Genomic Diversity of *Pseudoalteromonas atlantica* from Geographically Distant Deep Marine Basins

Erika N. Youngquist, Shane M. Hagen, Stephen M. Techtmann, Terry C. Hazen

Abstract

Microbes are extremely diverse and capable of catalyzing many functions. Currently, microbial species are defined though characterization of the microbe's metabolism and through sequencing of 16S rRNA. Recent studies suggest that microbes classified as the same species through 16S rRNA sequencing may demonstrate high genetic diversity. The goal of the study is to identify how much genomic diversity exists within a single species from different deep marine locations. Differences in the physical and chemical parameters of these locations may select for particular populations of microbes capable of growing in these environments. We isolated over 100 bacteria strains from the coasts of Australia, Angola, and Bermuda. Out of those isolates, we chose 32 *Pseudoalteromonas atlantica* isolates that have greater than 99% 16S rRNA identity. We sequenced the entire genomes of the 32 isolates and compared them to better understand the genomic heterogeneity within this species. We found that the *P. atlantica* isolates' 16S rRNA gene sequences were more than 99% similar, while the genomic makeup was much less than 99% similar. This work will demonstrate the utility of performing whole genome sequencing to differentiate between closely related taxa.

Introduction

Microbial taxa are often identified using phenotypic and genetic methods. Most commonly, taxonomic identification involves the sequencing of the 16S rRNA gene. The 16S rRNA gene sequence is used for bacterial species' taxonomic identification because it is a highly conserved genetic sequence found universally in bacteria for protein synthesis (5). Since it is highly conserved, it can be used to identify the genus and even species of bacteria. Thus, similarity of the 16S rRNA gene of two isolates often corresponds to taxonomic groupings (5). Bacteria are typically grouped into the same species using the cut off of 97% sequencing identity of the 16S rRNA gene (1, 2). Very closely related bacteria have been found in different environments throughout the world's oceans. The environmental parameters of these locations can vary greatly from location to location. Little is known about genetic difference between very closely related bacteria as well as how adaptation to growth under different environmental conditions affects the genomic makeup of microbes. However, previous studies revealed phenotypic and genotypic diversity of closely related species might occur (3, 4). Techtmann, et al. (2016) isolated oil-degrading *Colwellia psychrerythraea* from distant basins and isolates revealed great genetic differences after whole genome sequencing (6). Because of the previous *C. psychrerythraea* findings, we sought to investigate the genomic diversity of another oil-degrading species, *P. atlantica*, from distant deep marine locations. *Pseudoalteromonas* genus consists of gram-negative marine bacteria that are known to use hydrocarbon substrates and require sodium for growth (1, 7). The oil-degrading bacteria are of interest because of their ability to be used as a bioremediation solution for oil spills, such as the *Deepwater Horizon* spill of 2010 (4). Oil degrading bacteria such as *P. atlantica* and *C. psychrerythraea* can be crucial

for expediting the remediation of potential drilling accidents in the future (4). We found 32 *P. atlantica* strains out of the 100 isolates taken from three distant deep-sea basins. The Great Australian Bight (GAB) basin samples were taken south of Australia, the Western Atlantic basin samples were taken off the coast of Bermuda, and the Eastern Atlantic basin samples were taken off the coast of Angola. We wanted to determine if there is a high amount of genomic diversity between the *P. atlantica* isolates and if the physical and chemical factors of the environment contribute to the genetic differences. To do this, we sequenced the whole genome of 10 *P. atlantica* isolates and compared the Pan-genome tree results to the Phylogenetic tree results. Furthermore, we compared the optimal growth temperature, along with the Carbon, Nitrogen, and Osmolyte metabolisms of each *P. atlantica* isolate in this experiment. We hypothesize that the 16S rRNA and whole genome results will reveal significant genetic diversity between the *P. atlantica* isolates. In addition, we believe that the different physical and chemical environmental factors are influencing the genomic differences seen in *P. atlantica*. This work will help us to better understand how diverse a bacterial species is and how the environment is contributing to driving these differences.

Methods

Sampling

The water samples were taken from south of Australia in the Great Australian Bight basin, off the coast of Bermuda from the Western Atlantic basin, and off the coast of Angola from the Eastern Atlantic basin (Figure 1). These basins were selected because of their great distance from one another and their current patterns. These factors make the three basins isolated and ideal for our study. At each sampling site within the three basins, the temperature, salinity, and depth of the site were taken using the conductivity temperature depth device, which is a standard marine sampling device. These characteristics were recorded in tables located in the results section.

Isolations and DNA Extraction

The water samples taken from the different sites at each basin were grown on Marine Broth agar medium with crude oil. Then we picked colonies with similar morphologies from the agar and streaked them on either ONR7A or Marine Broth agar medium with crude oil. ONR7A media was chosen due to *P. atlantica's* optimal ability to metabolize it in previous studies (6). After growth on agar, we picked 99 isolated colonies to grow in their respective liquid media. Then we extracted DNA for each isolate after high optical densities were observed. DNA was extracted using the MoBio Ultraclean DNA extraction kit. To measure the quality of DNA and roughly estimate the DNA concentration, we measured the 260/280 ratio on a Nanodrop spectrophotometer (7).

Sequencing of the 16S rRNA gene

Polymerase Chain Reaction (PCR) was then performed to amplify the 16S sequence of each isolate using 1492R and 27F primers. We ran a gel electrophoresis on the PCR products to ensure the 16S gene was successfully amplified. Bands were compared to a ladder to observe an approximate size of 1,500 bp for the amplified product. The PCR product was cleaned using the zymo DNA clean and concentrate kit. DNA concentration of the product was determined using Qubit. Then the PCR product was sent to Walter's Life Science sequencing facility to obtain the 16S forward and reverse nucleotide sequences. We uploaded the sequence files in the application called Sequencher to assemble the forward and reverse reads into a contig. We edited the 16S contig of each isolate with Sequencher and then BLASTed the final edited sequence in the NCBI database using the nucleotide blast function. Once the identity of each isolate was confirmed, we chose 32 *Pseudoalteromonas* isolates that had greater than 99.5% 16S rRNA identity to one another.

Genome sequencing of isolates

Next, we prepared the 32 isolates for whole genome sequencing by beginning library prep. First, we performed high sensitivity Qubit readings on the genomic DNA to determine if the samples needed to be diluted. Once we obtained desired concentrations genome libraries were prepared using the Illumina NexteraXT kit. Genomic DNA was tagmented to break the DNA into small fragments. The fragmented DNA was PCR amplified to increase the concentration and add MiSeq adaptors and barcodes. Then we cleaned the samples with magnetic beads and ran a high sensitivity DNA assay with the Bioanalyzer. A few of the isolates' Bioanalyzer

data showed their DNA was not concentrated enough, so we concentrated the samples by evaporation using a rotovap. Next, we rechecked the concentrations with the Qubit and converted the ng/μL readings to nanomol. The samples were then pooled together and the Qubit concentration of the pool was checked and the final nanomol was determined. From this point on, the MiSeq protocol for whole genome sequencing was followed to clean the machine and finish prepping the pool.

