contig length | Longest contig | N50   | N75   | Average contig coverage | % Reads mapped |
|---------------------|----------|---------|--------------------|----------------|-------|-------|--------------------------|----------------|
| Desulfovibrio sp. NAR1 | 3502190  | 17      | 205643             | 627453         | 35371 | 249450 | 691.6                   | 99.8           |
| Desulfovibrio sp. NAR2 | 3907388  | 7       | 558198             | 1879556        | 1122917 | 1122917 | 942.1                   | 99.8           |
| Organism                        | Total bases | GC % | Coding bases % | Genes | CRISPR | PCG % | RCG % | tRNA genes % | rRNA genes % | other RNA genes % | Function PCG % | Hypothetical PCG % |
|--------------------------------|-------------|------|----------------|-------|--------|-------|-------|--------------|--------------|-------------------|----------------|--------------------|
| *D. sp. NAR1 (D)*              | 3508007     | 55.8 | 88.64          | 3279  | 2      | 98.08 | 1.92  | 0.15         | 1.56         | 0.21              | 80.82          | 17.26              |
| *D. sp. NAR2 (D)*              | 3907388     | 44.9 | 83.97          | 3474  | 0      | 96.69 | 3.31  | 0.29         | 2.88         | 0.14              | 78.27          | 18.42              |
| *D. desulfuricans ND132*       | 3858983     | 65.2 | 88.93          | 3534  | 1      | 98.19 | 1.81  | 17           | 1.56         | 0.08              | 79.2           | 18.99              |
| *D. piezophilus*               | 3646098     | 49.9 | 89.47          | 3431  | 0      | 97.76 | 2.24  | 0.26         | 1.69         | 0.29              | 67.3           | 30.46              |
| *D. desulfuricans subsp. aestuarii (D)* | 3551364   | 45.1 | 85.63          | 3251  | 0      | 96.55 | 3.45  | 0.43         | 2.86         | 0.15              | 77.39          | 19.16              |
| *D. alaskensis G20*            | 3730232     | 57.8 | 91.13          | 3874  | 1      | 97.68 | 2.32  | 0.31         | 1.7          | 0.31              | 59.42          | 38.26              |
| *D. vulgaris Miyazaki F*       | 4040304     | 67.1 | 83.35          | 3281  | 1      | 97.68 | 2.32  | 0.37         | 1.95         | 0                 | 72.66          | 25.02              |

**Table 2. Comparative genome statistics.** Comparative annotated genome statistics for *Desulfovibrio* sp. NAR1, *Desulfovibrio* sp. NAR2, and four of the most closely related *Desulfovibrio* genomes. The genome of *D. vulgaris* Miyazaki F was included as an additional point of reference, as neither isolate was very closely related to it. Isolate genome statistics are indicated by the bold text, and (D) indicates a draft genome.
Table 3. Comparison of gene presence involved in N metabolism. Comparative presence or absence of genes involved in N fixation, urea transport and catabolism, ammonium uptake, cellular N signaling, and nitrate metabolism for all *Desulfovibrio* genomes examined in this study. (*) Indicates the presence of one or more plasmids for that organism, numbers indicate the gene copy number for genes listed to the left, and (h) indicates that a homolog was found by hand. Genes were only tabulated if at least one representative was predicted to have that gene, in order to conserve space. In addition to what is shown here all representatives were assessed for the presence of alternate nitrogenases (*anf* and *vnf* gene families), urea amidolyases, assimilatory nitrate reductases (*nas* gene cluster), membrane-bound respiratory nitrate reductases (*nar* gene cluster), and periplasmic dissimilatory nitrate reductases (*nap* gene cluster), respiratory NO forming nitrite reductase (*nirS/nirK*), dissimilatory cytoplasmic NH$_4^+$ forming nitrite reductase (*nirBD*) and nitric oxide reductases (*nor* gene cluster).
| Organism                  | D. Langi | D. magnetotacticum | D. af. | D. af. A2 | D. af. TD | D. af. RA | D. af. DD | D. af. RD | D. af. RA2 | D. af. RA1 | D. af. RA1 | D. af. RA1 |
|--------------------------|----------|--------------------|--------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| D. sp. NAR1              | 1        | 1                  | 1      | 3         | -         | -         | -         | -         | -         | -         | -         | -         |
| D. sp. NAR2              | 1        | 1                  | 1      | 2         | 1         | 2         | 2         | 1         | 1         | 1         | 1         | 1         |
| D. africanus             | 1        | 1                  | 1      | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. aespoeensis           | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. alaskensis            | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. alcoholivorans        | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. alkalitolerans        | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. cf. magneticus        | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. cuneatus              | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. desulfuricans         | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. frigidus              | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. fructosovorans subsp. | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. gigas                 | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. hydrothermalis        | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. longus                | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. magneticus            | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. oxyclinae             | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. piezophilus           | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. putealis              | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. salexigens            | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. sp. A2                | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. sp. FW1012B*          | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. sp. U5L               | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. vulgaris str. DP4*    | 1        | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. vulgaris str. Hildenborough* | 1 | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |
| D. vulgaris str. Miyazaki F | 1 | 1                  | 1      | 2         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |

78
| Predicted nitrate/nitrite transport | MFS transporter, NNP family, nitrate/nitrite | Predicted dissimilatory nitrate reduction | Predicted respiratory nitrite reduction | Predicted respiratory nitric oxide reduction |
|-----------------------------------|---------------------------------------------|------------------------------------------|----------------------------------------|--------------------------------------------|
|                                   |                                              | napA | napC | napD | napE | napG | napH | napM |
| ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) |
|                                   | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) |
| ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) |
| ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) |
| ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) |
| ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) |
| ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) |
| ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) |
| ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) |
| ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) |
| ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) |
| ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) |
| ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) |
| ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) | ![image](https://via.placeholder.com/150) |
Supplemental Protocol 1

Complete DNA Extraction Protocol

Reagents Needed

1. Qiagen DNeasy Minikit for Plant Tissue
2. Phenol/chloroform/isoamyl alcohol (PCl) solution (25:24:1), molecular biology grade, pH 7.8-8.2
3. Chloroform/isoamyl alcohol (24:1), molecular biology grade
4. Elution Buffer (10mM Tris-HCl, pH 8.5)
5. Ammonium acetate, concentrated solution
6. 100% Ethanol
7. 70% Ethanol

Part 1-Modified Qiagen DNeasy Minikit for Plant Tissue

Using filter membrane or ~.30 g glass beads as starting materia

1. Add 400uL buffer AP1 to starting material in 2mL microcentrifuge tube.
2. If extracting from a filter, add .5um and .1um silica beads to each tube. If extracting from beads, add only .1um silica beads.
3. Add 4 uLs RNAse A to each tube
4. Vortex to mix
5. Incubate at 65°C for 10min, mixing 2-3 times during incubation by inverting tubes.
6. Bead beat each tube for 1 minute
7. Add 130uL of Buffer AP2 to lysate, vortex to mix
8. Incubate at 4°C for 5 minutes
9. Centrifuge lysate for minutes at 14,000 rpm
10. If filter is present, remove. If not, continue.

Part 2-Phenol/Chloroform/Isomyl Alcohol Extraction

Expect ~400uL of supernatant from step 10, split into two 2mL microcentrifuge tubes of 200uL each. If supernatant is less than 400uL, aliquot 200uL to one tube and bring up the volume of the 2nd tube to 200uL using elution buffer (10mM Tris-HCl, pH 8.5).

11. Add 200uL of phenol/chloroform/isoamyl alcohol solution (25:24:1) to tubes containing the supernatant.
12. Bead beat tubes 1min
13. Centrifuge tubes at 15,000 rpm for 5 minutes
14. Remove ~180uL of the top aqueous layer and place into a new 2mL microcentrifuge tube. Avoid picking up any of the phenol/chloroform/isoamyl alcohol phase. Set aside.
15. Add 200uL of EB to the first tube, still containing the phenol/chloroform isoamyl alcohol.
16. Repeat steps 12 and 13
17. Remove as much of the top aqueous layer as possible and add it to what is already in the second tube. Discard first tube.

**Part 3-Chloroform Back Extraction**

18. Add equal volumes (~800uL) chloroform/isoamyl alcohol solution (24:1) to second tube containing the aqueous layer
19. Bead beat 1min
20. Centrifuge at 15,000 for 5 minutes
21. Remove as much of the top aqueous solution as possible and place into a new 2mL microcentrifuge tube. Avoid picking up any of the chloroform/isoamyl alcohol phase.

**Part 4-Ethanol Precipitation**

At this stage samples originating from the same source, or samples that needed to be split in earlier stages because of large volumes, can be pooled together.

22. Combine any samples needing to be pooled into a new centrifuge tube
23. Add 0.1xvol 5M ammonium acetate
24. Add 2.5xvol 100% EtOH
25. Precipitate overnight at -20 °C
26. Centrifuge at 4 °C at maximum speed for 1hr
27. Decant supernatant
28. Add 750mL 70% EtOH
29. Centrifuge 5 min at maximum speed and pour off EtOH
30. Add 750mL 100% EtOH
31. Centrifuge 5 min at maximum speed and pipette out EtOH. Be careful not to disturb pellet (may or may not be visible)
32. Dry pellet 4-5 min
33. Elute DNA in 20-100uL of elution buffer or PCR-grade water, depending on desired DNA concentration

If contaminant EPS are still present, Part 4 may be repeated
Supplemental Protocol 2

Preparation of Anaerobic Glycerol Stocks

Reagents needed

Glycerol Solution
30% Glycerol
69.8% ddH₂O
.2% of 0.2% Resazurin solution

Preparation of Glycerol Vials

1. Prepare glycerol solution in a gas tight serum bottle. Place glycerol solution in 50°C water bath for 5 minutes and let cool while gassing with N₂. Flush and evacuate bottle.

2. Flush and evacuate all 10 ml serum vials to be used with N₂.

3. Transfer 1 ml of glycerol solution to each vial via syringe.

4. Flush and evacuate all vials repeatedly.

5. Autoclave 30 min.

Freezing Cells

1. Grow up a fresh culture of bacteria to high density.

2. Add a drop or two of an appropriate reducing agent (such as sodium sulfide) to vials until the resazurin turns colorless. Wrap label tape completely around vial.

3. Using aseptic technique, add 1 ml of culture by syringe to each vial. Quickly place vial into a dry ice/ethanol bath or liquid nitrogen.

4. Store frozen culture stocks at -80°C.
Figure S1. Calculations Used for Biomass Estimates

\[ C_T = G_T = \frac{DNA_{CT}}{G_M} \]

Where:

1. \( DNA_{TE} = \) Total Extracted DNA = DNA concentration of sample \( \times \) volume of sample
2. \( DNA_{GB} = \) DNA per gram of glass beads = \( DNA_{TE} \times \) mass of glass beads extracted
3. \( DNA_{CT} = \) DNA per culture tube = \( DNA_{GB} \times \) grams of glass beads in culture tube
4. \( Mass_{NP} = \) Average mass of nucleotide pair = 650 Daltons (g/mol) \( \times \) 1.1 \( \times \) 10\(^{-12}\) ng
5. \( G_S = \) Size of \( Desulfovibrio \) sp. NAR1 Genome = \( \sim \) 3,500,000 base pairs
6. \( G_M = \) Mass of genome = \( Mass_{NP} \times G_S = 0.00000385\) ng
7. \( G_T = \) Genomes per tube = \( DNA_{CT} \div G_M \)
8. \( C_T = \) Total cells = \( G_T \)

1. New England Biolabs (https://www.neb.com/tools-and-resources/usage-guidelines/nucleic-acid-data)
| Contig          | length | average coverage | total reads | reads in pairs | single reads |
|----------------|--------|------------------|-------------|----------------|--------------|
| DESnar1_contig1 | 62745.3| 649.6            | 218695.9    | 871350.0       | 1315609.0    |
| DESnar1_contig2 | 45609.6| 630.8            | 152770.5    | 611490.0       | 916215.0     |
| DESnar1_contig3 | 16985.7| 649.2            | 588479.0    | 234526.0       | 353953.0     |
| DESnar1_contig4 | 9397.0  | 632.8            | 32374.0     | 12612.0        | 19762.0      |
| DESnar1_contig5 | 35371.2| 755.5            | 1426648.0   | 574344.0       | 852304.0     |
| DESnar1_contig6 | 32778.8| 734.1            | 1308218.0   | 517404.0       | 790814.0     |
| DESnar1_contig7 | 13341.5| 815.6            | 592379.0    | 234126.0       | 358253.0     |
| DESnar1_contig8 | 20195.0| 777.3            | 86064.0     | 33994.0        | 52070.0      |
| DESnar1_contig9 | 12754.0| 909.1            | 625948.0    | 248282.0       | 377666.0     |
| DESnar1_contig10| 34007.1| 764.1            | 1408497.0   | 559778.0       | 848719.0     |
| DESnar1_contig11| 10300.6| 766.5            | 429425.0    | 170686.0       | 258739.0     |
| DESnar1_contig12| 8343.0 | 710.0            | 32539.0     | 13058.0        | 19481.0      |
| DESnar1_contig13| 36296.3| 725.4            | 1418561.0   | 563138.0       | 855423.0     |
| DESnar1_contig14| 24888.0| 665.6            | 890749.0    | 354300.0       | 536449.0     |
| DESnar1_contig15| 20442.9| 652.1            | 719542.0    | 282760.0       | 436782.0     |
| DESnar1_contig16| 1207.0 | 339.2            | 2258.0      | 1106.0         | 1152.0       |
| DESnar1_contig17| 1581.0 | 579.7            | 4998.0      | 2090.0         | 2908.0       |

Table S1. Contig Statistics NAR1. Table displaying statistics for individual contigs of the final draft genome assembly of *Desulfovibrio* sp. NAR1. The last three columns refer to reads that mapped back to the respective contigs.
**Table S2. Contig statistics NAR2.** Table displaying statistics for individual contigs of the final draft genome assembly of *Desulfovibrio* sp. NAR2. The last three columns refer to reads that mapped back to the respective contigs.