Phylogenetic Tree

A phylogenetic tree was also prepared to compare the phylogeny of these isolated using 16S rRNA sequence matches of all the isolates. First, the edited 16S sequences were pasted into TextWrangler to prepare a fasta file. Then, the fasta file of all the 16S rRNA sequences were aligned using muscle in MEGA6 (8, 11). Next, a neighbor-joining phylogenetic tree was made from the aligned sequences (11). The neighbor-joining phylogenetic tree was saved as a Newick file and then visualized in iTOL to create the final phylogenetic tree.

Pan-genome tree

The complete genome sequences from Illumina MiSeq readings were saved and assembled using SPAdes to form contigs of the genomes. Then, the contigs were compared to each other and to other known Pseudoalteromonas species (14). The files were uploaded into CMG Biotools to create a pan-genome tree (14). Multiple outliers' whole genome sequences were also uploaded from NCBI database and included as a known species for comparison.

Isolate Selection

The 10 *P. atlantica* isolates used for phenotypic experimentation were chosen from the 32 *Pseudoalteromonas* isolates in the Pan-genomic tree in Figure 2. The isolates were chosen in order to capture an isolate from each basin from all of the different branches of the Pan-genomic tree. This would allow us to see what physical and genetic characteristics caused the isolates from the same basin to fall into different branches of the whole genome tree. Furthermore, it can reveal similar properties that isolates from different basins share. In the first branch of the tree in Figure 1, GAB NS 16 A, GAB NB 9D, GAB 2/3 16 C, Angola 30, and Angola 22 were the isolates chosen. There were no Bermuda (CST) isolates that fell within the first branch (Figure 2). For the second major branch, there were many branch points and branches within it; therefore, at least one isolate was taken from each of the smaller clades. Angola 9 and 7 were chosen because they flanked the large and only group of CST isolates (Figure 2). These isolates could help determine the physical similarities between the Angola and Bermuda isolates and reveal the differences between all of the Angola isolates. Lastly, GAB NS 16C was chosen since it was one of two Australia isolates that was part of its own clade (Figure 2).

Optimal Temperature Growth Curves

After selecting the 10 isolates, the glycerol freezer stock solution of each was used to grow a fresh culture to perform optimal temperature growth curves. The isolates were grown in Marine Broth at 4, 8, 12, 16, and 21°C. These temperatures were selected based on the temperatures the isolates were pulled from in their respective basins as shown in Table 1. Moreover, the temperatures were chosen to

allow a wide range for the isolates to grow in and to help better determine their optimal temperature growth range. Each isolate was grown in 10 mL of Marine broth. After the cells reached a log phase between 0.300 and 0.400 optical density of 600 nm, 1mL of the culture was pipetted into 9 mL of fresh media for a 10% dilution. Once in the fresh media, absorbance (optical density of 600 nm) readings were recorded at a set time interval by using a spectrophotometer, and there were three replicates of each isolate. The spectrophotometer was blanked with non-inoculated Marine broth frequently to ensure the machine was well calibrated. Once a growth curve was generated for each isolate at the five different temperatures, the isolates were grown in fresh media and then 1 mL of culture was pulled at the time they reached log phase. The 1 mL of log-phase culture was then transferred to 9 mL of fresh media and final growth curve readings were taken for each isolate in each temperature with three replicates. For both sets of growth curve readings, a large batch of each isolate was grown so the starting optical density would be the same for all temperature readings. The final optimal growth curve results were graphed to determine the optimal temperature for each isolate and the graphs were included in results.

Isolate	Depth (m)	Temperature °C	Salinity
1: GAB NS 16A	200	12.57	35.12
2: GAB NS 16C	200	12.57	35.12
3: GAB 2/3 16C	1650	2.62	34.62
4: GAB NB 9D	1901	2.40	34.69
8: Angola 30 19 2/3	850	4.90	34.60
7: Angola 22 24 2/3	960	4.36	34.62
6: Angola 9 19b	1250	4.20	34.80
5: Angola 7 24b	1450	3.92	34.93
10: CST 1 05-1200	1197	5.62	35.06
9: CST 2 03-5000	5005	2.12	34.85

Table 1: This table shows the basin, depth, temperature, and salinity each isolate was pulled from in the ocean. The measurements were taken with the conductivity temperature depth device. The temperatures in this chart were used to determine which temperatures were selected for the optimal growth curves.

Phenotype Microarray (PM) Plates

The PM plates tested on the isolates were Carbon PM1, Nitrogen PM3B, and Osmolyte PM9. A glycerol stock sample of each of the 10 isolates was used to grow fresh cultures for PM plate analysis at 14°C. The cultures were grown up to log phase, approximately .300 to .350 OD, and then 27mL of ONR7A media, 3 mL of culture, and 300µL of dye A was mixed for each isolate to have 2 replicates of each metabolic source. 100 µL of the mixture was pipetted into each well of the PM plates using a multipipetter. Once the PM plates were inoculated, they incubated for one week at 14°C. The plate concentrations were read with the Biolog Omnilog, which is a high throughput assay for screening metabolism on many substrates at once to compare and contrast the phenotypes (3). The plates were read at four separate time plots; time 0, time 24 hours, time 72 hours, and 1 week to ensure any delayed growth was detected. The plates were also read visually with dye A at each time

point, and any detection of purple in the wells was considered a positive metabolic source.

Results

P. atlantica isolates were sampled from distant basins

The three basins selected for sampling are very distant sites. The water samples were taken south of Australia from the Great Australian Bight basin, off the coast of Bermuda from the Western Atlantic basin, and off the coast of Angola from the Eastern Atlantic basin as shown in the map below.

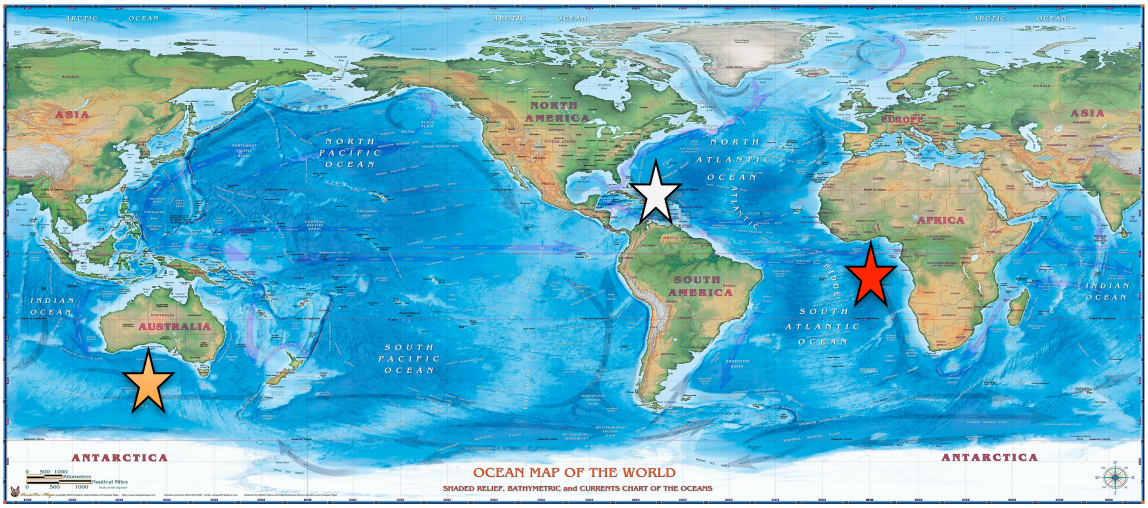

Figure 1: The stars on the map indicate the sampling sites with orange representing the Great Australian Bite basin, white representing the Western Atlantic basin, and red representing the Eastern Atlantic basin.

Isolates were obtained from environments with different geochemistry

The depth of the samples collected from the three different basins varies greatly from 200 m to over 5,000 m. In addition, the temperature of the sites ranged from 2°C to 12°C. There were also significant dissolved oxygen differences between the sites. However, only minor salinity differences were observed.

Isolate	Basin	Depth (m)	Temp (°C)	Salinity	Dissolved Oxygen
CST-03-5000	Western Atlantic	5005.049	2.1253	34.8582	8.18
CST-05-1200	Western Atlantic	1197.356	5.6213	35.0644	6.81
GAB-SS-09-NB	Great Australian Bight	1901	2.398	34.691	5.49
GAB-SS-16-NS	Great Australian Bight	200	12.574	35.123	7.98
GAB-SS-16-2/3	Great Australian Bight	1650	2.621	34.624	5.45
19-2/3	Eastern Atlantic (Angola)	850	4.9	34.6	2.18
19-NB	Eastern Atlantic (Angola)	1250	4.2	34.8	2.89
24-2/3	Eastern Atlantic (Angola)	960	4.366	34.618	3.06
24-NB	Eastern Atlantic (Angola)	1450	3.925	34.93	4.48

Table 2: Geochemical parameters of sampling sites. Temperature and depth vary greatly between each site. The measurements were taken with the conductivity temperature depth device.

P. atlantica isolates are present in distant deep sea basins

Of the over 100 isolates that were obtained, 32 isolates were classified as *P. atlantica.* The 32 *P. atlantica* isolates chosen were found to have greater than 99.5% 16S rRNA identity, despite the fact that they were obtained from three geographically distant basins. The 12 isolates from Angola consisted of 2 from site 24 near bottom (1450 m), 4 from site 24 at 2/3 depth (960 m), 3 from site 19 near bottom (1250 m), and 3 from site 19 at 2/3 depth (850 m). The 8 Bermuda isolates consisted of 7 from site 1,200 m and 1 from site 5,000 m, which was adjacent to the Puerto Rico Trench. The Australia samples were comprised of 9 isolates total with 2 isolates from station 16 2/3 depth (1650 m), 1 from station 9 near bottom (1901 m), and 6 from station 16 near surface (200 m). The 16S rRNA percent identity shows the *P. atlantica* isolates are the same species with little to no 16S rRNA genetic variance between the isolates from distant locations.

Isolate	16S Percent ID
Angola2_13	99.5
Angola2_18	99.64
Angola2_20	99.78
Angola2_22	99.93
Angola2_26	99.5
Angola2_27	99.79
Angola2_28	99.79
Angola2_30	99.86
Angola2_31	99.79
Angola2_32	99.93
Angola2_4	99.64
Angola2_7	100
CST_1	99.29
CST_2	99.07
CST_3	99.17

Isolate	16S Percent ID
CST_4	99.57
CST_5	99.22
CST_6	99.43
CST_7	99.64
CST_9	99.36
GAB_23_16A	99.86
GAB_23_16C	99.79
GAB_NB_9D	99.79
GAB_NS_16A	99.86
GAB_NS_16C	99.57
GAB_NS_16E	99.86
GAB_NS_16G	99.86
GAB_NS_16H	99.86
GAB_NS_9D	99.79

Table 3: 16S rRNA Percent Identity of *Pseudoalteromonas* isolates taken from three basins. All of the isolates have greater than 99.5 % 16S relatedness, indicating they are the same species, *P. atlantica.*

The P. atlantica isolates show no genetic diversity of 16S rRNA phylogeny

A phylogenetic tree was built with the 16S rRNA sequences of our *P. atlantica* isolates along with select non-*Pseudoalteromonas* spp. isolated from other environments. The 16S rRNA phylogenetic tree shows no genetic diversity exists between the *P. atlantica* isolates on the 16S genetic level.

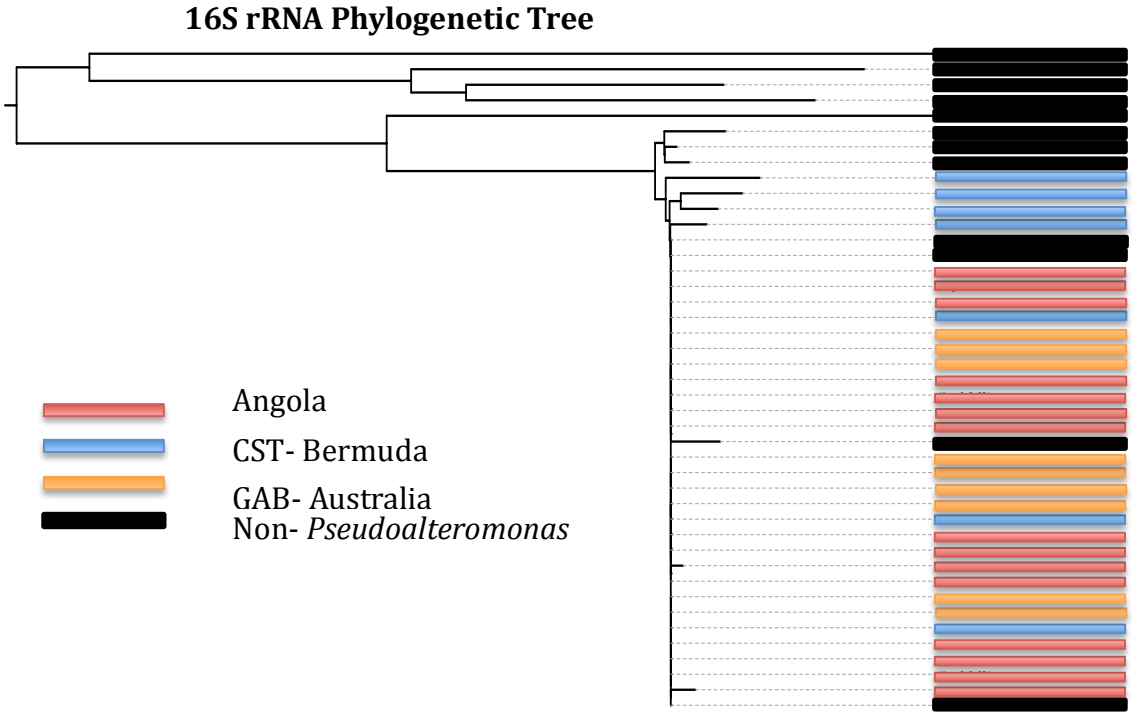

Figure 2: 16S rRNA Phylogenetic tree shows that there are no geographic patterns for the *P. atlantica* isolates. There are no significant genetic differences on the 16S rRNA level.