| Contig         | length  | average coverage | total reads | reads in pairs | single reads |
|----------------|---------|------------------|-------------|----------------|--------------|
| DESnar2_contig1| 23563.0 | 858.2            | 104701.0    | 102634.0       | 2067.0       |
| DESnar2_contig2| 1122304.0 | 788.2           | 4619612.0    | 4555582.0      | 64030.0      |
| DESnar2_contig3| 1879303.0 | 705.1           | 6896898.0    | 6803450.0      | 93448.0      |
| DESnar2_contig4| 669951.0 | 789.4            | 2752059.0    | 2714686.0      | 37373.0      |
| DESnar2_contig5| 179092.0 | 825.2            | 770800.0    | 759902.0       | 10898.0      |
| DESnar2_contig6| 11013.0 | 752.8            | 43358.0     | 41914.0        | 1444.0       |
| DESnar2_contig7| 22162.0 | 1875.7           | 220441.0    | 212364.0       | 8077.0       |
| Sequence | Database Retrieved | RefSeq | DSM | ATCC |
|----------|-------------------|--------|-----|------|
| Desulfovibrio acrylicus | IMG | NZ_AU213000000.0 | 10141 | ---- |
| Desulfovibrio aerogenes str. Aspg12 | GenBank | NC_014844.1 | 10631 | 700646 |
| Desulfovibrio africanus str. PCS | IMG | NZ_AOSV0000000.0 | ---- | ---- |
| Desulfovibrio africanus str. Walvis Bay | GenBank | NC_016629.1 | ---- | 19997 |
| Desulfovibrio alaskensis str. DSM 16109 | IMG | NZ_AWQW0000000.1 | 16109 | ---- |
| Desulfovibrio alaskensis str. G20 | GenBank | NC_007519.1 | ---- | ---- |
| Desulfovibrio alkalitolerans | IMG | NZ_JNIA0000000.1 | 5433 | ---- |
| Desulfovibrio alkalitolerans | IMG | NZ_ATBH0000000.1 | 16529 | ---- |
| Desulfovibrio amorphus | IMG | NZ_AUMA0000000.1 | 12254 | ---- |
| Desulfovibrio c. magnus | IMG | NZ_ZAGG0000000.0 | ---- | ---- |
| Desulfovibrio caenuat | IMG | NZ_AUCY0000000.1 | 11391 | ---- |
| Desulfovibrio desulfuricans str. ND132 | GenBank | NC_016803.1 | ---- | ---- |
| Desulfovibrio desulfuricans subsp. aestuarii | IMG | NZ_ARQF0000000.1 | 17919 | 29577 |
| Desulfovibrio desulfuricans subsp. desulfuricans | IMG | NZ_ATUG0000000.1 | 642 | 29577 |
| Desulfovibrio desulfuricans subsp. desulfuricans str. ATCC 27774 | GenBank | NC_011883.1 | ---- | 27774 |
| Desulfovibrio frigidus | IMG | NZ_JQNL0000000.1 | 17176 | ---- |
| Desulfovibrio fructosivorans str. AL | IMG | NZ_AECZ0000000.1 | ---- | ---- |
| Desulfovibrio gigan | GenBank | NC_024444.1 | 1382 | 19364 |
| Desulfovibrio hydrogenophilus str. AM13 | GenBank | NC_020055.1 | 14728 | ---- |
| Desulfovibrio longus | IMG | NZ_ATV0000000.1 | 6739 | 51456 |
| Desulfovibrio magneticus str. RS1 | GenBank | NC_012796.1 | ---- | ---- |
| Desulfovibrio magneticus str. RS1 plasmid pDMC1 | GenBank | NC_012797.1 | ---- | ---- |
| Desulfovibrio magneticus str. RS1 plasmid pDMC2 | GenBank | NC_012798.1 | ---- | ---- |
| Desulfovibrio oreasimae | IMG | NZ_AQXE0000000.1 | 11498 | ---- |
| Desulfovibrio pycnophilus | GenBank | NC_020499.1 | 21447 | ---- |
| Desulfovibrio piger | IMG | NZ_AKUX0000000.1 | 13098 | ---- |
| Desulfovibrio putidatus | IMG | NZ_AUBQ0000000.1 | 16056 | ---- |
| Desulfovibrio salinifontes | GenBank | NC_012881.1 | 2638 | ---- |
| Desulfovibrio sp. A2 | IMG | NZ_A半夜0000000.1 | ---- | ---- |
| Desulfovibrio sp. FW/102B | IMG | NZ_CMB00136.1 | ---- | ---- |
| Desulfovibrio sp. FW/102B plasmid pFW/1010 | IMG | NZ_CMB00136.1 | ---- | ---- |
| Desulfovibrio sp. FW/102B plasmid pFW/10102 | IMG | NZ_CMB00136.1 | ---- | ---- |
| Desulfovibrio sp. USL | IMG | NZ_AHMC0000000.1 | ---- | ---- |
| Desulfovibrio vulgaris str. DP4 | IMG | NC_008731.1 | ---- | ---- |
| Desulfovibrio vulgaris str. DP4 plDMC1 | IMG | NC_008731.1 | ---- | ---- |
| Desulfovibrio vulgaris str. Hildenborough | GenBank | NC_002937.1 | 29579 | ---- |
| Desulfovibrio vulgaris str. Hildenborough plasmid pDV | GenBank | NC_005863.1 | 29579 | ---- |
| Desulfovibrio vulgaris str. Miyazaki F | IMG | NC_011769.1 | ---- | ---- |
| Desulfovibrio vulgaris str. RCH1 | IMG | NC_017311.1 | ---- | ---- |
| Desulfovibrio vulgaris str. RCH1 plasmid pDEVAL1 | IMG | NC_017311.1 | ---- | ---- |
| Desulfovibrio zosterae | IMG | NZ_AUDC0000000.1 | 11974 | ---- |

Table S3. *Desulfovibrio* genomes used in comparison. Table displaying organism names, source databases, RefSeq accession numbers, and culture collection information for all additional *Desulfovibrio* genomes examined in this study.
| Gene name | Gene product                                         |
|-----------|------------------------------------------------------|
| infB      | translation initiation factor-2                     |
| rplB      | 50S ribosomal protein L2                             |
| rplD      | 50S ribosomal protein L4                             |
| rplE      | 50S ribosomal protein L5                             |
| rplF      | 50S ribosomal protein L6                             |
| rplK      | 50S ribosomal protein L11                            |
| rplN      | 50S ribosomal protein L14                            |
| lepA      | GTP binding protein LepA                             |
| pheS      | phenylalanyl-tRNA synthetase, alpha subunit          |
| rpsB      | 30S ribosomal protein S2                             |
| rpsC      | 30S ribosomal protein S3                             |
| rpsE      | 30S ribosomal protein S5                             |
| rpsG      | 30S ribosomal protein S7                             |
| rpsH      | 30S ribosomal protein S8                             |
| rpsl      | 30S ribosomal protein S9                             |
| rpsK      | 30S ribosomal protein S11                            |
| rpsL      | 30S ribosomal protein S12                            |
| rpsM      | 30S ribosomal protein S13                            |
| rpsQ      | 30S ribosomal protein S17                            |
| rpsS      | 30S ribosomal protein S19                            |

Table S4. Genes used to construct multigene phylogeny. Information regarding genes used to construct our multigene phylogeny, including gene name and product.
CONCLUSION

N-fixation by heterotrophic diazotrophs residing in estuarine sediments has recently been shown to be a significant source of N to the overall N budget of these ecosystems under certain environmental conditions (Fulweiler et al., 2007). Sulfate reducing bacteria, such as members of the *Desulfovibrionaceae*, have been found to be one of the dominant groups of diazotrophs in a variety of estuarine systems including the sediments of the Narragansett Bay (Brown and Jenkins, 2014; Fulweiler et al., 2013), Eckernförde Bay (Baltic Sea) (Bertics et al., 2012), and Chesapeake Bay (Burns et al., 2002). The habitats in which these sulfate reducers dominate sediment N-fixation are typically replete for sources of combined N, the presence of which is believed to suppress nitrogen fixation in the environment. Although it has been known for some time that *Desulfovibrio* are capable of fixing N (Sisler & ZoBell, 1951), questions regarding controls on their N fixation behavior, such as why they fix N in seemingly N rich environments, and other questions about their autecology, have remained largely unanswered, in part due to a lack of cultivated representatives.

This thesis sought to couple bacterial genome analysis and comparative genomics with N fixation rate measurements under various combined N conditions in order to examine the response of nitrogenase activity to the presence of different N sources, and link that to the genomic potential (i.e. gene content) of members of the *Desulfovibrionaceae*. Nitrogen fixation rates for pure cultures of *D*. sp. NAR1 grown under a variety of combined N conditions were examined using the acetylene reduction assay (Capone, 1993). Cultures without any reactive N source were found to reduce acetylene to ethylene, with peak ethylene production occurring between 5-10
days after inoculation. Cultures with added 12 mM nitrate were also found to reduce acetylene to ethylene, with peak ethylene production occurring 10 days after inoculation and with rates that were comparable to the reactive-N free control. Cultures with added 12 mM urea were found to reduce acetylene to ethylene, again with peak ethylene production occurring 10 days after inoculation but with rates that were slightly higher than the reactive-N free control and nitrate sets. Increased NA for NAR1 in the presence of urea is something that was observed in earlier pilot experiments, and is a result for which there is currently no explanation or known cause. Cultures with added 12 mM ammonia were not found to reduce acetylene to ethylene, exhibiting no measureable NA, which agrees with what is widely accepted regarding the behavior of the nitrogenase enzyme (Mackerras, 1986). The genomes of both NAR1 and NAR2 isolates were further found to lack genes necessary for catabolism of either nitrate or urea, but were found to possess genes necessary for assimilating and catabolizing ammonia and deaminating amino acids, which agrees with our ARA based observations. Additionally, these gene-level observations held true for the majority of the Desulfovibrio genomes that were assessed in comparison to our environmental isolates, suggesting that an inability to utilize forms of reactive N other than ammonia may be characteristic of the Desulfovibrio. This inability could in turn be part of why we observe N fixation by this group even in environments that are not limited for sources of combined N, such as Narragansett Bay.

In addition to providing insight regarding the N fixation behavior of members of the Desulfovibrio, this study also sought to add to the body of knowledge and collection of representative genomes for this genus by sequencing and annotating the
genomes of two novel *Desulfovibrio* isolates, *Desulfovibrio* sp. NAR1 and *Desulfovibrio* sp. NAR2, and characterizing additional aspects of their physiology and genome content outside of that which directly pertained to N fixation. The NAR1 isolate was found to produce a biofilm under nearly all growth conditions, where NAR2 exhibited no biofilm formation. The biofilm formation in NAR1 was also reflected at the genome level, with an evolutionary gene network revealing genes found only in NAR1 that were involved in sensing and synthesizing c-di-GMP, a cell-to-cell signaling molecule that has been shown to be involved in biofilm formation and persistence (Römling et al., 2013; Hunter et al., 2014; Bellows et al., 2012). This could have important implications for future work involving *D.* sp. NAR1, or other biofilm forming *Desulfovibrio*, as the mechanism and genes involved in biofilm formation for this genus are not well characterized. The NAR1 isolate was also found to possess genes for mercury methylation, an environmentally harmful process that is primarily restricted to anaerobic sulfate reducing bacteria (Compeau & Bartha, 1985; Gilmour et al., 1992). Both isolates were found to have a significant number of genes involved in bacterial defense and heavy metal resistance, which is common for members of the *Desulfovibrio*, however, the presence of these genes may still provide a specific advantage for survival in Narragansett Bay sediments.

This study is meaningful to increase our understanding of the N metabolism and nitrogenase activity of members of the *Desulfovibrionaceae*, who are one of the primary groups responsible for heterotrophic N fixation in Narragansett Bay, RI sediments (Fulweiler et al., 2013). It is also advantageous to add to the body of molecular knowledge regarding this genus, as they are important to several
environmental processes and have additional industry applications (Fulweiler et al., 2013; Parks et al., 2013; Muyzer & Stams, 2008; Hulshoff et al., 1998; Martins et al., 2009; Gutierrez-Sanchez et al., 2011). By increasing our overall understanding of this important diazotrophic group through detailed examination of cultivated representatives and their N fixation profiles, and comparing those results to environmental field observations, we can gain a more comprehensive view of factors controlling prevalence and N fixation behavior of this group. This will then allow for better predictions of how N fixation by this group will respond to changing environmental conditions and could ultimately serve to better inform policies regarding environmental waste and water management.

References:

Bellows, L. E., Koestler, B. J., Karaba, S. M., Waters, C. M., & Lathem, W. W. (2012). Hfq-dependent, co-ordinate control of cyclic diguanylate synthesis and catabolism in the plague pathogen Yersinia pestis. *Molecular Microbiology, 86*, 661–674.

Bertics, V. J., Löscher, C. R., Salonen, I., Dale, A. W., Schmitz, R. A., & Treude, T. (2012). Occurrence of benthic microbial nitrogen fixation coupled to sulfate reduction in the seasonally hypoxic Eckernförde Bay, Baltic Sea. *Biogeosciences Discussions, 9*(6), 6489–6533.

Brown, S. M., & Jenkins, B. D. (2014). Profiling gene expression to distinguish the likely active diazotrophs from a sea of genetic potential in marine sediments. *Environmental Microbiology.* doi:10.1111/1462-2920.12403

Burns, J. A, Zehr, J. P., & Capone, D. G. (2002). Nitrogen-fixing phylotypes of Chesapeake Bay and Neuse River estuarial sediments. *Microbial Ecology, 44*(4), 336–343.

Capone, D. (1993). Determination of Nitrogenase Activity in Aquatic Samples Using the Acetylene Reduction Procedure. In P. Kemp, J. Cole, B. Sherr, & E. Sherr (Eds.), *Handbook of methods in aquatic microbial ecology*. Lewis Publ: Boca Raton, FL.
Compeau, G. C., & Bartha, R. (1985). Sulfate-reducing bacteria: principal methylators of mercury in anoxic estuarine sediment. *Applied and Environmental Microbiology, 50*(2), 498–502.

Fulweiler, R. W., Brown, S. M., Nixon, S. W., & Jenkins, B. D. (2013). Evidence and a conceptual model for the co-occurrence of nitrogen fixation and denitrification in heterotrophic marine sediments. *Marine Ecology Progress Series, 482*, 57–68.

Fulweiler, R. W., Nixon, S. W., Buckley, B. A., & Granger, S. L. (2007). Reversal of the net dinitrogen gas flux in coastal marine sediments. *Nature, 448*(7150), 180–182.

Gilmour, C. C., Henry, E. A., & Mitchell, R. (1992). Sulfate stimulation of mercury methylation in freshwater sediments. *Environmental Science & Technology, 26*(11), 2281–2287.

Gutiérrez-Sánchez, C., Olea, D., Marques, M., Fernández, V. M., Pereira, I. A. C., Vélez, M., & De Lacey, A. L. (2011). Oriented Immobilization of a Membrane-Bound Hydrogenase onto an Electrode for Direct Electron Transfer. *Langmuir, 27*(10), 6449–6457.

Hulshoff Pol, L., Lens, P. L., Stams, A. M., & Lettinga, G. (1998). Anaerobic treatment of sulphate-rich wastewaters. *Biodegradation, 9*(3-4), 213–224.

Hunter, J. L., Severin, G. B., Koestler, B. J., & Waters, C. M. (2014). The Vibrio cholerae diguanylate cyclase VCA0965 has an AGDEF active site and synthesizes cyclic di-GMP. *BMC Microbiology, 14*, 22.

Mackerras, A. H., & Smith, G. D. (1986). Evidence for direct repression of nitrogenase by ammonia in the cyanobacterium Anabaena cylindrica. *Biochemical and Biophysical Research Communications, 134*(2), 835–44.

Martins, M., Faleiro, M. L., Barros, R. J., Veríssimo, A. R., Barreiros, M. A., & Costa, M. C. (2009). Characterization and activity studies of highly heavy metal resistant sulphate-reducing bacteria to be used in acid mine drainage decontamination. *Journal of Hazardous Materials, 166*(2-3), 706–713.

Muyzer, G., & Stams, A. J. M. (2008). The ecology and biotechnology of sulphate-reducing bacteria. *Nature Reviews. Microbiology, 6*(6), 441–454.

Parks, J. M., Johs, A., Podar, M., Bridou, R., Hurt, R. a, Smith, S. D., … Liang, L. (2013). The genetic basis for bacterial mercury methylation. *Science (New York, N.Y.), 339*(6125), 1332–1335.
Römling, U., Galperin, M. Y., & Gomelsky, M. (2013). Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiology and Molecular Biology Reviews: MMBR, 77*, 1–52.