***Whole genome tree confirms more genetic diversity exists between P. atlantica* isolates**

After creating the final pan-genomic tree and comparing the results to the 16S rRNA phylogenetic tree, we concluded a much higher amount of genetic diversity exists between the *P. atlantica* isolates. 10 of the *P. atlantica* isolates were chosen from the whole genome tree as specified in the methods for further phenotypic testing of optimal temperature and metabolism.

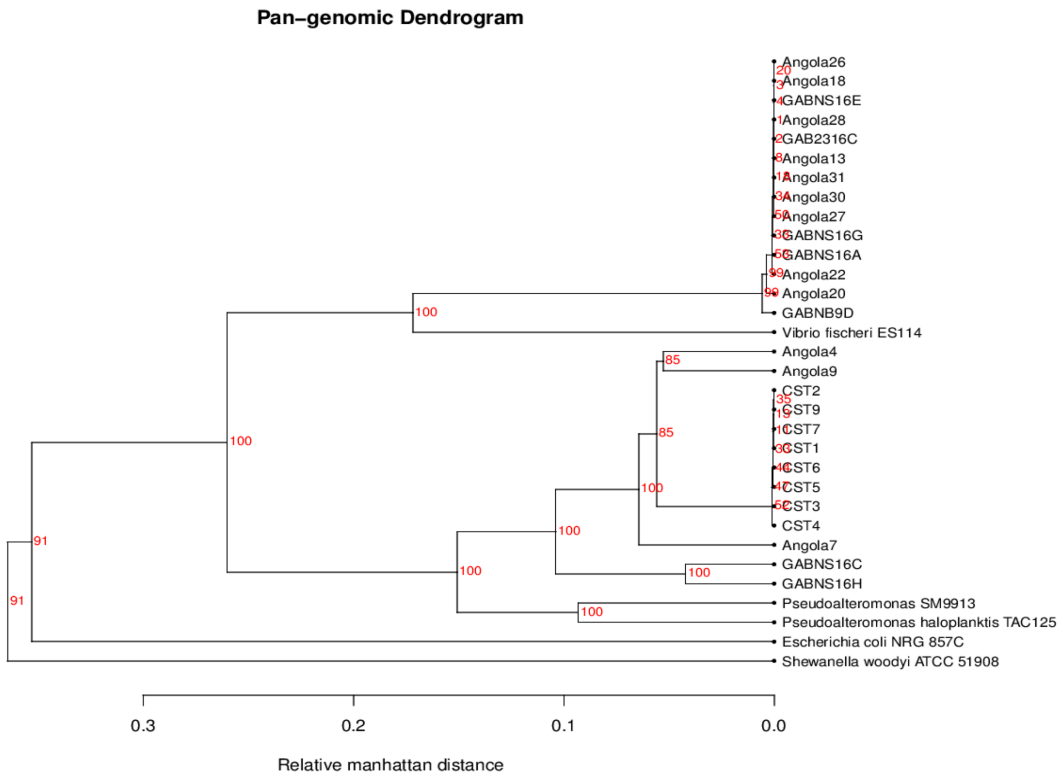

Figure 3: This Pan-genomic tree represents the 32 *P. atlantica* isolates that were compared to their 16S rRNA phylogenetic tree. From this whole genome tree, 10 isolates were selected that fell into different clades in order to determine the phenotypic differences that account for the genomic diversity.

Optimal temperature growth curve results reveal optimal growth range for all isolates is between 12°C and 21°C

Each graph includes the five separate temperature growth curves for one isolate. The results reveal that all isolates preferred a temperature range between 12°C and 21°C. Since all isolates preferred the same temperature range, it is harder to conclude any major definitive genomic differences that might correlate with preferred temperature. It is important to note, however, that some isolates were able to grow very well in temperatures that other isolates were not such as 4°C and 8°C. These phenotypic differences might explain some minor genomic differences between the isolates.

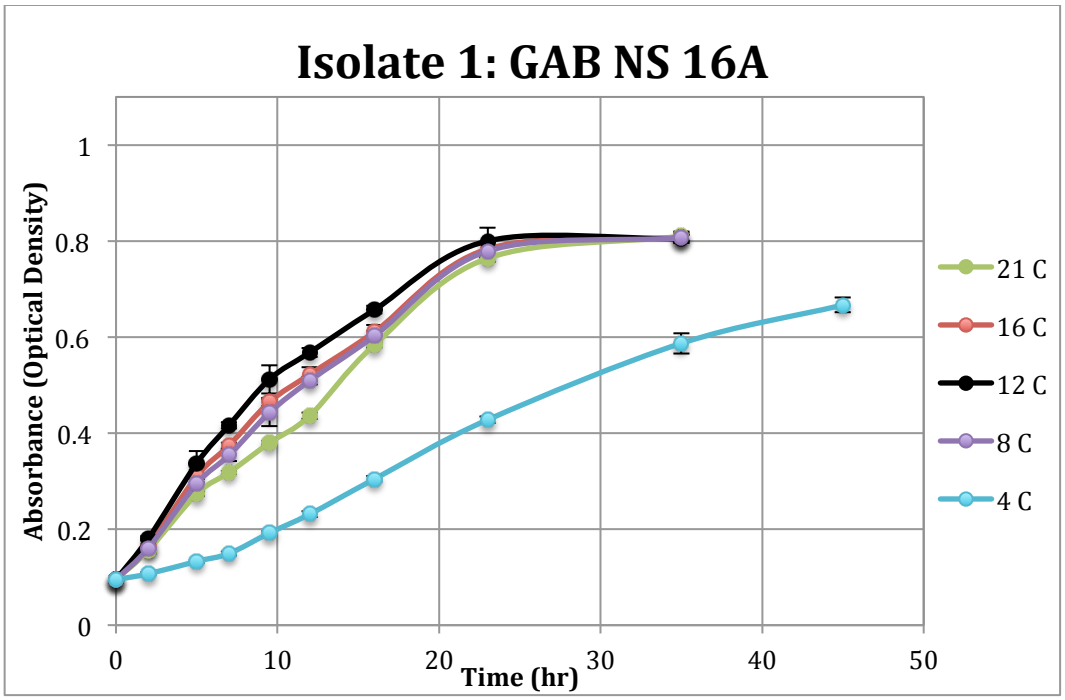

Figure 4: Optimal growth temperature for Isolate 1 is 12°C. Growth in 4°C was very slow for isolate 1 and did not reach a high OD. Growth in 8 °C was much better for Isolate 1 than many other isolates.

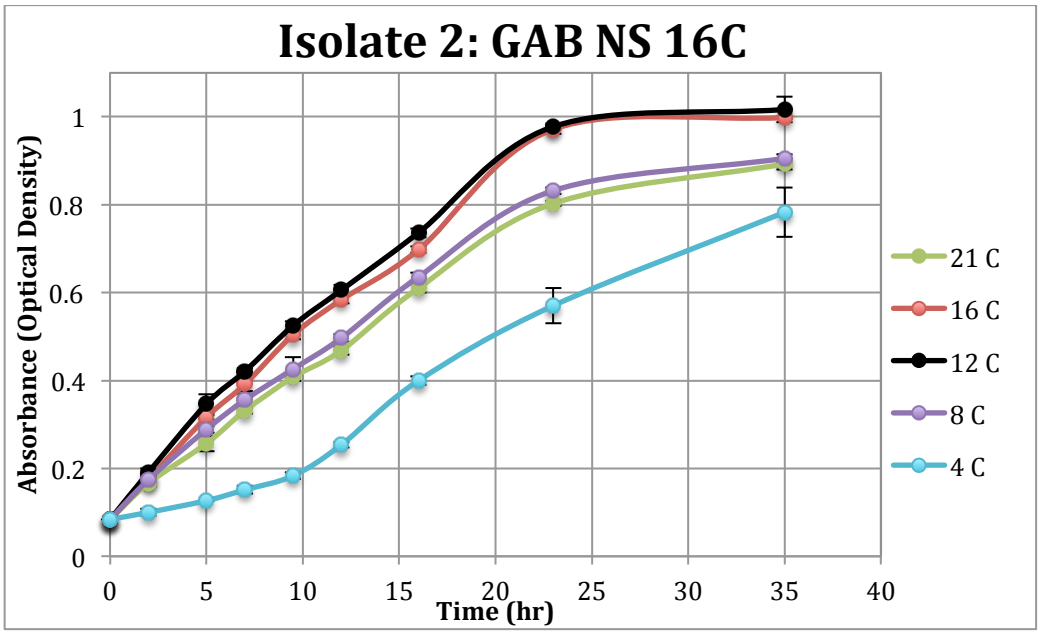

Figure 5: Optimal growth temperature for Isolate 2 is 12°C. Growth in 4°C was much slower and did not reach a high OD; Growth in 8°C was much better for Isolate 2 than most isolates.

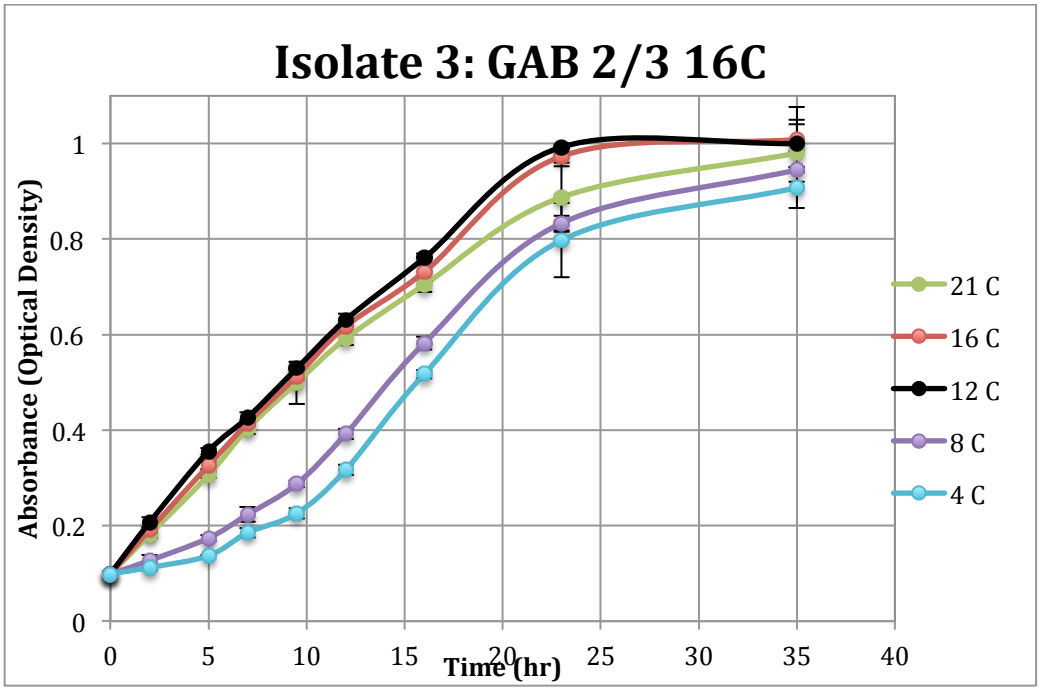

Figure 6: Optimal growth temperature for Isolate 3 is 12°C. Growth in 4°C and 8°C is much better than most isolates.

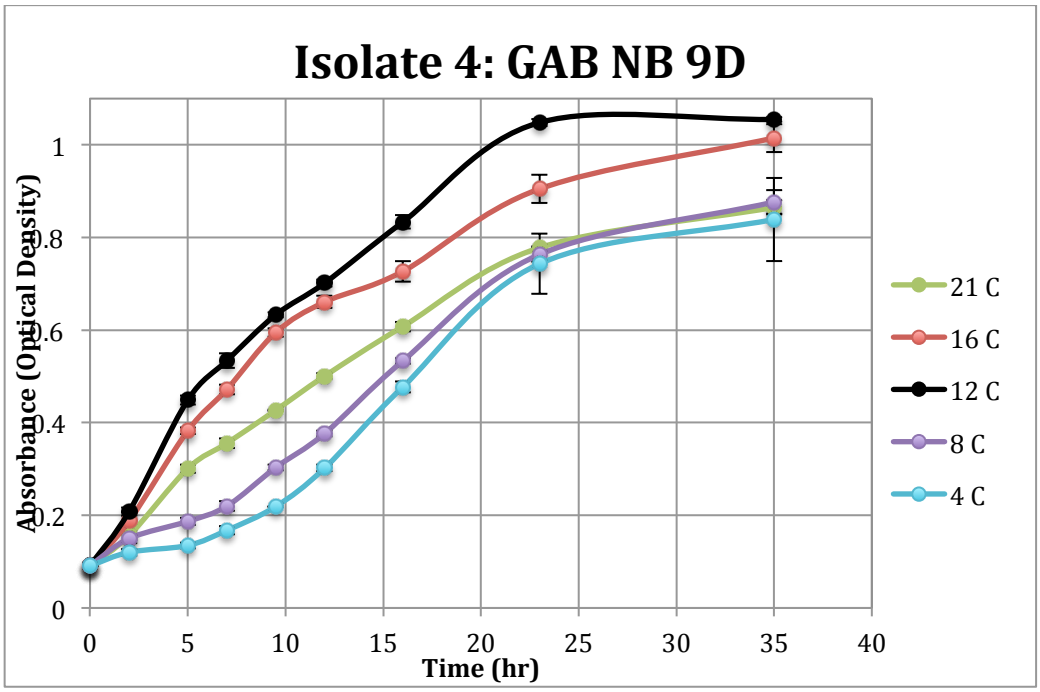

Figure 7: Optimal growth temperature for Isolate 4 is 12°C. Growth in 16°C was not as high as other Isolates. Isolate 4 shows the strongest affinity for 12°C compared to the other temperatures it was grown in and the growth curves of all the other isolates.