Sisler, F. D., & ZoBell, C. E. (1951). Nitrogen Fixation by Sulfate-reducing Bacteria Indicated by Nitrogen/Argon Ratios. *Science, 113*(2940), 511–512.
### Appendix A. List of proteins contained in Desulfovibrio sp. NAR1 filtered network.

| Network protein number | Contig assignment | Functional annotation |
|------------------------|------------------|-----------------------|
| 3_DESnar1              | 13               | 1,4-alpha-glucan branching enzyme |
| 5_DESnar1              | 12               | 1-acetyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) |
| 11_DESnar1             | 6                | 2'-5' RNA ligase |
| 14_DESnar1             | 15               | 2,3-bisphosphoglycerate-independent phosphoglycerate mutase, archaeal type (EC 5.4.2.1) |
| 20_DESnar1             | 13               | 2-heptaprenyl-1,4-naphthoquinone methyltransferase (EC 2.1.1.163) |
| 21_DESnar1             | 10               | 2-heptaprenyl-1,4-naphthoquinone methyltransferase (EC 2.1.1.163) |
| 22_DESnar1             | 7                | 2-hydroxycromene-2-carboxylate isomerase/DsbA-like thioredoxin domain |
| 27_DESnar1             | 2                | 2-oxoglutarate oxoreductase, alpha subunit (EC 1.2.7.3) |
| 34_DESnar1             | 3                | 2-oxoglutarate oxoreductase, delta subunit, putative (EC 1.2.7.3) |
| 44_DESnar1             | 1                | 3-demethylubiquine-9 3-O-methyltransferase |
| 56_DESnar1             | 13               | 3-polypropenyl-4-hydroxybenzoate carboxy-lyase UbcX (EC 4.1.1.-) |
| 58_DESnar1             | 1                | 4-amino-6-deoxy-N-Acetyl-D-hexosaminyl-(Lip carrier) acetyltransferase |
| 60_DESnar1             | 1                | 4-hydroxy-2-oxovalerate aldolase (EC 4.1.3.39) |
| 64_DESnar1             | 2                | 4-hydroxybenzoyl-CoA thioesterase family active site |
| 67_DESnar1             | 6                | 4Fe-4S binding domain/4Fe-4S dicluster domain |
| 69_DESnar1             | 10               | 4Fe-4S dicluster domain |
| 70_DESnar1             | 2                | 4Fe-4S ferredoxin, iron-sulfur binding domain protein |
| 76_DESnar1             | 10               | 5-formyltetrahydrofolate cyclo-ligase (EC 6.3.3.2) |
| 81_DESnar1             | 10               | 6-phosphogluconolactonase (EC 3.1.1.31), euukaryotic type |
| 83_DESnar1             | 6                | 7,8 dihydropteroyl synthase (methyloterpoin) |
| 84_DESnar1             | 9                | 7,8 dihydropteroyl synthase (methyloterpoin) |
| 85_DESnar1             | 1                | 8-amino-7-oxononanoate synthase (EC 2.3.1.47) |
| 87_DESnar1             | 6                | 16S rRNA processing protein RimM |
| 88_DESnar1             | 6                | 16S rRNA processing protein RimM |
| 89_DESnar1             | 6                | 115 kDa surface antigen |
| 90_DESnar1             | 7                | 115 kDa surface antigen precursor |
| 92_DESnar1             | 15               | 18K peptoglycan-associated outer membrane lipoprotein; Peptoglycan-associated lipoprotein precursor |
| 103_DESnar1            | 1                | A/G-specific adenine glycosylase (EC 3.2.2.-) |
| 105_DESnar1            | 10               | AAA domain containing protein |
| 106_DESnar1            | 1                | ABC transporter system, permease component YbhR |
| 107_DESnar1            | 1                | ABC transporter system, permease component YbhS |
| 109_DESnar1            | 5                | ABC transporter ATP-binding protein |
| 110_DESnar1            | 14               | ABC transporter related |
| 114_DESnar1            | 1                | ABC transporter multirug efflux pump, fused ATP-binding domains |
| 117_DESnar1            | 6                | ABC transporter related |
| 125_DESnar1            | 14               | ABC Transporter ATP-binding protein |
| 129_DESnar1            | 6                | ABC Transporter permease protein |
| 131_DESnar1            | 13               | ABC-type amino ac transport, signal transduction systems, periplasmic component/domain |
| 132_DESnar1            | 13               | ABC-type amino ac transport, signal transduction systems, periplasmic component/domain |
| 133_DESnar1            | 1                | ABC-type amino ac transport, signal transduction systems, periplasmic component/domain |
| 134_DESnar1            | 15               | ABC-type amino ac transport, signal transduction systems, periplasmic component/domain |
| 135_DESnar1            | 13               | ABC-type amino ac transport/signal transduction systems, periplasmic component/domain |
| 136_DESnar1            | 2                | ABC-type amino ac transport/signal transduction systems, periplasmic component/domain |
| 137_DESnar1            | 2                | ABC-type amino ac transport/signal transduction systems, periplasmic component/domain |
| 138_DESnar1            | 5                | ABC-type amino ac transport/signal transduction systems, periplasmic component/domain |
| 139_DESnar1            | 14               | ABC-type amino ac transport/signal transduction systems, periplasmic component/domain |
| 140_DESnar1            | 10               | ABC-type amino ac transport/signal transduction systems, periplasmic component/domain |
| 141_DESnar1            | 15               | ABC-type amino ac transport/signal transduction systems, periplasmic component/domain |
| 142_DESnar1            | 15               | ABC-type amino ac transport/signal transduction systems, periplasmic component/domain |
| 144_DESnar1            | 2                | ABC-type amino ac transport/signal transduction systems, periplasmic component/domain |
| 146_DESnar1            | 1                | ABC-type branched-chain amino ac transport systems, periplasmic component |
| 147_DESnar1            | 10               | ABC-type branched-chain amino ac transport systems, periplasmic component |
| 150_DESnar1            | 1                | ABC-type metal ion transport system, periplasmic component/surface adhesin |
| 151_DESnar1            | 1                | ABC-type metal ion transport system, periplasmic component/surface adhesin |
| 152_DESnar1            | 1                | ABC-type multirug transport system, permease component |
| 153_DESnar1            | 1                | ABC-type nitrate/sulfonate/bicarbonate transport systems, periplasmic components |
| 156_DESnar1            | 1                | ABC-type polysaccharide/polyl phosphate export systems, permease component |
| 157_DESnar1            | 13               | ABC-type sugar transport system, ATPase component |
| 158_DESnar1            | 1                | ABC-type sugar transport system, periplasmic binding protein YcjN |
160_DESnar1 13 ABC-type transport system involved in resistance to organic solvents
162_DESnar1 1 ABC-type tungstate transport system, ATP-binding protein
166_DESnar1 1 Abortive infection C-terminus
167_DESnar1 1 Acetate kinase (EC 2.7.2.1)
180_DESnar1 5 Acetyl-CoA synthetase (ADP-forming) alpha chain (EC 6.2.1.13)
187_DESnar1 10 Acetylmornithine aminotransferase (EC 2.6.1.11)
190_DESnar1 6 Acetyltransferase YpeA
191_DESnar1 6 Acetyltransferases
192_DESnar1 1 Acetyltransferases
193_DESnar1 2 Acetyltransferases
196_DESnar1 2 ACP-hemolysin acyltransferase (hemolysin-activating protein)
197_DESnar1 6 Acriflavin resistance protein
199_DESnar1 14 Acriflavin resistance protein
200_DESnar1 7 Acriflavin resistance protein
201_DESnar1 9 Acriflavin resistance protein
203_DESnar1 14 Activator of (R)-2-hydroxyglutaryl-CoA dehydratase
206_DESnar1 13 Acyl-ACP thioesterase
209_DESnar1 6 Acylphosphate phosphohydrolase (EC 3.6.1.7), putative
210_DESnar1 5 Acylphosphate phosphohydrolase (EC 3.6.1.7), putative
211_DESnar1 1 Acylphosphate phosphohydrolase (EC 3.6.1.7), putative
216_DESnar1 2 Adenylylsulfate kinase (EC 2.7.1.25)
218_DESnar1 9 ADP-heptose synthase (EC 2.7.-.-) / D-glycero-beta-D-manno-heptose 7-phosphate kinase
235_DESnar1 15 ADP-heptose--lipooligosaccharide heptosyltransferase II (EC 2.4.1.-)
240_DESnar1 6 Alanine racemase (EC 5.1.1.1)
243_DESnar1 14 Alcohol dehydrogenase (EC 1.1.1.1)
245_DESnar1 2 Aldose 1-epimerase (EC 5.1.3.3)
250_DESnar1 5 Alkaline phosphatase (EC 3.1.3.1)
255_DESnar1 10 Alkyl hydroperoxide reductase subunit C-like protein
256_DESnar1 5 Aldehyde dehydrogenase, iron-sulfur subunit
257_DESnar1 5 Alpha/beta hydrolase family
260_DESnar1 9 Alpha/beta hydrolase family
262_DESnar1 2 Amohydrolase family protein
266_DESnar1 13 Amino acid ABC transporter substrate-binding protein, PAAT family (TC 3.A.1.3.-)
267_DESnar1 1 Amino acid ABC transporter substrate-binding protein, PAAT family (TC 3.A.1.3.-)
269_DESnar1 15 Amino acid ABC transporter substrate-binding protein, PAAT family (TC 3.A.1.3.-)
271_DESnar1 10 Amino acid ABC Transporter ATP-binding protein
272_DESnar1 10 Amino acid ABC Transporter periplasmic amino acid-binding protein
273_DESnar1 13 Amino acid ABC Transporter periplasmic precursor
274_DESnar1 10 Amino acid ABC Transporter permease protein
279_DESnar1 13 amino acid/amine ABC transporter substrate-binding protein, HAAT family (TC 3.A.1.4.-)
281_DESnar1 9 Amylosaccharomannan lyase (EC 4.1.3.38)
293_DESnar1 2 Anaerobic dehydrogenases, typically selenocysteine-containing
300_DESnar1 10 Ankyrin repeats
305_DESnar1 12 Apolipoprotein N-acyltransferase (EC 2.3.1.-) / Copper homeostasis protein CufE
307_DESnar1 5 Arabinoheptose efflux permease
309_DESnar1 6 AraC-type DNA-binding domain-containing proteins
311_DESnar1 14 AraC-type DNA-binding domain-containing proteins
312_DESnar1 10 AraC-type DNA-binding domain-containing proteins
313_DESnar1 11 AraC-type DNA-binding domain-containing proteins
320_DESnar1 13 Arsenical pump-driving ATPase (EC 3.6.3.16)
321_DESnar1 13 Arsenical resistance operon repressor
322_DESnar1 14 Arsenical resistance operon repressor
| DESnar1 | 325 | Arylesterase precursor (EC 3.1.1.2) |
|---------|-----|-----------------------------------|
| DESnar1 | 326 | Arylsulfotransferase (ASST)        |
| DESnar1 | 327 | AsnA protein                      |
| DESnar1 | 328 | asparagine synthase (glutamine-hydrolyzing) |
| DESnar1 | 329 | Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4) |
| DESnar1 | 330 | ATP-dependent DNA helicase pcra (EC 3.6.1.1) |
| DESnar1 | 331 | ATP-dependent exoDNAse (exonuclease V) alpha subunit - helicase superfamily I member |
| DESnar1 | 332 | ATP-dependent helicase            |
| DESnar1 | 333 | ATP-dependent helicase            |
| DESnar1 | 334 | ATP-dependent transcriptional regulator-like protein |
| DESnar1 | 335 | ATP-utilizing enzyme of the PP-loop superfamily |
| DESnar1 | 336 | ATPase                            |
| DESnar1 | 337 | ATPase component STY3233 of energizing module of queuosine-regulated ECF transporter |
| DESnar1 | 338 | ATPase YjeE, predicted to have essential role in cell wall biosynthesis |
| DESnar1 | 339 | Autolysis histine kinase LytS      |
| DESnar1 | 340 | Bacterial RNase P class A         |
| DESnar1 | 341 | Bacteriophage CI repressor helix-turn-helix domain |
| DESnar1 | 342 | Bacteriophage tail assembly protein |
| DESnar1 | 343 | Baseplate assembly protein J       |
| DESnar1 | 344 | Beta-hexosaminase (EC 3.2.1.52)   |
| DESnar1 | 345 | Beta-propeller domains of methanol dehydrogenase type |
| DESnar1 | 346 | Bifunctional protein: zinc-containing alcohol dehydrogenase; quinone oxoreductase |
| DESnar1 | 347 | binding-protein-dependent transport systems inner membrane component |
| DESnar1 | 348 | Biopolymer transport protein ExbD/ToICl |
| DESnar1 | 349 | Biotin-(acetyl-CoA carboxylase) ligase |
| DESnar1 | 350 | Biotin-protein ligase (EC 6.3.4.15) |
| DESnar1 | 351 | Branched-chain amino acid ABC Transporter amino acid-binding protein (TC 3.A.1.4.1) |
| DESnar1 | 352 | Butyrate kinase (EC 2.7.2.7)      |
| DESnar1 | 353 | CacA/CxxGCC family protein        |
| DESnar1 | 354 | CaaX prenyl protease 1, putative  |
| DESnar1 | 355 | Capsul polysaccharide ABC Transporter ATP-binding protein KpsT |
| DESnar1 | 356 | Capsul polysaccharide export system inner membrane protein KpsE |
| DESnar1 | 357 | Capsul polysaccharide export system periplasmic protein KpsD |
| DESnar1 | 358 | Capsul polysaccharide export system protein KpsC |
| DESnar1 | 359 | Capsul polysaccharide export system protein KpsS |
| DESnar1 | 360 | Carbon monoxide dehydrogenase CooS subunit (EC 1.2.99.2) |
| DESnar1 | 361 | Carbon monoxide-responsive transcriptional activator CooA |
| DESnar1 | 362 | Carbon starvation protein A        |
| DESnar1 | 363 | Carbon starvation protein, predicted membrane protein |
| DESnar1 | 364 | Cardiolipin synthase               |
| DESnar1 | 365 | Cardiolipin synthetase (EC 2.7.8.-) |
| DESnar1 | 366 | Catalase (EC 1.11.1.6)/Peroxase (EC 1.11.1.7) |
| DESnar1 | 367 | Cation/multrug efflux pump         |
| DESnar1 | 368 | Cell division inhibitor            |
| DESnar1 | 369 | Cell division protein FtsQ         |
| DESnar1 | 370 | Cell wall-associated hydrolases (invasion-associated proteins) |
| DESnar1 | 371 | Cephalosporin hydroxylase          |
| DESnar1 | 372 | CgEB family protein                |
| DESnar1 | 373 | CgEB family protein                |
| DESnar1 | 374 | Channel-forming transporter/cytolysins activator of TpsB family |
| DESnar1 | 375 | Chemotaxis protein CheD            |
| DESnar1 | 376 | Chemotaxis protein cheV (EC 2.7.3.-) |
| DESnar1 | 377 | Chemotaxis protein methyltransferase CheR (EC 2.1.1.80) |
| DESnar1 | 378 | Chemotaxis regulator-transmits chemoreceptor signals to flagellar motor components CheY |
| DESnar1 | 379 | Chemotaxis regulator-transmits chemoreceptor signals to flagellar motor components CheY |
| DESnar1 | 380 | Chemotaxis signal transduction protein |
| DESnar1 | 381 | Chitin catabolic cascade sensor histine kinase ChiS |
| DESnar1 | 382 | Choline-glycine betaine transporter |
| DESnar1 | 383 | Chromosome (plasm) partitioning protein ParB |
| DESnar1 | 384 | Chromosome initiation inhibitor    |
Citrate synthase (si) (EC 2.3.3.1)
Class III cytochrome C family
CO dehydrogenase accessory protein CooC (nickel insertion)
Cobalt-precorrin-6 synthase, anaerobic
Cobalt-precorrin-6 synthase, anaerobic
CoD-related protein, GAK system
COG0398: uncharacterized membrane protein/PF00070 family
COG1242: Predicted Fe-S oxdoreductase
COG1720: Uncharacterized conserved protein
COG2740: Predicted nucleic-ac-binding protein implicated in transcription termination
COG2827: putative endonuclease containing a URI domain
COG4123: Predicted O-methyltransferase
Cold-shock DNA-binding protein family
Competence protein F homolog, phosphoribosyltransferase domain; protein YhgH
Conserved hypothetical protein
Conserved protein, permease-related
CTP:molybdopterin cytylyltransferase
CTP:molybdopterin cytylyltransferase
Cupin 2, conserved barrel domain protein
Cyclic nucleote-binding protein
Cyclolysin secretion ATP-binding protein
Cysteine desulfurase (EC 2.8.1.7)
Cytochrome B561
cytochrome c family protein
Cytochrome c family protein
Cytochrome c-type biogenesis protein DelD, protein-disulfe reductase (EC 1.8.1.8)
Cycloisyn secretion ATP-binding protein
Cysteine desulfurase (EC 2.8.1.7)
Cytochrome B561
cytochrome c family protein
Cytochrome c family protein
Cytochrome c family protein
Cytochrome c family protein
Cytochrome c family protein
Cytosine methyltransferase (EC 2.1.1.37)
Cytosine methyltransferase (EC 2.1.1.37)
Cytosine methyltransferase (EC 2.1.1.37)
Cycloisyn secretion ATP-binding protein
Cytochrome c3
Cytochrome C553 (soluble cytochrome f)
D-alanyl-D-alanine carboxypeptidase
D-glycerol-D-manno-heptose 1,7-bisphosphate phosphatase
DegT/Dmr/HyrC1/Sfr S ammnotransferase family protein
Deoxyguanosinephosphate triphosphohydrolase (EC 3.1.5.1)
diguanylate cyclase (GGDEF) domain
diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)
diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)
diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)
diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)
diHdrofrolate synthase (EC 6.3.2.12) / Folypolyglutamate synthase (EC 6.3.2.17)
Dimucleo-utilizing enzymes involved in molybdopterin and thiamine biosynthesis family 2
Dimucleo-utilizing enzymes involved in molybdopterin and thiamine biosynthesis family 2
DNA internalization-related competence protein ComEC/Rec2
DNA topoisomerase IV subunit A
DNA-binding protein, CopG family
DNA-binding protein, CopG family
DNA-binding response regulator LuxR family
DNA-cytosine methyltransferase (EC 2.1.1.37)
DNA-cytosine methyltransferase (EC 2.1.1.37)
DNA-cytosine methyltransferase (EC 2.1.1.37)
Dolichol-phosphate mannosyltransferase (EC 2.4.1.83)
Dolichyl-phosphate-mannose-protein mannosyltransferase family protein
Domain of unknown function (DUF303)
Domain of unknown function (DUF1837)
Domain of unknown function (DUF1987)
dTDP-4-dehydrohamnose 3,5-epimerase (EC 5.1.3.13)
dTDP-4-dehydrohamnose 3,5-epimerase (EC 5.1.3.13)
dTDP-4-dehydrohamnose reductase (EC 1.1.1.133)
dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)
EBNA-1 nuclear protein
electron transfer flavoprotein beta subunit-like protein
Electron transport complex protein RosC
751_DESnar1 5  Electron transport complex protein RnfC
762_DESnar1 6  Endonuclease/exonuclease/phosphatase
770_DESnar1 5  Energy-conserving hydrogenase (ferredoxin), subunit F
773_DESnar1 1  Epoxide hydrolase (EC 3.3.2.9)
774_DESnar1 2  ERF superfamily
775_DESnar1 5  EriK/YbiS/YcfS/YnhG family protein
787_DESnar1 14 Exopolyphosphatase (EC 3.6.1.11)
788_DESnar1 14 Exopolyphosphatase
791_DESnar1 13 Extracellular ligand-binding receptor
792_DESnar1 5  Extracellular ligand-binding receptor
793_DESnar1 6  Extracellular Mutant; Ecm15p
794_DESnar1 14 Extracellular solute binding protein ScrB
795_DESnar1 14 Extracellular solute-binding protein, family 3/GGDEF domain protein
796_DESnar1 13 extracellular solute-binding protein, putative
800_DESnar1 1  Fatty ac desaturase
801_DESnar1 1  Fatty-ac desaturase (EC 3.6.1.11)
808_DESnar1 1  Fatty-ac desaturase (EC 1.14.99.-)
813_DESnar1 1  Fe2+/Zn2+ uptake regulation proteins
814_DESnar1 1  Ferredoxin
815_DESnar1 3  Ferredoxin
816_DESnar1 14 Ferredoxin
820_DESnar1 14 Ferredoxin
821_DESnar1 14 Ferredoxin
826_DESnar1 1  Ferric iron ABC Transporter ATP-binding protein
828_DESnar1 2  Ferric serophore transport system, periplasmic binding protein TonB
831_DESnar1 1  Ferric uptake regulation protein FUR
832_DESnar1 14 Ferric uptake regulation protein FUR
840_DESnar1 1  FIG000233: metal-dependent hydrolase
843_DESnar1 9  FIG000605: protein co-occurring with transport systems (COG1739)
848_DESnar1 14 FIG004453: protein YceG like
850_DESnar1 9  FIG004694: Hypothetical protein
851_DESnar1 13 FIG022606: AAA ATPase
854_DESnar1 1  FIG049111: Hypothetical protein in pyoverdin gene cluster
855_DESnar1 3  FIG057253: Fe-S oxoreductase
856_DESnar1 13 FIG059250: hypothetical protein
862_DESnar1 3  FIG146085: 3'-to-5' oligoribonuclease A, Bacillus type
865_DESnar1 1  FIG00440303: hypothetical protein
870_DESnar1 2  FIG00602277: hypothetical protein
871_DESnar1 3  FIG00602295: hypothetical protein
874_DESnar1 3  FIG00602413: hypothetical protein
875_DESnar1 15 FIG00602413: hypothetical protein
876_DESnar1 10 FIG00602432: hypothetical protein
884_DESnar1 9  FIG00602496: hypothetical protein
886_DESnar1 2  FIG00602510: hypothetical protein
887_DESnar1 2  FIG00602543: hypothetical protein
888_DESnar1 15 FIG00602569: hypothetical protein
895_DESnar1 15 FIG00602578: hypothetical protein
897_DESnar1 2  FIG00602945: hypothetical protein
899_DESnar1 2  FIG00602963: hypothetical protein
900_DESnar1 1  FIG00603294: hypothetical protein
901_DESnar1 3  FIG00603589: hypothetical protein
916_DESnar1 15 FIG00603666: hypothetical protein
917_DESnar1 2  FIG00603746: hypothetical protein
922_DESnar1 5  FIG00603852: hypothetical protein
929_DESnar1 2  FIG00604212: hypothetical protein
930_DESnar1 5  FIG00604258: hypothetical protein