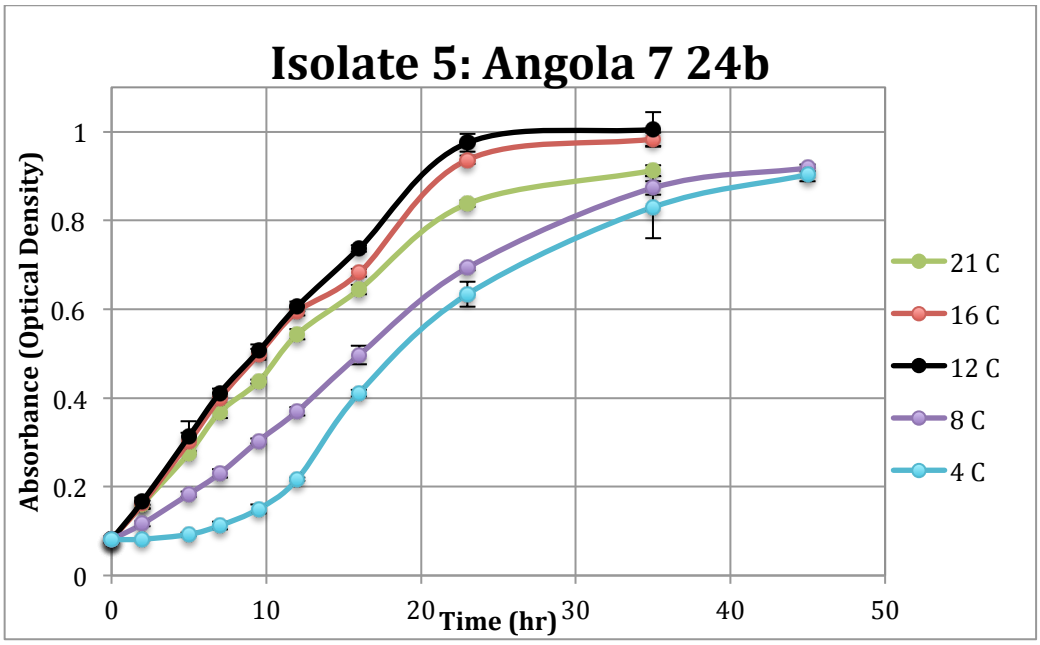

Figure 8: Optimal growth temperature for Isolate 5 is 12°C with 16°C close behind. Growth in all temperatures was strong for Isolate 5.

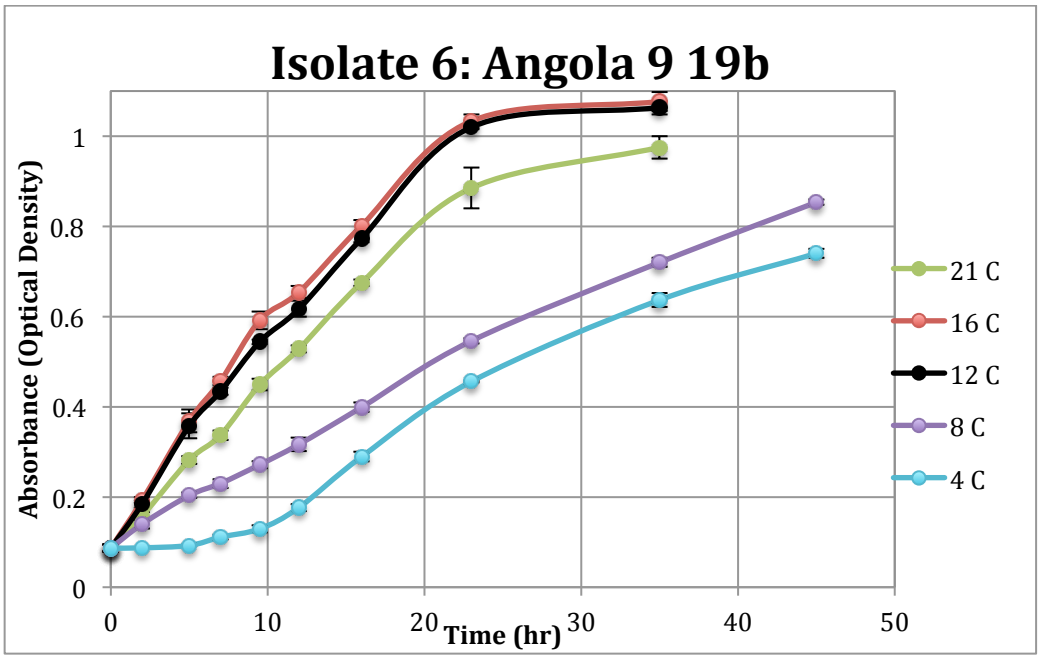

Figure 9: Optimal growth temperature for Isolate 6 is 16°C. Growth in 4 and 8°C is slow for Isolate 6.

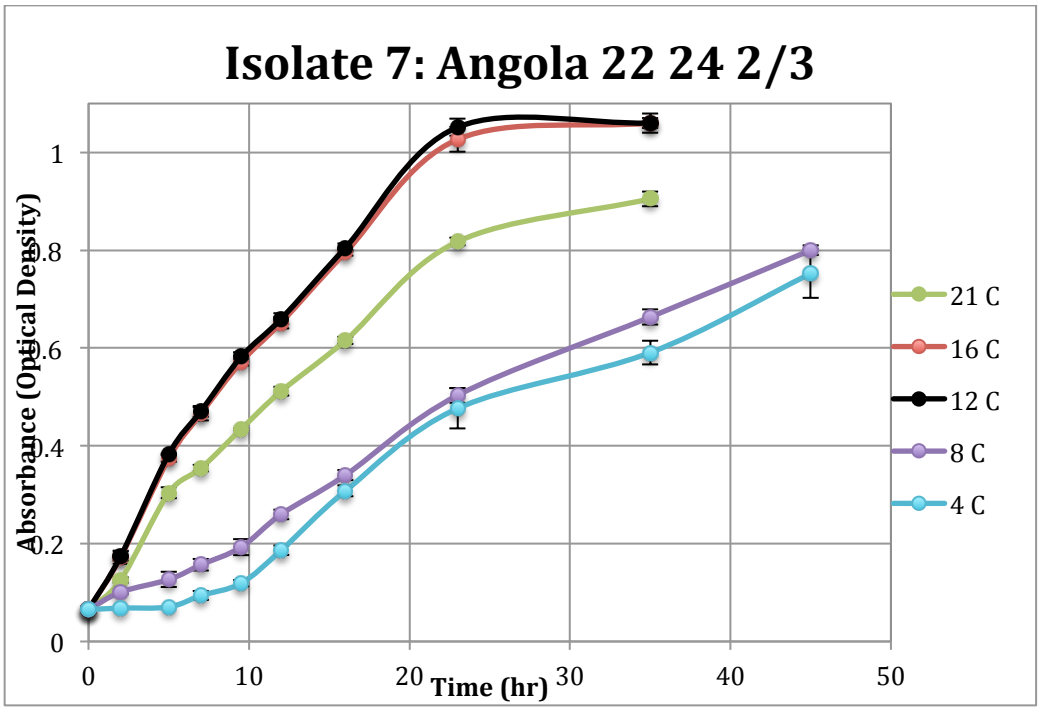

Figure 10: Optimal growth temperature for Isolate 7 is between 12°C and 16°C . Growth in 4 and 8°C is slow for Isolate 7.

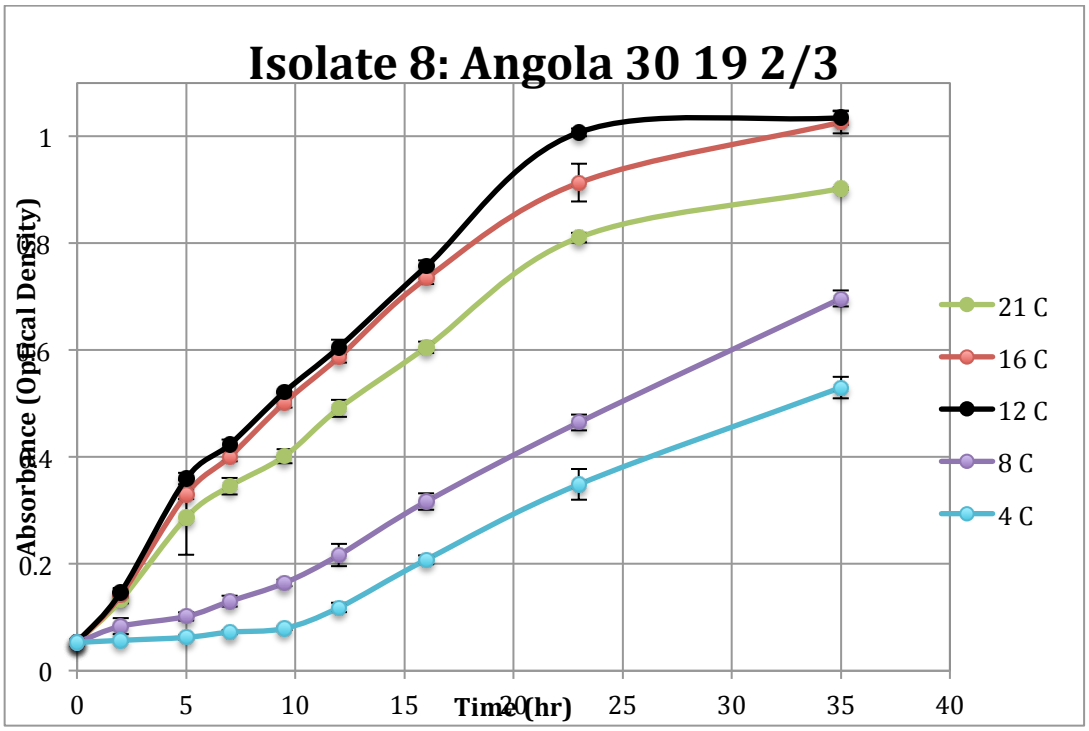

Figure 11: Optimal growth temperature for Isolate 8 is 12°C with 16°C very close behind. Growth in 4 and 8°C is very slow for Isolate 8.

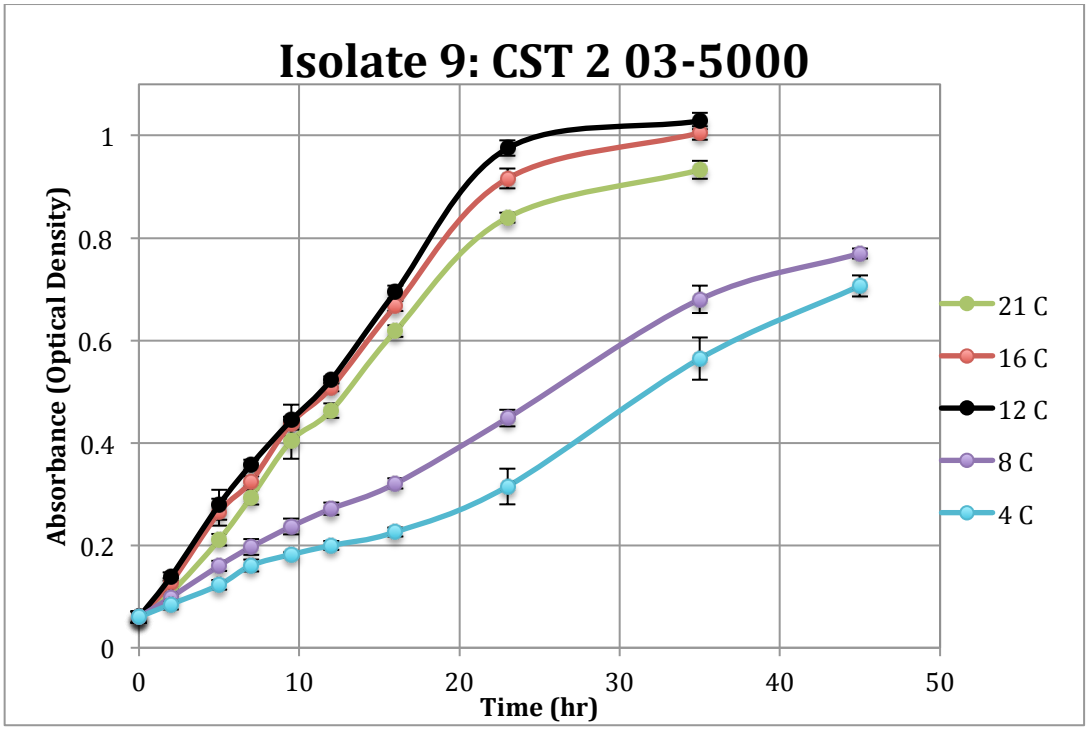

Figure 12: Optimal growth temperature for Isolate 9 is between 12°C with 16°C very close behind. Growth in 4 and 8°C is slow for Isolate 9.