937_DESnar1 3  FIG00604763: hypothetical protein
942_DESnar1 1  FIG00605110: hypothetical protein
945_DESnar1 5  FIG00605436: hypothetical protein
947_DESnar1 5  FIG00605486: hypothetical protein
Flagellar assembly protein FliH
Flagellar basal-body P-ring formation protein FlgA
Flagellar basal-body rod protein FlgB
Flagellar hook protein FlgE
Flagellar hook-probasal body complex protein FlfE
Flagellar motor length control protein FlgK
Flagellar motor rotation protein MotB
Flagellar protein FlgJ (peptoglycan hydrolase) (EC 3.2.1.2)
Flagellin protein FlaA
Flavodoxin domain/4Fe-4S binding domain
Flp pilus assembly protein TadD, contains TPR repeat
FOG: GGDEF domain
FOG: HEAT repeat
FOG: HPt domain
FOG: PAS/PAC domain
FOG: TPR repeat, SEL1 subfamily
Foldase protein PrsA precursor (EC 5.2.1.8)
Formate dehydrogenase chain D (EC 1.2.1.2)
Formate dehydrogenase formation protein FdhE
Formate dehydrogenase-O, major subunit (EC 1.2.1.2)
Fosfomycin resistance protein FosX
Fructose-2,6-bisphosphatase
Glucose-1-phosphate cytylyltransferase (EC 2.7.7.33)
Glutamine ABC Transporter periplasmic glutamine-binding
Glutamate dehydrogenase, family 13 domain protein
Glycosyl transferase family 2
Glycosyl transferase, family 2
Glycosyl transferase family 2
Glycosyl transferase group 1
Glycosyl transferase, family 2
| Gene Name | Value | Description |
|-----------|-------|-------------|
| 1286_DESnar1 | 13 | hypothetical protein |
| 1288_DESnar1 | 13 | hypothetical protein |
| 1289_DESnar1 | 13 | hypothetical protein |
| 1290_DESnar1 | 13 | hypothetical protein |
| 1291_DESnar1 | 13 | hypothetical protein |
| 1292_DESnar1 | 13 | hypothetical protein |
| 1293_DESnar1 | 13 | hypothetical protein |
| 1294_DESnar1 | 13 | hypothetical protein |
| 1295_DESnar1 | 13 | hypothetical protein |
| 1296_DESnar1 | 13 | hypothetical protein |
| 1297_DESnar1 | 13 | hypothetical protein |
| 1298_DESnar1 | 13 | hypothetical protein |
| 1299_DESnar1 | 13 | hypothetical protein |
| 1301_DESnar1 | 13 | hypothetical protein |
| 1302_DESnar1 | 13 | hypothetical protein |
| 1303_DESnar1 | 13 | hypothetical protein |
| 1304_DESnar1 | 13 | hypothetical protein |
| 1305_DESnar1 | 13 | hypothetical protein |
| 1306_DESnar1 | 13 | hypothetical protein |
| 1307_DESnar1 | 13 | hypothetical protein |
| 1308_DESnar1 | 13 | hypothetical protein |
| 1309_DESnar1 | 13 | hypothetical protein |
| 1310_DESnar1 | 13 | hypothetical protein |
| 1311_DESnar1 | 13 | hypothetical protein |
| 1312_DESnar1 | 13 | hypothetical protein |
| 1313_DESnar1 | 13 | hypothetical protein |
| 1314_DESnar1 | 13 | hypothetical protein |
| 1315_DESnar1 | 13 | hypothetical protein |
| 1316_DESnar1 | 13 | hypothetical protein |
| 1317_DESnar1 | 13 | hypothetical protein |
| 1318_DESnar1 | 13 | hypothetical protein |
| 1319_DESnar1 | 13 | hypothetical protein |
| 1320_DESnar1 | 13 | hypothetical protein |
| 1321_DESnar1 | 13 | hypothetical protein |
| 1322_DESnar1 | 13 | hypothetical protein |
| 1323_DESnar1 | 13 | hypothetical protein |
| 1324_DESnar1 | 13 | hypothetical protein |
| 1325_DESnar1 | 13 | hypothetical protein |
| 1326_DESnar1 | 13 | hypothetical protein |
| 1327_DESnar1 | 13 | hypothetical protein |
| 1328_DESnar1 | 13 | hypothetical protein |
| 1329_DESnar1 | 13 | hypothetical protein |
| 1330_DESnar1 | 13 | hypothetical protein |
| 1331_DESnar1 | 13 | hypothetical protein |
| 1332_DESnar1 | 13 | hypothetical protein |
| 1333_DESnar1 | 13 | hypothetical protein |
| 1334_DESnar1 | 13 | hypothetical protein |
| 1335_DESnar1 | 13 | hypothetical protein |
| 1336_DESnar1 | 13 | hypothetical protein |
| 1337_DESnar1 | 13 | hypothetical protein |
| 1338_DESnar1 | 13 | hypothetical protein |
| 1339_DESnar1 | 13 | hypothetical protein |
| 1340_DESnar1 | 13 | hypothetical protein |
| 1341_DESnar1 | 13 | hypothetical protein |
| 1342_DESnar1 | 13 | hypothetical protein |
| 1343_DESnar1 | 13 | hypothetical protein |
| 1344_DESnar1 | 13 | hypothetical protein |
| 1345_DESnar1 | 13 | hypothetical protein |
| 1346_DESnar1 | 13 | hypothetical protein |
| 1347_DESnar1 | 13 | hypothetical protein |
| 1348_DESnar1 | 13 | hypothetical protein |
| 1349_DESnar1 | 13 | hypothetical protein |
| 1350_DESnar1 | 13 | hypothetical protein |
| 1351_DESnar1 | 13 | hypothetical protein |
| 1352_DESnar1 | 13 | hypothetical protein |
| 1353_DESnar1 | 13 | hypothetical protein |
| 1354_DESnar1 | 13 | hypothetical protein |
| 1355_DESnar1 | 13 | hypothetical protein |
| 1356_DESnar1 | 13 | hypothetical protein |
| 1357_DESnar1 | 13 | hypothetical protein |
| 1358_DESnar1 | 13 | hypothetical protein |
| 1359_DESnar1 | 13 | hypothetical protein |
| 1360_DESnar1 | 13 | hypothetical protein |
| 1361_DESnar1 | 13 | hypothetical protein |
| 1362_DESnar1 | 13 | hypothetical protein |
| 1363_DESnar1 | 13 | hypothetical protein |
| 1364_DESnar1 | 13 | hypothetical protein |
| 1365_DESnar1 | 13 | hypothetical protein |
| 1366_DESnar1 | 13 | hypothetical protein |
| 1367_DESnar1 | 13 | hypothetical protein |
| 1368_DESnar1 | 13 | hypothetical protein |
| 1369_DESnar1 | 13 | hypothetical protein |
| 1370_DESnar1 | 13 | hypothetical protein |
| 1371_DESnar1 | 13 | hypothetical protein |
| 1372_DESnar1 | 13 | hypothetical protein |
| 1373_DESnar1 | 13 | hypothetical protein |
| 1374_DESnar1 | 13 | hypothetical protein |
| 1375_DESnar1 | 13 | hypothetical protein |
| 1376_DESnar1 | 13 | hypothetical protein |
| 1377_DESnar1 | 13 | hypothetical protein |
| 1378_DESnar1 | 13 | hypothetical protein |
| 1379_DESnar1 | 13 | hypothetical protein |
| 1380_DESnar1 | 13 | hypothetical protein |
| 1381_DESnar1 | 13 | hypothetical protein |
| 1382_DESnar1 | 13 | hypothetical protein |
| 1383_DESnar1 | 13 | hypothetical protein |
| 1384_DESnar1 | 13 | hypothetical protein |
| 1385_DESnar1 | 13 | hypothetical protein |
| 1386_DESnar1 | 13 | hypothetical protein |
| 1387_DESnar1 | 13 | hypothetical protein |
| 1388_DESnar1 | 13 | hypothetical protein |
| 1389_DESnar1 | 13 | hypothetical protein |
| 1390_DESnar1 | 13 | hypothetical protein |
| 1391_DESnar1 | 13 | hypothetical protein |
| 1392_DESnar1 | 13 | hypothetical protein |
| 1393_DESnar1 | 13 | hypothetical protein |
| 1394_DESnar1 | 13 | hypothetical protein |
| 1395_DESnar1 | 13 | hypothetical protein |
| 1396_DESnar1 | 13 | hypothetical protein |
| 1397_DESnar1 | 13 | hypothetical protein |
| 1398_DESnar1 | 13 | hypothetical protein |
| 1399_DESnar1 | 13 | hypothetical protein |
| 1400_DESnar1 | 13 | hypothetical protein |
| 1401_DESnar1 | 13 | hypothetical protein |
| 1402_DESnar1 | 13 | hypothetical protein |
| 1403_DESnar1 | 13 | hypothetical protein |
| 1404_DESnar1 | 13 | hypothetical protein |
| 1405_DESnar1 | 13 | hypothetical protein |
| Gene Name     | Value | Description      |
|--------------|-------|------------------|
| 1627_DSNar1  | 1     | hypothetical protein |
| 1628_DSNar1  | 1     | hypothetical protein |
| 1632_DSNar1  | 1     | hypothetical protein |
| 1633_DSNar1  | 1     | hypothetical protein |
| 1634_DSNar1  | 1     | hypothetical protein |
| 1635_DSNar1  | 1     | hypothetical protein |
| 1636_DSNar1  | 1     | hypothetical protein |
| 1637_DSNar1  | 1     | hypothetical protein |
| 1638_DSNar1  | 1     | hypothetical protein |
| 1639_DSNar1  | 1     | hypothetical protein |
| 1640_DSNar1  | 1     | hypothetical protein |
| 1641_DSNar1  | 1     | hypothetical protein |
| 1642_DSNar1  | 1     | hypothetical protein |
| 1643_DSNar1  | 1     | hypothetical protein |
| 1644_DSNar1  | 1     | hypothetical protein |
| 1645_DSNar1  | 1     | hypothetical protein |
| 1646_DSNar1  | 1     | hypothetical protein |
| 1647_DSNar1  | 1     | hypothetical protein |
| 1648_DSNar1  | 1     | hypothetical protein |
| 1649_DSNar1  | 1     | hypothetical protein |
| 1650_DSNar1  | 1     | hypothetical protein |
| 1651_DSNar1  | 1     | hypothetical protein |
| 1652_DSNar1  | 1     | hypothetical protein |
| 1653_DSNar1  | 1     | hypothetical protein |
| 1654_DSNar1  | 1     | hypothetical protein |
| 1655_DSNar1  | 1     | hypothetical protein |
| 1656_DSNar1  | 1     | hypothetical protein |
| 1657_DSNar1  | 1     | hypothetical protein |
| 1658_DSNar1  | 1     | hypothetical protein |
| 1659_DSNar1  | 1     | hypothetical protein |
| 1660_DSNar1  | 1     | hypothetical protein |
| 1661_DSNar1  | 1     | hypothetical protein |
| 1662_DSNar1  | 1     | hypothetical protein |
| 1663_DSNar1  | 1     | hypothetical protein |
| 1664_DSNar1  | 1     | hypothetical protein |
| 1665_DSNar1  | 1     | hypothetical protein |
| 1666_DSNar1  | 1     | hypothetical protein |
| 1667_DSNar1  | 1     | hypothetical protein |
| 1668_DSNar1  | 1     | hypothetical protein |
| 1669_DSNar1  | 1     | hypothetical protein |
| 1670_DSNar1  | 1     | hypothetical protein |
| 1671_DSNar1  | 1     | hypothetical protein |
| 1672_DSNar1  | 1     | hypothetical protein |
| 1673_DSNar1  | 1     | hypothetical protein |
| 1674_DSNar1  | 1     | hypothetical protein |
| 1675_DSNar1  | 1     | hypothetical protein |
| 1676_DSNar1  | 1     | hypothetical protein |
| 1677_DSNar1  | 1     | hypothetical protein |
| 1678_DSNar1  | 1     | hypothetical protein |
| 1679_DSNar1  | 1     | hypothetical protein |
| 1680_DSNar1  | 1     | hypothetical protein |
| 1681_DSNar1  | 1     | hypothetical protein |
| 1682_DSNar1  | 1     | hypothetical protein |
| 1683_DSNar1  | 1     | hypothetical protein |
| 1684_DSNar1  | 1     | hypothetical protein |
| 1685_DSNar1  | 1     | hypothetical protein |
| 1686_DSNar1  | 1     | hypothetical protein |
| 1687_DSNar1  | 1     | hypothetical protein |
| 1688_DSNar1  | 1     | hypothetical protein |
| 1689_DSNar1  | 1     | hypothetical protein |
| 1690_DSNar1  | 1     | hypothetical protein |
| 1691_DSNar1  | 1     | hypothetical protein |
| 1692_DSNar1  | 1     | hypothetical protein |
| 1693_DSNar1  | 1     | hypothetical protein |
| 1694_DSNar1  | 1     | hypothetical protein |
| 1695_DSNar1  | 1     | hypothetical protein |
| 1696_DSNar1  | 1     | hypothetical protein |
| 1697_DSNar1  | 1     | hypothetical protein |
| 1698_DSNar1  | 1     | hypothetical protein |
| 1699_DSNar1  | 1     | hypothetical protein |
| 1700_DSNar1  | 1     | hypothetical protein |
| 1701_DSNar1  | 1     | hypothetical protein |
| 1702_DSNar1  | 1     | hypothetical protein |
| 1703_DSNar1  | 1     | hypothetical protein |
| 1704_DSNar1  | 2     | hypothetical protein |
| 1705_DSNar1  | 2     | hypothetical protein |
| 1706_DSNar1  | 2     | hypothetical protein |
| 1707_DSNar1  | 2     | hypothetical protein |
| 1708_DSNar1  | 2     | hypothetical protein |
| Accession | Value | Name               |
|-----------|-------|--------------------|
| 1875_DESna1 | 5     | hypothetical protein |
| 1877_DESna1 | 5     | hypothetical protein |
| 1878_DESna1 | 5     | hypothetical protein |
| 1879_DESna1 | 5     | hypothetical protein |
| 1881_DESna1 | 5     | hypothetical protein |
| 1882_DESna1 | 5     | hypothetical protein |
| 1883_DESna1 | 5     | hypothetical protein |
| 1884_DESna1 | 5     | hypothetical protein |
| 1885_DESna1 | 5     | hypothetical protein |
| 1886_DESna1 | 5     | hypothetical protein |
| 1887_DESna1 | 5     | hypothetical protein |
| 1888_DESna1 | 5     | hypothetical protein |
| 1889_DESna1 | 5     | hypothetical protein |
| 1890_DESna1 | 5     | hypothetical protein |
| 1891_DESna1 | 5     | hypothetical protein |
| 1892_DESna1 | 5     | hypothetical protein |
| 1893_DESna1 | 5     | hypothetical protein |
| 1894_DESna1 | 5     | hypothetical protein |
| 1895_DESna1 | 5     | hypothetical protein |
| 1896_DESna1 | 5     | hypothetical protein |
| 1897_DESna1 | 5     | hypothetical protein |
| 1898_DESna1 | 5     | hypothetical protein |
| 1900_DESna1 | 5     | hypothetical protein |
| 1901_DESna1 | 5     | hypothetical protein |
| 1902_DESna1 | 5     | hypothetical protein |
| 1905_DESna1 | 5     | hypothetical protein |
| 1906_DESna1 | 5     | hypothetical protein |
| 1907_DESna1 | 5     | hypothetical protein |
| 1909_DESna1 | 5     | hypothetical protein |
| 1910_DESna1 | 5     | hypothetical protein |
| 1911_DESna1 | 5     | hypothetical protein |
| 1912_DESna1 | 5     | hypothetical protein |
| 1913_DESna1 | 5     | hypothetical protein |
| 1914_DESna1 | 5     | hypothetical protein |
| 1915_DESna1 | 5     | hypothetical protein |
| 1916_DESna1 | 5     | hypothetical protein |
| 1918_DESna1 | 5     | hypothetical protein |
| 1919_DESna1 | 5     | hypothetical protein |
| 1920_DESna1 | 5     | hypothetical protein |
| 1921_DESna1 | 5     | hypothetical protein |
| 1922_DESna1 | 5     | hypothetical protein |
| 1924_DESna1 | 5     | hypothetical protein |
| 1925_DESna1 | 5     | hypothetical protein |
| 1926_DESna1 | 5     | hypothetical protein |
| 1927_DESna1 | 5     | hypothetical protein |
| 1928_DESna1 | 5     | hypothetical protein |
| 1929_DESna1 | 5     | hypothetical protein |
| 1930_DESna1 | 5     | hypothetical protein |
| 1931_DESna1 | 5     | hypothetical protein |
| 1933_DESna1 | 5     | hypothetical protein |
| 1939_DESna1 | 5     | hypothetical protein |
| 1975_DESna1 | 14    | hypothetical protein |
| 1977_DESna1 | 14    | hypothetical protein |
| 1978_DESna1 | 14    | hypothetical protein |
| 1980_DESna1 | 14    | hypothetical protein |
| 1981_DESna1 | 14    | hypothetical protein |
| 1982_DESna1 | 14    | hypothetical protein |
| 1983_DESna1 | 14    | hypothetical protein |
| 1984_DESna1 | 14    | hypothetical protein |
| 1985_DESna1 | 14    | hypothetical protein |
| 1986_DESna1 | 14    | hypothetical protein |
| 1988_DESna1 | 14    | hypothetical protein |
| Year | DESnar1 | Length | Type          |
|------|---------|--------|---------------|
| 1989 | DESnar1 | 14     | hypothetical protein |
| 1990 | DESnar1 | 14     | hypothetical protein |
| 1991 | DESnar1 | 14     | hypothetical protein |
| 1992 | DESnar1 | 14     | hypothetical protein |
| 1993 | DESnar1 | 14     | hypothetical protein |
| 1994 | DESnar1 | 14     | hypothetical protein |
| 1995 | DESnar1 | 14     | hypothetical protein |
| 1997 | DESnar1 | 14     | hypothetical protein |
| 1999 | DESnar1 | 14     | hypothetical protein |
| 2000 | DESnar1 | 14     | hypothetical protein |
| 2001 | DESnar1 | 14     | hypothetical protein |
| 2002 | DESnar1 | 14     | hypothetical protein |
| 2003 | DESnar1 | 14     | hypothetical protein |
| 2004 | DESnar1 | 14     | hypothetical protein |
| 2006 | DESnar1 | 14     | hypothetical protein |
| 2008 | DESnar1 | 14     | hypothetical protein |
| 2009 | DESnar1 | 14     | hypothetical protein |
| 2010 | DESnar1 | 14     | hypothetical protein |
| 2011 | DESnar1 | 14     | hypothetical protein |
| 2013 | DESnar1 | 14     | hypothetical protein |
| 2015 | DESnar1 | 14     | hypothetical protein |
| 2018 | DESnar1 | 14     | hypothetical protein |
| 2019 | DESnar1 | 14     | hypothetical protein |
| 2021 | DESnar1 | 14     | hypothetical protein |
| 2023 | DESnar1 | 14     | hypothetical protein |
| 2024 | DESnar1 | 14     | hypothetical protein |
| 2026 | DESnar1 | 14     | hypothetical protein |
| 2027 | DESnar1 | 14     | hypothetical protein |
| 2029 | DESnar1 | 14     | hypothetical protein |
| 2030 | DESnar1 | 14     | hypothetical protein |
| 2031 | DESnar1 | 14     | hypothetical protein |
| 2050 | DESnar1 | 7      | hypothetical protein |
| 2051 | DESnar1 | 7      | hypothetical protein |
| 2052 | DESnar1 | 7      | hypothetical protein |
| 2054 | DESnar1 | 7      | hypothetical protein |
| 2055 | DESnar1 | 7      | hypothetical protein |
| 2056 | DESnar1 | 7      | hypothetical protein |
| 2057 | DESnar1 | 7      | hypothetical protein |
| 2058 | DESnar1 | 7      | hypothetical protein |
| 2060 | DESnar1 | 7      | hypothetical protein |
| 2061 | DESnar1 | 7      | hypothetical protein |
| 2062 | DESnar1 | 7      | hypothetical protein |
| 2063 | DESnar1 | 7      | hypothetical protein |
| 2064 | DESnar1 | 7      | hypothetical protein |
| 2065 | DESnar1 | 7      | hypothetical protein |
| 2066 | DESnar1 | 7      | hypothetical protein |
| 2067 | DESnar1 | 7      | hypothetical protein |
| 2068 | DESnar1 | 7      | hypothetical protein |
| 2069 | DESnar1 | 7      | hypothetical protein |
| 2070 | DESnar1 | 7      | hypothetical protein |
| 2071 | DESnar1 | 7      | hypothetical protein |
| 2072 | DESnar1 | 7      | hypothetical protein |
| 2073 | DESnar1 | 7      | hypothetical protein |
| 2074 | DESnar1 | 7      | hypothetical protein |
| 2075 | DESnar1 | 7      | hypothetical protein |
| 2076 | DESnar1 | 7      | hypothetical protein |
| 2079 | DESnar1 | 7      | hypothetical protein |
| 2080 | DESnar1 | 7      | hypothetical protein |
| 2081 | DESnar1 | 7      | hypothetical protein |
| 2082 | DESnar1 | 7      | hypothetical protein |
| 2083 | DESnar1 | 7      | hypothetical protein |
| 2084 | DESnar1 | 7      | hypothetical protein |
| 2085 | DESnar1 | 7      | hypothetical protein |
| Accession | Position | Description         |
|-----------|----------|---------------------|
| DESnar1   | 209      | hypothetical protein|
| DESnar1   | 2109     | hypothetical protein|
| DESnar1   | 2110     | hypothetical protein|
| DESnar1   | 2112     | hypothetical protein|
| DESnar1   | 2116     | hypothetical protein|
| DESnar1   | 2117     | hypothetical protein|
| DESnar1   | 2118     | hypothetical protein|
| DESnar1   | 2121     | hypothetical protein|
| DESnar1   | 2122     | hypothetical protein|
| DESnar1   | 2125     | hypothetical protein|
| DESnar1   | 2126     | hypothetical protein|
| DESnar1   | 2127     | hypothetical protein|
| DESnar1   | 2128     | hypothetical protein|
| DESnar1   | 2129     | hypothetical protein|
| DESnar1   | 2132     | hypothetical protein|
| DESnar1   | 2133     | hypothetical protein|
| DESnar1   | 2134     | hypothetical protein|
| DESnar1   | 2135     | hypothetical protein|
| DESnar1   | 2136     | hypothetical protein|
| DESnar1   | 2137     | hypothetical protein|
| DESnar1   | 2138     | hypothetical protein|
| DESnar1   | 2142     | hypothetical protein|
| DESnar1   | 2143     | hypothetical protein|
| DESnar1   | 2144     | hypothetical protein|
| DESnar1   | 2152     | hypothetical protein|
| DESnar1   | 2153     | hypothetical protein|
| DESnar1   | 2154     | hypothetical protein|
| DESnar1   | 2155     | hypothetical protein|
| DESnar1   | 2156     | hypothetical protein|
| DESnar1   | 2157     | hypothetical protein|
| DESnar1   | 2158     | hypothetical protein|
| DESnar1   | 2159     | hypothetical protein|
| DESnar1   | 2160     | hypothetical protein|
| DESnar1   | 2161     | hypothetical protein|
| DESnar1   | 2162     | hypothetical protein|
| DESnar1   | 2164     | hypothetical protein|
| DESnar1   | 2165     | hypothetical protein|
| DESnar1   | 2166     | hypothetical protein|
| DESnar1   | 2167     | hypothetical protein|
| DESnar1   | 2170     | hypothetical protein|
| DESnar1   | 2171     | hypothetical protein|
| DESnar1   | 2173     | hypothetical protein|
| DESnar1   | 2174     | hypothetical protein|
| DESnar1   | 2175     | hypothetical protein|
| DESnar1   | 2176     | hypothetical protein|
| DESnar1   | 2178     | hypothetical protein|
| DESnar1   | 2179     | hypothetical protein|
| DESnar1   | 2180     | hypothetical protein|
| DESnar1   | 2182     | hypothetical protein|
| DESnar1   | 2184     | hypothetical protein|
| DESnar1   | 2185     | hypothetical protein|
| DESnar1   | 2186     | hypothetical protein|
| DESnar1   | 2187     | hypothetical protein|
| DESnar1   | 2188     | hypothetical protein|
| DESnar1   | 2189     | hypothetical protein|
| DESnar1   | 2192     | hypothetical protein|
| DESnar1   | 2194     | hypothetical protein|
| DESnar1   | 2195     | hypothetical protein|
| DESnar1   | 2197     | hypothetical protein|
| DESnar1   | 2198     | hypothetical protein|
| DESnar1   | 2199     | hypothetical protein|
| DESnar1   | 2200     | hypothetical protein|
| Gene ID | Description |
|--------|-------------|
| 2201  | hypothetical protein |
| 2202  | hypothetical protein |
| 2204  | hypothetical protein |
| 2205  | hypothetical protein |
| 2206  | hypothetical protein |
| 2208  | hypothetical protein |
| 2212  | hypothetical protein |
| 2213  | hypothetical protein |
| 2214  | hypothetical protein |
| 2215  | hypothetical protein |
| 2216  | hypothetical protein |
| 2217  | hypothetical protein |
| 2218  | hypothetical protein |
| 2220  | hypothetical protein |
| 2223  | hypothetical protein |
| 2227  | hypothetical protein |
| 2228  | hypothetical protein |
| 2229  | hypothetical protein |
| 2230  | hypothetical protein |
| 2231  | hypothetical protein |
| 2232  | hypothetical protein |
| 2233  | hypothetical protein |
| 2234  | hypothetical protein |
| 2235  | hypothetical protein |
| 2236  | hypothetical protein |
| 2237  | hypothetical protein |
| 2238  | hypothetical protein |
| 2239  | hypothetical protein |
| 2240  | hypothetical protein |
| 2241  | hypothetical protein |
| 2242  | hypothetical protein |
| 2243  | hypothetical protein |
| 2244  | hypothetical protein |
| 2245  | hypothetical protein |
| 2246  | hypothetical protein |
| 2247  | hypothetical protein |
| 2248  | hypothetical protein |
| 2249  | hypothetical protein |
| 2250  | hypothetical protein |
| 2251  | hypothetical protein |
| 2252  | hypothetical protein |
| 2253  | hypothetical protein |
| 2254  | hypothetical protein |
| 2255  | hypothetical protein |
| 2256  | hypothetical protein |
| 2257  | hypothetical protein |
| 2258  | hypothetical protein |
| 2259  | hypothetical protein |
| 2260  | hypothetical protein |
| 2261  | hypothetical protein |
| 2262  | hypothetical protein |
| 2263  | hypothetical protein |
| 2264  | hypothetical protein |
| 2265  | hypothetical protein |
| 2266  | hypothetical protein |
| 2268  | hypothetical protein |
| 2269  | hypothetical protein |
| 2270  | hypothetical protein |
| 2271  | hypothetical protein |
| 2272  | hypothetical protein |
| 2273  | hypothetical protein |
| 2274  | hypothetical protein |
| 2275  | hypothetical protein |
| 2276  | hypothetical protein |
| 2277  | hypothetical protein |
| 2278  | hypothetical protein |
| 2279  | hypothetical protein |
| 2280  | hypothetical protein |
| 2281  | hypothetical protein |
| 2282  | hypothetical protein |
| 2283  | hypothetical protein |
| 2284  | hypothetical protein |
| 2285  | hypothetical protein |
| 2286  | hypothetical protein |
| 2287  | hypothetical protein |
| 2288  | hypothetical protein |
| 2289  | hypothetical protein |
| 2290  | hypothetical protein |
| 2291  | hypothetical protein |
| 2292  | hypothetical protein |
| 2293  | hypothetical protein |
| 2294  | hypothetical protein |
| 2295  | hypothetical protein |
| 2296  | hypothetical protein |
| 2297  | hypothetical protein |
| 2298  | hypothetical protein |
| 2299  | hypothetical protein |
| 2300  | hypothetical protein |
| 2301  | hypothetical protein |
| 2302  | hypothetical protein |
| 2303  | hypothetical protein |
| 2304  | hypothetical protein |
| 2305  | hypothetical protein |
| 2306  | hypothetical protein |
| 2307  | hypothetical protein |
| 2308  | hypothetical protein |
| 2309  | hypothetical protein |
| 2310  | hypothetical protein |
| 2311  | hypothetical protein |
| 2312  | hypothetical protein |
| 2313  | hypothetical protein |
| 2314  | hypothetical protein |
| 2315  | hypothetical protein |
| 2316  | hypothetical protein |
| 2317  | hypothetical protein |
| 2318  | hypothetical protein |
| 2319  | hypothetical protein |
| 2320  | hypothetical protein |
| 2340  | Hypothetical protein Cj1505c |
| 2341  | hypothetical protein YaeJ |
| 5614  | Imazole glycerol phosphate synthase amotransferase |
| 5616  | Imazole glycerol phosphate synthase cyclase |
| 5618  | Imazolonepropionase and related amohydrolases |
| 5621  | Import inner membrane translocase, subunit Tim44 |
| 5622  | Inactive homolog of metal-dependent proteases |
| 5631  | Inositol-1-monophosphatase (EC 3.1.3.25) |
| Accession | Gene Name                      | Description                                                                 |
|-----------|-------------------------------|------------------------------------------------------------------------------|
| 5634_DESnar1 | 10  | Integral membrane protein, DUF6                                            |
| 5636_DESnar1 | 1   | Integrase                                                                  |
| 5637_DESnar1 | 2   | Integrase                                                                  |
| 5638_DESnar1 | 2   | Integrase                                                                  |
| 5639_DESnar1 | 5   | Integrase                                                                  |
| 5641_DESnar1 | 1   | Integrase family protein                                                  |
| 5645_DESnar1 | 15  | iron (III) ABC transporter ATP-binding                                   |
| 5648_DESnar1 | 13  | Iron(III) dicitrate-binding protein                                       |
| 5651_DESnar1 | 3   | Iron-sulfur cluster assembly scaffold protein                             |
| 5653_DESnar1 | 10  | Iron-sulfur cluster binding protein                                       |
| 5654_DESnar1 | 1   | Iron-sulfur cluster-binding protein                                       |
| 5656_DESnar1 | 14  | iron-sulfur cluster-binding protein                                       |
| 5658_DESnar1 | 2   | iron-sulfur flavoprotein                                                  |
| 5660_DESnar1 | 13  | iron-sulfur flavoprotein, putative                                         |
| 5662_DESnar1 | 6   | Isochorismatase (EC 3.3.2.1)                                              |
| 5669_DESnar1 | 6   | Kinesin light chain                                                        |
| 5670_DESnar1 | 1    | L-serine dehydratase (EC 4.3.1.17)                                        |
| 5687_DESnar1 | 10  | Lhr-like helicases                                                         |
| 5688_DESnar1 | 1   | Lip A export ATP-binding/permease protein MsbA                             |
| 5689_DESnar1 | 9   | Lip carrier : UDP-N-acetylgalactosaminyltransferase                       |
| 5695_DESnar1 | 13  | Lipopolysaccharide biosynthesis chain length determinant                  |
| 5702_DESnar1 | 6   | lipoprotein, putative                                                      |
| 5703_DESnar1 | 2   | lipoprotein, putative                                                      |
| 5707_DESnar1 | 9   | lipoprotein, putative                                                      |
| 5710_DESnar1 | 14  | Long-chain acyl-CoA synthetases (AMP-forming)                              |
| 5711_DESnar1 | 1   | Long-chain fatty ac transport protein                                      |
| 5713_DESnar1 | 1   | Long-chain-fatty-ac-CoA ligase (EC 6.2.1.3)                                |
| 5741_DESnar1 | 10  | LSU ribosomal protein L25p                                                 |
| 5747_DESnar1 | 13  | LSU ribosomal protein L32P                                                 |
| 5753_DESnar1 | 2   | LysE type translocator                                                    |
| 5756_DESnar1 | 16  | Lysine exporter protein (LYSE/YGGA)                                        |
| 5758_DESnar1 | 6   | Lysophospholipase L1 and related esterases                                |
| 5761_DESnar1 | 5   | M. jannaschii predicted coding region MJ0455                               |
| 5762_DESnar1 | 3   | Mac 1                                                                     |
| 5764_DESnar1 | 2   | Magnesium and cobalt efflux protein CorC                                  |
| 5766_DESnar1 | 13  | major facilitator superfamily MFS 1                                       |
| 5768_DESnar1 | 10  | major facilitator superfamily MFS 1                                       |
| 5771_DESnar1 | 1   | Malonyl CoA-acyl carrier protein transacylase                              |
| 5773_DESnar1 | 1   | Manganese transport system ATP-binding protein                            |
| 5775_DESnar1 | 3   | Mannose-1-phosphate guanylyltransferase (GDP)                              |
| 5777_DESnar1 | 9   | MATE efflux family protein                                                 |
| 5779_DESnar1 | 6   | Membrane fusion protein of RND family multrug efflux pump                 |
| 5780_DESnar1 | 2   | Membrane fusion protein of RND family multrug efflux pump                 |
| 5781_DESnar1 | 13  | membrane protein                                                           |
| 5782_DESnar1 | 6   | membrane protein                                                           |
| 5783_DESnar1 | 14  | Membrane protein                                                           |
| 5789_DESnar1 | 13  | membrane protein, putative                                                |
| 5790_DESnar1 | 3   | membrane protein, putative                                                |
| 5793_DESnar1 | 5   | membrane protein, putative                                                |
| 5796_DESnar1 | 10  | Membrane proteins related to metalloendopeptidases                         |
| 5798_DESnar1 | 1   | Membrane-associated lipoprotein                                           |
| 5800_DESnar1 | 6   | Membrane-bound lysozyme inhibitor of c-type lysozyme                       |
| 5802_DESnar1 | 1   | Membrane-bound lytic murine transglycosylase A precursor                  |
| 5803_DESnar1 | 1   | Membrane-bound lytic murine transglycosylase B precursor                  |
| 5807_DESnar1 | 1   | Membrane-bound lytic murine transglycosylase E (EC 3.2.1.-)               |
| 5809_DESnar1 | 7   | Membrane-fusion protein                                                   |
| 5811_DESnar1 | 6   | Menaquinone via futalosine step 2                                         |
| 5814_DESnar1 | 2   | Mercuric ion reductase (EC 1.16.1.1)                                      |
| 5815_DESnar1 | 6   | metal dependent phosphohydrolase                                          |
| 5818_DESnar1 | 11  | Metal-dependent hydrolase (EC 3.-.-.-)                                    |
| 5819_DESnar1 | 1   | Metal-dependent hydrolase                                                 |
| 5821_DESnar1 | 5   | Metal-dependent hydrolases of the beta-lactamase superfamily II           |
| Gene ID   | Description                                                                 |
|----------|------------------------------------------------------------------------------|
| 5822_DESnar1 | Metal-dependent phosphohydrolase                                              |
| 5823_DESnar1 | Metal-dependent phosphohydrolase                                              |
| 5828_DESnar1 | Metallophosphoesterase                                                       |
| 5835_DESnar1 | Methionine transporter MetT                                                  |
| 5839_DESnar1 | Methyl-accepting chemotaxis protein                                          |
| 5840_DESnar1 | Methyl-accepting chemotaxis protein                                          |
| 5841_DESnar1 | Methyl-accepting chemotaxis protein                                          |
| 5842_DESnar1 | Methyl-accepting chemotaxis protein                                          |
| 5843_DESnar1 | Methyl-accepting chemotaxis protein                                          |
| 5844_DESnar1 | Methyl-accepting chemotaxis protein                                          |
| 5845_DESnar1 | Methyl-accepting chemotaxis protein                                          |
| 5846_DESnar1 | Methyl-accepting chemotaxis protein                                          |
| 5847_DESnar1 | Methyl-accepting chemotaxis protein                                          |
| 5849_DESnar1 | Methyl-accepting chemotaxis protein                                          |
| 5850_DESnar1 | Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)         |
| 5851_DESnar1 | Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)         |
| 5852_DESnar1 | Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)         |
| 5853_DESnar1 | Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)         |
| 5854_DESnar1 | Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)         |
| 5855_DESnar1 | Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)         |
| 5856_DESnar1 | Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)         |
| 5857_DESnar1 | Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)         |
| 5858_DESnar1 | Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)         |
| 5859_DESnar1 | Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)         |
| 5860_DESnar1 | Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)         |
| 5861_DESnar1 | Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)         |
| 5862_DESnar1 | Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)         |
| 5863_DESnar1 | Methylase involved in ubiquinone/maenquinone biosynthesis                     |
| 5865_DESnar1 | Methylated-DNA--protein-cysteine methyltransferase (EC 2.1.1.63)             |
| 5866_DESnar1 | Methylated-DNA--protein-cysteine methyltransferase-related protein           |
| 5871_DESnar1 | Methyltransferase                                                            |
| 5872_DESnar1 | Methyltransferase Rd1 738                                                    |
| 5873_DESnar1 | Methyltransferase, FkbM family domain protein                                |
| 5874_DESnar1 | Methyltransferase, putative                                                  |
| 5880_DESnar1 | Mll3258 protein                                                             |
| 5881_DESnar1 | Mlr1851 protein                                                             |
| 5882_DESnar1 | Mlr1851 protein                                                             |
| 5883_DESnar1 | MobD                                                                         |
| 5884_DESnar1 | Mobile element protein                                                       |
| 5885_DESnar1 | Mobile element protein                                                       |
| 5889_DESnar1 | Molybdenum ABC Transporter periplasmic molybdenum-binding protein            |
| 5890_DESnar1 | Molybdenum ABC Transporter periplasmic molybdenum-binding protein            |
| 5891_DESnar1 | Molybdenum cofactor biosynthesis protein MoaA                                 |
| 5892_DESnar1 | Molybdenum transport ATP-binding protein ModC (TC 3.A.1.8.1)                 |
| 5893_DESnar1 | Molybdenum transport ATP-binding protein ModC (TC 3.A.1.8.1)                 |
| 5890_DESnar1 | Molybdenum transport ATP-binding protein ModC (TC 3.A.1.8.1)                 |
| 5892_DESnar1 | Molybdenum transport system permease protein ModB (TC 3.A.1.8.1)             |
| 5894_DESnar1 | Molybdenum-pterin binding domain                                              |
| 5902_DESnar1 | Molybdenum-pterin binding domain protein/site-specific recombinase           |
| 5913_DESnar1 | Molybdopterin biosynthesis protein MoeA                                      |
| 5915_DESnar1 | Molybdopterin oxoeductase, membrane subunit                                  |
| 5916_DESnar1 | Molybdopterin-guanine dinucleote biosynthesis protein A                       |
| 5919_DESnar1 | Molybdopterin-guanine dinucleote biosynthesis protein MobB                   |
| 5921_DESnar1 | MotA/TolQ/ExbB proton channel family protein                                 |
| 5923_DESnar1 | multi-sensor hybr histine kinase                                              |
| 5924_DESnar1 | multi-sensor hybr histine kinase                                              |
| 5935_DESnar1 | N-acetyl sugar amotransferase                                                 |
| 5936_DESnar1 | N-Acetyl-D-glucosamine ABC transport system, sugar-binding protein           |
| 5938_DESnar1 | N-acetylglucosamine-1-phosphate uryltransferase                              |
| 5941_DESnar1 | N-acetylmannosaminyltransferase (EC 2.4.1.187)                               |
| 5944_DESnar1 | N-Acetylneuraminic cytylyltransferase (EC 2.7.7.43)                          |
| ID   | Description                                                                                       | EC Numbers |
|------|--------------------------------------------------------------------------------------------------|------------|
| 5945 | N-Acetylmuramic acid cytoyltransferase (EC 2.7.7.43)                                            |            |
| 5946 | N-Acetylmuramic acid cytoyltransferase (EC 2.7.7.43)                                            |            |
| 5947 | N-acetylmuraminate synthase (EC 2.5.1.56)                                                       |            |
| 5949 | N-carbamoylputrescine amase (3.5.1.53)                                                           |            |
| 5959 | NAD binding oxoreductase                                                                         |            |
| 5960 | NAD kinase (EC 2.7.1.23)                                                                         |            |
| 5964 | NAD(P)H oxoreductase YRKL (EC 1.6.99.-)                                                          |            |
| 5965 | NAD(P)H-nitrite reductase                                                                         |            |
| 5966 | NAD(P)HX epimerase / NAD(P)HX dehydratase                                                        |            |
| 5986 | NADPH:quinone reductase                                                                           |            |
| 5989 | NHL domain/cytochrome c family protein                                                           |            |
| 5994 | Nicotinamide-adenylyltransferase, NadM family                                                    |            |
| 5995 | Nicotinate-nucleotide adenyllyltransferase (EC 2.7.7.18)                                         |            |
| 6013 | Nitrogenase FeMo-cofactor synthesis FeS core scaffold                                            |            |
| 6014 | Nitroreductase                                                                                    |            |
| 6017 | Nitroreductase family protein                                                                     |            |
| 6023 | Nucleoside triphosphate pyrophosphohydrolase MazG                                                |            |
| 6024 | Nucleoside-diphosphate-sugar epimerases                                                          |            |
| 6025 | Nucleoside-diphosphate-sugar epimerases                                                           |            |
| 6032 | O-methyltransferase involved in polykete biosynthesis                                              |            |
| 6033 | O-methyltransferase, family 2                                                                     |            |
| 6034 | O-methyltransferase, family 2                                                                     |            |
| 6035 | Octaheme tetrathionate reductase                                                                  |            |
| 6036 | Octaheme tetrathionate reductase                                                                  |            |
| 6050 | Ornithine cyclodeaminase (EC 4.3.1.12)                                                            |            |
| 6053 | Osmosensitive K+ channel histidine kinase KdpD                                                    |            |
| 6054 | Osmotically inducible protein Y precursor                                                         |            |
| 6055 | Outer membrane autotransporter barrel                                                             |            |
| 6056 | Outer membrane efflux protein                                                                     |            |
| 6058 | Outer membrane lipoprotein                                                                       |            |
| 6059 | Outer membrane lipoprotein omp16 precursor                                                        |            |
| 6060 | Outer membrane protein                                                                           |            |
| 6062 | Outer membrane protein                                                                           |            |
| 6063 | Outer membrane protein H precursor                                                               |            |
| 6064 | Outer membrane protein H precursor                                                               |            |
| 6066 | Outer membrane receptor for ferrienterochelin and colicins                                       |            |
| 6067 | Outer membrane receptor proteins, mostly Fe transport                                             |            |
| 6068 | Outer membrane receptor proteins, mostly Fe transport                                             |            |
| 6072 | Oxoreductase (EC 1.1.1.1)                                                                         |            |
| 6074 | oxoreductase, Gfo/h/MocA family                                                                  |            |
| 6076 | Oxoreductase, short-chain dehydrogenase/reductase family                                          |            |
| 6079 | Pantoate--beta-alanine ligase (EC 6.3.2.1)                                                        |            |
| 6081 | PAI2 family protein                                                                               |            |
| 6082 | PAI2 superfamily                                                                                 |            |
| 6083 | Para-aminobenzoate synthase, amotransferase component                                            |            |
| 6084 | Para-aminobenzoate synthase, aminase component                                                    |            |
| 6085 | Paraquat-inducible protein B                                                                      |            |
| 6086 | Parvalin-like peptidyl-prolyl isomerase                                                           |            |
| 6088 | PAS domain S-box                                                                                 |            |
| 6089 | PAS domain S-box                                                                                 |            |
| 6090 | PAS domain S-box                                                                                 |            |
| 6091 | PAS domain S-box                                                                                 |            |
| 6092 | PAS domain S-box                                                                                 |            |
| 6093 | PAS domain S-box                                                                                 |            |
| 6094 | PAS domain S-box                                                                                 |            |
| 6095 | PAS domain S-box                                                                                 |            |
| 6096 | PAS domain S-box                                                                                 |            |
| 6098 | PAS domain S-box/diguanulate cyclase (GGDEF) domain                                              |            |
| 6099 | PAS modulated sigma54 specific transcriptional regulator                                          |            |
| 6100 | PDZ domain protein                                                                               |            |
| ID     | Description                                                                 | Type                  | Class       | Subclass                  |
|--------|------------------------------------------------------------------------------|-----------------------|-------------|---------------------------|
| 6104   | Peptase S24, S26A and S26B                                                     | Peptase               | M16 family  |                           |
| 6105   | Peptase, M16 family                                                          | Peptase               | M23/M315    | family                    |
| 6108   | Peptase, M23/M315 family                                                      | Peptase               | M23/M315    | family                    |
| 6110   | Peptase, M23/M315 family                                                      | Peptase               | M23/M315    | family                    |
| 6122   | Peptyl-prolyl cis-trans isomerase (EC 5.2.1.8)                               | Peptase               | M16 family  |                           |
| 6124   | Peptyl-prolyl cis-trans isomerase PpiA precursor (EC 5.2.1.8)                | Peptase               | M23/M315    | family                    |
| 6126   | Peptyl-prolyl cis-trans isomerase ppIC (EC 5.2.1.8)                          | Peptase               | M23/M315    | family                    |
| 6130   | Periplasmic amino acid-binding protein-related protein                        | Peptase               | M16 family  |                           |
| 6131   | Periplasmic aromatic aldehyde oxidoreductase                                  | Peptase               | M23/M315    | family                    |
| 6132   | Periplasmic binding protein, putative                                         | Peptase               | M23/M315    | family                    |
| 6133   | Periplasmic chaperone for outer membrane proteins Skp                        | Peptase               | M23/M315    | family                    |
| 6134   | Periplasmic component of the Tol biopolymer transport system                 | Peptase               | M23/M315    | family                    |
| 6138   | Periplasmic protein involved in polysaccharate export                        | Peptase               | M23/M315    | family                    |
| 6143   | Permease of the drug/metabolite transporter (DMT) superfamily                | Peptase               | M23/M315    | family                    |
| 6145   | Permease of the drug/metabolite transporter (DMT) superfamily                | Peptase               | M23/M315    | family                    |
| 6151   | Phage baseplate assembly protein W                                           | Peptase               | M23/M315    | family                    |
| 6153   | Phage DNA packaging protein, Nu1 subunit of terminase                         | Peptase               | M23/M315    | family                    |
| 6155   | Phage integrase                                                               | Peptase               | M23/M315    | family                    |
| 6156   | Phage major tail tube protein                                                 | Peptase               | M23/M315    | family                    |
| 6157   | Phage portal protein                                                          | Peptase               | M23/M315    | family                    |
| 6158   | Phage protein                                                                 | Peptase               | M23/M315    | family                    |
| 6159   | Phage protein D                                                               | Peptase               | M23/M315    | family                    |
| 6160   | Phage protein U                                                               | Peptase               | M23/M315    | family                    |
| 6161   | Phage regulatory protein, rha family                                          | Peptase               | M23/M315    | family                    |
| 6162   | Phage Rha protein                                                             | Peptase               | M23/M315    | family                    |
| 6164   | Phage tail protein                                                            | Peptase               | M23/M315    | family                    |
| 6165   | Phage tail protein, P2 protein I family                                       | Peptase               | M23/M315    | family                    |
| 6166   | Phage tail sheath monomer                                                     | Peptase               | M23/M315    | family                    |
| 6167   | Phage terminase, large subunit                                                | Peptase               | M23/M315    | family                    |
| 6168   | Phage-related protein                                                         | Peptase               | M23/M315    | family                    |
| 6169   | Phage-related tail protein                                                    | Peptase               | M23/M315    | family                    |
| 6175   | Phenylacetate-coenzyme A ligase (EC 6.2.1.30)                                 | Peptase               | M23/M315    | family                    |
| 6177   | Phenylacetic ac degradation protein PaaD, thioesterase                         | Peptase               | M23/M315    | family                    |
| 6198   | Phosphoenolpyruvate synthase (EC 2.7.9.2)                                     | Peptase               | M23/M315    | family                    |
| 6205   | Phosphoesterase, PA-phosphatase related                                       | Peptase               | M23/M315    | family                    |
| 6209   | Phosphoglycerate mutase family                                                | Peptase               | M23/M315    | family                    |
| 6210   | Phosphoglycerate mutase family protein                                        | Peptase               | M23/M315    | family                    |
| 6212   | Phospholip methyltransferase                                                   | Peptase               | M23/M315    | family                    |
| 6214   | Phosphomethylpyrimidine kinase                                                 | Peptase               | M23/M315    | family                    |
| 6215   | Phosphonate ABC transporter phosphate-binding                                  | Peptase               | M23/M315    | family                    |
| 6233   | Phosphosulfolactate phosphohydrolase and related enzymes                      | Peptase               | M23/M315    | family                    |
| 6236   | Phosphotransferase system                                                     | Peptase               | M23/M315    | family                    |
| 6239   | Polyferredoxin NapH (periplasmic nitrate reductase)                            | Peptase               | M23/M315    | family                    |
| 6240   | Polymyxin resistance protein                                                  | Peptase               | M23/M315    | family                    |
| 6244   | Positive regulator of CheA protein activity                                    | Peptase               | M23/M315    | family                    |
| 6247   | Possible ATLS1-like light-inducible protein                                    | Peptase               | M23/M315    | family                    |
| 6250   | Possible sterol desaturase                                                     | Peptase               | M23/M315    | family                    |
| 6252   | Potassium efflux system KefA protein                                          | Peptase               | M23/M315    | family                    |
| 6258   | Predicted acyltransferases                                                    | Peptase               | M23/M315    | family                    |
| 6266   | Predicted dehydrogenases and related proteins                                 | Peptase               | M23/M315    | family                    |
| 6270   | Predicted Fe-S-cluster oxoreductase                                            | Peptase               | M23/M315    | family                    |
| 6273   | Predicted glycosyltransferases                                                 | Peptase               | M23/M315    | family                    |
| 6276   | Predicted hydrolases or acyltransferases                                      | Peptase               | M23/M315    | family                    |
| 6279   | Predicted membrane fusion protein (MFP)                                        | Peptase               | M23/M315    | family                    |
| 6280   | Predicted membrane protein                                                    | Peptase               | M23/M315    | family                    |
| 6281   | Predicted metal-dependent hydrolase                                           | Peptase               | M23/M315    | family                    |
| 6286   | Predicted O-methyltransferase                                                  | Peptase               | M23/M315    | family                    |
| 6292   | Predicted regulator of STY3230 transporter operon                             | Peptase               | M23/M315    | family                    |
| 6297   | Predicted signal transduction protein                                          | Peptase               | M23/M315    | family                    |
Predicted Zn-dependent peptases
Probable 2-phosphosulfolactate phosphatase (EC 3.1.3.71)
Probable ATP/GTP binding protein
probable beta-D-galactosase
Probable Co/Zn/Cd efflux system membrane fusion protein
probable exported protein STY2149
probable glycosyl transferase
probable membrane protein YPO22915
probable methyltransferase
Probable poly(beta-D-mannuronate) O-acetylase (EC 2.3.1.-)
probable selenium-dependent hydroxylase accessory protein YqeC
probable tail fiber assembly protein
PROBABLE TRANSMEMBRANE PROTEIN
Probable two-component sensor
Protein containing aminopeptase domain
Protein of unknown function (DUF1847)
Protein of unknown function (DUF3108)
Protein of unknown function CGGC region
protein of unknown function DUF45
protein of unknown function DUF606
 Protein-N(5)-glutamine methyltransferase PrmC
Pseudaminic ac biosynthesis protein PseA
pseudaminic ac biosynthesis-associated protein PseG
2 Pseudourylate synthases, 23S RNA-specific
PTS system, mannose-specific I component (EC 2.7.1.69)
Putative activity regulator of membrane protease YbbK
Putative amino ac ABC Transporter periplasmic amino ac-binding protein
Putative cytochrome C
Putative cytoplasmic protein
putative efflux protein, MATE family
putative exonuclease
putative glycosyltransferase
Putative hemagglutinin/hemolysin-related protein
Putative L-asparaginase (EC 3.5.1.1) (L-asparagine amohydrolase)
Putative L-xylulose-5-phosphate 3-epimerase
Putative lipoprotein
putative LPS biosynthesis protein WbpG
putative membrane protein
putative membrane protein
putative membrane protein
Putative metal chaperone, involved in Zn homeostasis, GTPase
Putative methionine Transporter NhaC family (TC 2.A.35.1.-)
Putative Nudix hydrolase YfcD (EC 3.6.-.-)
PUTATIVE PERIPLASMIC PROTEIN
Putative periplasmic protein YibQ
Putative PQQ enzyme repeat
Putative PQQ enzyme repeat
putative protein
putative secreted lipase
putative sulfonate/nitrate transport system substrate-binding protein
Pyroxamine S'-phosphate oxase-related, FMN-binding
Pyrroline-5-carboxylate reductase (EC 1.5.1.2)
radical activating enzyme
Radical SAM domain protein
Radical SAM domain protein
Radical SAM domain protein
radical SAM/Cys-rich domain protein
Related to MCBG protein
Resolvase
Respiratory arsenate reductase cytoplasmic chaperone
| Gene ID | Description |
|---------|-------------|
| 6464_DESnar1 | Respiratory arsenate reductase FeS subunit |
| 6468_DESnar1 | Response regulator |
| 6470_DESnar1 | Response regulator |
| 6476_DESnar1 | Response regulator |
| 6477_DESnar1 | Response regulator |
| 6479_DESnar1 | Response regulator containing a CheY-like receiver domain |
| 6484_DESnar1 | Response regulator of zinc sigma-54-dependent |
| 6488_DESnar1 | Response regulator of zinc sigma-54-dependent |
| 6500_DESnar1 | Response regulator receiver protein |
| 6505_DESnar1 | Response regulator receiver protein |
| 6507_DESnar1 | Response regulator receiver protein |
| 6511_DESnar1 | Restriction endonuclease |
| 6514_DESnar1 | Rhodanese-related sulfurtransferase |
| 6516_DESnar1 | Rhombo family protein |
| 6540_DESnar1 | Ribosomal large subunit pseudourine synthase B |
| 6541_DESnar1 | Ribosomal large subunit pseudourine synthase C |
| 6542_DESnar1 | Ribosomal large subunit pseudourine synthase D |
| 6571_DESnar1 | RNA-binding protein Jag |
| 6583_DESnar1 | RsbR, positive regulator of sigma-B |
| 6598_DESnar1 | SAM-dependent methyltransferases |
| 6601_DESnar1 | Sensor histidine kinase |
| 6605_DESnar1 | Sensor histidine kinase |
| 6610_DESnar1 | Sensor histidine kinase |
| 6611_DESnar1 | Sensor histidine kinase |
| 6613_DESnar1 | Sensor histidine kinase |
| 6618_DESnar1 | Sensory box histidine kinase/Response regulator |
| 6619_DESnar1 | Sensory box histidine kinase/Response regulator |
| 6621_DESnar1 | Sensory box/GGDEF family protein |
| 6622_DESnar1 | Sensory transduction regulatory protein |
| 6624_DESnar1 | Ser/Thr protein phosphatase family |
| 6627_DESnar1 | Serine phosphatase RsbU, regulator of sigma subunit |
| 6629_DESnar1 | Serine phosphatase RsbU, regulator of sigma subunit |
| 6631_DESnar1 | Serine protease precursor MucD/AlgY associated with sigma factor RpoE |
| 6635_DESnar1 | Similar to ATPase (AAA+ superfamily) |
| 6639_DESnar1 | Similar to CDP-glucose 4,6-dehydratase (EC 4.2.1.45) |
| 6640_DESnar1 | Similar to TadZ/CpaE, associated with Flp pilus assembly |
| 6644_DESnar1 | Similar to carbon monoxide dehydrogenase |
| 6645_DESnar1 | Similar to CDP-glucose 4,6-dehydratase (EC 4.2.1.45) |
| 6646_DESnar1 | Similar to TadZ/CpaE, associated with Flp pilus assembly |
| 6647_DESnar1 | Sigma-54 dependent transcriptional regulator/Response regulator |
| 6648_DESnar1 | Sigma-54 dependent transcriptional regulator/Response regulator |
| 6652_DESnar1 | Sigma-54 dependent transcriptional regulator/Response regulator |
| 6653_DESnar1 | Sigma-70, region 4 |
| 6656_DESnar1 | Site-specific recombinase XerD |
| 6657_DESnar1 | Site-specific recombinase XerD |
| 6658_DESnar1 | Site-specific recombinases, DNA invertase Pin homologs |
| 6670_DESnar1 | Sodium-dependent phosphate transporter |
| 6671_DESnar1 | SOS-response transcriptional repressors |
| 6721_DESnar1 | Staphylococcus nuclease (SNase) domain |
| Accession | Count |
|-----------|-------|
| 6722_DESnar1 | 1 |
| 6723_DESnar1 | 15 |
| 6724_DESnar1 | 6 |
| 6726_DESnar1 | 3 |
| 6729_DESnar1 | 1 |
| 6737_DESnar1 | 5 |
| 6741_DESnar1 | 5 |
| 6745_DESnar1 | 2 |
| 6752_DESnar1 | 1 |
| 6753_DESnar1 | 3 |
| 6754_DESnar1 | 1 |
| 6755_DESnar1 | 1 |
| 6756_DESnar1 | 1 |
| 6757_DESnar1 | 10 |
| 6758_DESnar1 | 10 |
| 6759_DESnar1 | 14 |
| 6761_DESnar1 | 13 |
| 6762_DESnar1 | 2 |
| 6763_DESnar1 | 2 |
| 6764_DESnar1 | 2 |
| 6766_DESnar1 | 13 |
| 6767_DESnar1 | 13 |
| 6768_DESnar1 | 14 |
| 6771_DESnar1 | 14 |
| 6780_DESnar1 | 10 |
| 6788_DESnar1 | 10 |
| 6802_DESnar1 | 5 |
| 6804_DESnar1 | 6 |
| 6805_DESnar1 | 1 |
| 6806_DESnar1 | 5 |
| 6807_DESnar1 | 14 |
| 6808_DESnar1 | 14 |
| 6809_DESnar1 | 14 |
| 6810_DESnar1 | 14 |
| 6811_DESnar1 | 7 |
| 6812_DESnar1 | 10 |
| 6813_DESnar1 | 10 |
| 6814_DESnar1 | 6 |
| 6818_DESnar1 | 10 |
| 6820_DESnar1 | 14 |
| 6821_DESnar1 | 14 |
| 6823_DESnar1 | 2 |
| 6825_DESnar1 | 11 |
| 6826_DESnar1 | 2 |
| 6827_DESnar1 | 2 |
| 6828_DESnar1 | 13 |
| 6830_DESnar1 | 6 |
| 6831_DESnar1 | 6 |
| 6835_DESnar1 | 11 |
| 6846_DESnar1 | 1 |
| 6847_DESnar1 | 9 |
| 6848_DESnar1 | 15 |
| 6850_DESnar1 | 2 |
| 6851_DESnar1 | 6 |
| 6852_DESnar1 | 1 |
| 6853_DESnar1 | 2 |
| 6854_DESnar1 | 5 |
| 6855_DESnar1 | 14 |
| 6856_DESnar1 | 14 |
| 6858_DESnar1 | 7 |
| 6859_DESnar1 | 10 |
| 6860_DESnar1 | 10 |

| Element Description | Accession | Count |
|---------------------|-----------|-------|
| Structural elements; Cell Exterior; surface polysaccharides/antigens | 6722_DESnar1 | 1 |
| Suc2_YeiO_YrdC_YwIC family protein | 6723_DESnar1 | 15 |
| Substrate-specific component BioY of biotin ECF transporter | 6726_DESnar1 | 3 |
| Succinate dehydrogenase subunit C (EC 1.3.5.1) | 6729_DESnar1 | 1 |
| Sulfate adenylyltransferase subunit 2 (EC 2.7.7.4) | 6737_DESnar1 | 5 |
| Sulfate permease | 6741_DESnar1 | 5 |
| Sulfite reductase assimilatory-type (EC 1.8.7.-) | 6745_DESnar1 | 2 |
| Sulfotransferase | 6752_DESnar1 | 1 |
| Sulfotransferase domain | 6753_DESnar1 | 3 |
| Sulfotransferase domain | 6754_DESnar1 | 1 |
| Sulfotransferase domain | 6755_DESnar1 | 1 |
| Sulfur carrier protein adenylyltransferase ThiF | 6756_DESnar1 | 1 |
| Sulfur carrier protein adenylyltransferase ThiF | 6757_DESnar1 | 10 |
| Superfamily I DNA/RNA helicase protein | 6758_DESnar1 | 10 |
| Superfamily II DNA and RNA helicases | 6759_DESnar1 | 14 |
| Surface lipoprotein | 6761_DESnar1 | 13 |
| Survival protein SurA precursor (Peptidyl-prolyl cis-trans isomerase SurA) (EC 5.2.1.8) | 6762_DESnar1 | 2 |
| Tail fiber protein | 6763_DESnar1 | 2 |
| Tail protein I | 6764_DESnar1 | 2 |
| Tetraacyldisaccharide 4'-kinase (EC 2.7.1.130) | 6766_DESnar1 | 13 |
| Tetranomycin polykete synthesis O-methyltransferase tcmP | 6767_DESnar1 | 13 |
| Tetranomycin polykete synthesis O-methyltransferase tcmP | 6768_DESnar1 | 14 |
| Tetraacrylate repeat | 6771_DESnar1 | 14 |
| ThiJ_PfP family protein | 6780_DESnar1 | 10 |
| Thioredoxin family protein | 6788_DESnar1 | 10 |
| TonB family C-terminal domain | 6802_DESnar1 | 5 |
| TonB-dependent receptor | 6804_DESnar1 | 6 |
| TonB-dependent receptor | 6805_DESnar1 | 1 |
| TonB-dependent receptor | 6806_DESnar1 | 5 |
| TonB-dependent receptor | 6807_DESnar1 | 14 |
| TonB-dependent receptor | 6808_DESnar1 | 14 |
| TonB-dependent receptor | 6809_DESnar1 | 14 |
| TonB-dependent receptor | 6810_DESnar1 | 14 |
| TonB-dependent receptor | 6811_DESnar1 | 7 |
| TonB-dependent receptor | 6812_DESnar1 | 10 |
| TonB-dependent receptor | 6813_DESnar1 | 10 |
| TonB-dependent receptor | 6814_DESnar1 | 6 |
| TPR domain protein | 6818_DESnar1 | 10 |
| TPR domain protein | 6820_DESnar1 | 14 |
| TPR domain protein, putative component of TonB system | 6821_DESnar1 | 14 |
| TPR domain protein, putative component of TonB system | 6823_DESnar1 | 2 |
| TPR repeat repeat-containing protein | 6825_DESnar1 | 11 |
| TPR repeat containing exported protein; Putative periplasmic protein | 6826_DESnar1 | 2 |
| TPR repeat containing exported protein; Putative periplasmic protein | 6827_DESnar1 | 2 |
| TPR repeat precursor | 6828_DESnar1 | 13 |
| TPR repeat-containing protein | 6830_DESnar1 | 6 |
| TPR repeat-containing protein | 6831_DESnar1 | 6 |
| TPR repeat-containing protein | 6835_DESnar1 | 11 |
| Transcriptional regulator | 6846_DESnar1 | 1 |
| Transcriptional regulator | 6847_DESnar1 | 9 |
| Transcriptional regulator | 6848_DESnar1 | 15 |
| Transcriptional regulator | 6850_DESnar1 | 2 |
| Transcriptional regulator YbiH, TetR family | 6851_DESnar1 | 6 |
| Transcriptional regulator AraC family | 6852_DESnar1 | 1 |
| Transcriptional regulator AraC family | 6853_DESnar1 | 2 |
| Transcriptional regulator AraC family | 6854_DESnar1 | 5 |
| Transcriptional regulator AraC family | 6855_DESnar1 | 14 |
| Transcriptional regulator AraC family | 6856_DESnar1 | 14 |
| Transcriptional regulator AraC family | 6858_DESnar1 | 7 |
| Transcriptional regulator AraC family | 6859_DESnar1 | 10 |
| Transcriptional regulator AraC family | 6860_DESnar1 | 10 |
Transcriptional regulator AraC family

Transcriptional regulator AraC family

Transcriptional regulator AraC family

Transcriptional regulator AsnC family

Transcriptional regulator GntR family

Transcriptional regulator GntR family

Transcriptional regulator IclR family

Transcriptional regulator LysR family

Transcriptional regulator LysR family

Transcriptional regulator LysR family

Transcriptional regulator MarR family/Acetyltransferase(GNAT)

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transcriptional regulator MarR family

Transporter LysE family

Transporter LysE family

Transporter LysE family

Transporter putative

Transposase and inactivated derivatives

Transposase and inactivated derivatives

Transposase and inactivated derivatives

Transposase and inactivated derivatives

Transposase and inactivated derivatives

Transposase and inactivated derivatives

Transposase and inactivated derivatives

Transposase and inactivated derivatives

Transposase Tn15 transposition protein tnsA

Transposase Tn15 transposition protein tnsA

Transposase Tn15 transposition protein tnsB

Transposase Tn15 transposition protein tnsB

Transposase Tn15 transposition protein tnsC

Transposase Tn15 transposition protein tnsD

Transposase Tn15 transposition protein tnsD

Transposase Tn15 transposition protein tnsE

Transposase Tn15 transposition protein tnsE

TRAP transporter solute receptor TAXI family precursor

TRAP transporter solute receptor unknown substrate

TRAP-type C4-dicarboxylate transport system, periplasmic

TRAP-type C4-dicarboxylate transport system, large permease

TRAP-type C4-dicarboxylate transport system, periplasmic
| Gene       | Description                                                                 |
|------------|-----------------------------------------------------------------------------|
| TRAP       | TRAP-type transport system, small permease component                         |
| rRNA       | rRNA (5-methylaminomethyl-2-thiourate)-methyltransferase                     |
| rRNA       | rRNA nucleotyltransferase (EC 2.7.7.21) (EC 2.7.7.25)                        |
| rRNA       | rRNA threonylcarbamoyl adenosine modification protein YjeE                   |
| rRNA       | rRNA(Ile)-lysine synthetase                                                 |
| tRNA       | tRNA (5-methylaminomethyl-2-thiourylate)-methyltransferase (EC 2.1.1.61)    |
| tRNA       | tRNA nucleotyltransferase (EC 2.7.7.21) (EC 2.7.7.25)                       |
| tRNA       | tRNA threonylcarbamoyl adenosine modification protein YjeE                   |
| tRNA       | tRNA(Ile)-lysine synthetase                                                 |
| tRNA       | tRNA(Ile)-lysine synthetase, N-terminal domain                              |
| Trypsin-Like Proteases | Trypsin-like serine proteases, typically periplasmic |
| Tryptophan Synthetase | Tryptophan synthase beta chain like (EC 4.2.1.20) |
| Tungsten-Containing Enzyme | Tungsten-containing aldehyde:ferrodoxin oxoreductase (EC 1.2.7.5) |
| Tungsten-Containing Enzyme | Tungsten-containing aldehyde:ferrodoxin oxoreductase (EC 1.2.7.5) |
| Twin-arginine translocation protein TatB | Twin-arginine translocation protein TatB |
| Twin-arginine translocation protein TatC | Twin-arginine translocation protein TatC |
| Two-component HyrBR | Two-component hyrBR sensor and regulator |
| Two-component HyrBR | Two-component hyrBR sensor and regulator |
| Two-component HyrBR | Two-component HyrBR sensor and regulator |
| Type I secretion system ABC transporter HlyB family | Type I secretion system ABC transporter HlyB family |
| Type II R/M system | Type II R/M system (EC 2.1.1.72) |
| Twin-arginine translocation protein TatB | Twin-arginine translocation protein TatB |
| Twin-arginine translocation protein TatC | Twin-arginine translocation protein TatC |
| Two-component hyrBR sensor and regulator | Two-component hyrBR sensor and regulator |
| Two-component hyrBR sensor and regulator | Two-component hyrBR sensor and regulator |
| Two-component hyrBR sensor and regulator | Two-component hyrBR sensor and regulator |
| Type I secretion system ABC transporter HlyB family | Type I secretion system ABC transporter HlyB family |
| Type II R/M system | Type II R/M system (EC 2.1.1.72) |
| Twin-arginine translocation protein TatB | Twin-arginine translocation protein TatB |
| Twin-arginine translocation protein TatC | Twin-arginine translocation protein TatC |
| Two-component hyrBR sensor and regulator | Two-component hyrBR sensor and regulator |
| Two-component hyrBR sensor and regulator | Two-component hyrBR sensor and regulator |
| Two-component hyrBR sensor and regulator | Two-component hyrBR sensor and regulator |
| Type I secretion system ABC transporter HlyB family | Type I secretion system ABC transporter HlyB family |
| Two-component hyrBR sensor and regulator | Two-component hyrBR sensor and regulator |
| Two-component hyrBR sensor and regulator | Two-component hyrBR sensor and regulator |
| Type I secretion system ABC transporter HlyB family | Type I secretion system ABC transporter HlyB family |

The table continues with similar information, listing various enzymes and proteins with their associated EC numbers and functions. The table is extensive, covering a wide range of biological processes and mechanisms.
VCBS repeat

Vitamin B12 ABC Transporter permease component BtuC

Xanthine and CO dehydrogenases maturation factor, XdhC/CoxF family

Xanthine and CO dehydrogenases maturation factor, XdhC/CoxF family

Xanthine and CO dehydrogenases maturation factor, XdhC/CoxF family

Xanthine dehydrogenase iron-sulfur subunit (EC 1.17.1.4)

Xanthine dehydrogenase, molybdenum binding subunit (EC 1.17.1.4)

YbbL ABC transporter ATP-binding protein

YefM protein (antitoxin to YoeB)

YoeB toxin protein

Zinc ABC Transporter periplasmic-binding protein ZnuA

Zinc resistance-associated protein

Zn-dependent protease with chaperone function PA4632