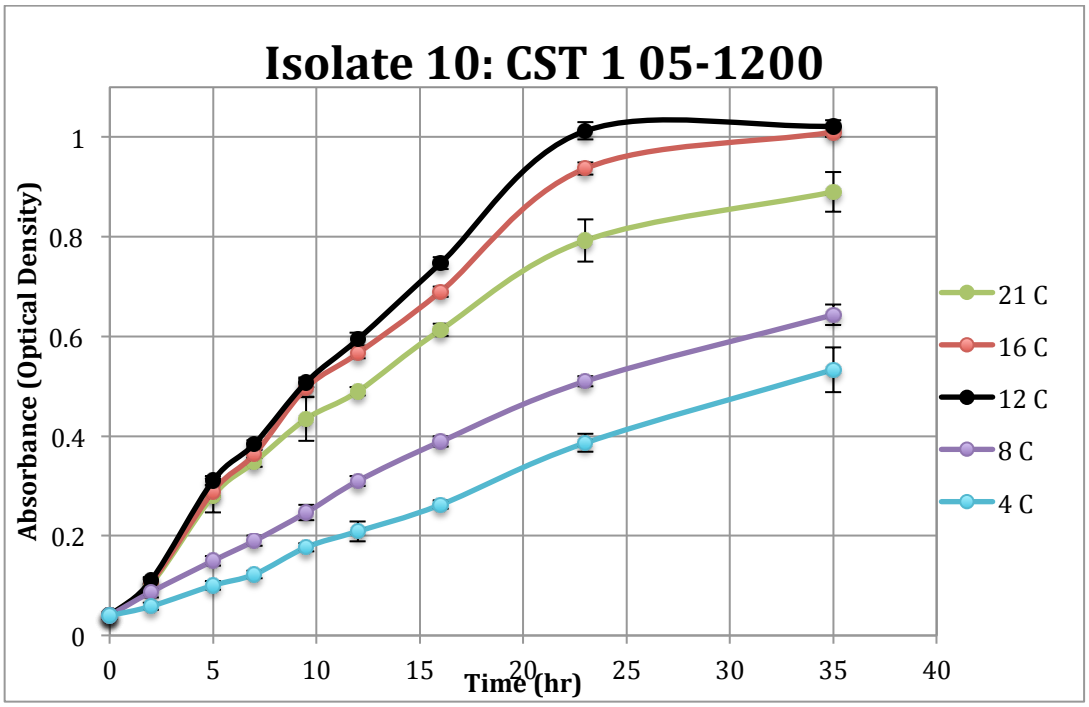

Figure 13: Optimal growth temperature for Isolate 10 is between 12°C. Growth in 4 and 8°C is slow for Isolate 10.

Phenotype Microarray of Carbon, Nitrogen, and Osmolyte sources reveal great metabolic diversity between P. atlantica isolates

The Phenotype Microarray (PM) plate results show the differences and similarities of metabolism between the 10 isolates in Carbon, Nitrogen, and Osmolyte sources. All 10 *P. atlantica* isolates showed differences in their ability to metabolize the three nutrient sources. These differences can account for some of the genetic diversity of the isolates since each metabolic pathway has its own unique genetic sequence.

Carbon PM1 Plate

A1 Negative Control	A2 L-Arabinose	A3 N-Acetyl-DGlucosamine	A4 D-Saccharic Acid	A5 Succinic Acid	A6 D-Galactose	A7 L-Aspartic Acid	A8 L-Proline	A9 D-Alanine	A10 D-Trehalose	A11 D-Mannose	A12 Dulcitol
B1 D-Serine	B2 D-Sorbitol	B3 Glycerol	B4 L-Fucose	B5 D-GlucuronicAcid	B6 D-Gluconic Acid	B7 D,LaGlycerolPhosphate	B8 D-Xylose	B9 L-Lactic Acid	B10 Formic Acid	B11 D-Mannitol	B12 L-Glutamic Acid
C1 D-Glucose-6Phosphate	C2 D-GalactonicAcid-g-Lactone	C3 D,L-Malic Acid	C4 D-Ribose	C5 Tween 20	C6 L-Rhamnose	C7 D-Fructose	C8 Acetic Acid	C9 a-D-Glucose	C10 Maltose	C11 D-Melibiose	C12 Thymidine
D-1 L-Asparagine	D2 D-Aspartic Acid	D3 D-Glucosaminic Acid	D4 1,2-Propanediol	D5 Tween 40	D6 a-Keto-Glutaric Acid	D7 a-Keto-Butyric Acid	D8 aMethyl-DGalactoside	D9 a-D-Lactose	D10 Lactulose	D11 Sucrose	D12 Uridine
E1 L-Glutamine	E2 m-Tartaric Acid	E3 DGlucose1Phosphate	E4 DFructose6Phosphate	E5 Tween 80	E6 aHydroxyGlutaric Acid-g-Lactone	E7 a-Hydroxy Butyric Acid	E8 bMethyl-DGlucoside	E9 Adonitol	E10 Maltotriose	E11 2-Deoxy Adenosine	E12 Adenosine
F1 Glycyl-L-Aspartic Acid	F2 Citric Acid	F3 m-Inositol	F4 D-Threonine	F5 Fumaric Acid	F6 Bromo Succinic Acid	F7 Propionic Acid	F8 Mucic Acid	F9 Glycolic Acid	F10 Glyoxylic Acid	F11 D-Cellobiose	F12 Inosine
G1 Glycyl-LGlutamic Acid	G2 Tricarballylic Acid	G3 L-Serine	G4 L-Threonine	G5 L-Alanine	G6 L-Alanyl-Glycine	G7 Acetoacetic Acid	G8 N-Acetyl-bDMannosamine	G9 Mono Methyl Succinate	G10 Methyl Pyruvate	G11 D-Malic Acid	G12 L-Malic Acid
H1 Glycyl-L-Proline	H2 pHydroxyPhenyl Acetic Acid	H3 mHydroxyPhenyl Acetic Acid	H4 Tyramine	H5 D-Psicose	H6 L-Lyxose	H7 Glucuronamide	H8 Pyruvic Acid	H9 LGalactonicAcid-gLactone	H10 DGalacturonic Acid	H11 Phenylethylamine	H12 2-Aminoethanol

Figure 14: The differences in Carbon metabolism between the isolates is shown with blue indicating similarities where all isolates grew, orange indicating at least one or more isolate's metabolism was different, and white indicating no metabolism for any isolate.

Nitrogen PM3B Plate

A1 Negative Control	A2 Ammonia	A3 Nitrite	A4 Nitrate	A5 Urea	A6 Biuret	A7 L-Alanine	A8 L-Arginine A9	A9 L-Asparagine	A10 L-Aspartic Acid	A11 L-Cysteine	A12 L-Glutamic Acid
B1 L-Glutamine	B2 Glycine	B3 L-Histidine	B4 L-Isoleucine	B5 L-Leucine	B6 L-Lysine	B7 L-Methionine	B8 L-Phenylalanine	B9 L-Proline	B10 L-Serine	B11 L-Threonine	B12 L-Tryptophan
C1 L-Tyrosine	C2 L-Valine	C3 D-Alanine	C4 D-Asparagine	C5 D-Aspartic Acid	C6 D-Glutamic Acid	C7 D-Lysine	C8 D-Serine	C9 D-Valine	C10 L-Citrulline	C11 L-Homoserine	C12 L-Ornithine
D-1 NAcetylL Glutamic Acid	D2 NPhthaloylLGlutamic Acid	D3 L-Pyroglutamic Acid	D4 Hydroxylamine	D5 Methylamine	D6 N-Amylamine	D7 N-Butylamine	D8 Ethylamine	D9 Ethanolamine	D10 Ethylenediamine	D11 Putrescine	D12 Agmatine
E1 Histamine	E2 b-Phenylethylamine	E3 Tyramine	E4 Acetamide	E5 Formamide	E6 Glucuronamide	E7 D,L-Lactamide	E8 D-Glucosamine	E9 D-Galactosamine	E10 D-Mannosamine	E11 N-Acetyl-DGlucosamine	E12 NAcetyl DGalactosamine
F1 NAcetyl DMannosamine	F2 Adenine	F3 Adenosine	F4 Cytidine	F5 Cytosine	F6 Guanine	F7 GuanosineF8Thymine	F8 Thymine	F9 Thymidine	F10 Uracil	F11 Uridine	F12 Inosine
G1 Xanthine	G2 Xanthosine	G3 Uric Acid	G4 Alloxan	G5 Allantoin	G6 Parabanic Acid	G7 D,L-a-Amino-NButyric Acid	G8 gAmino-NButyric Acid	G9 eAmino NCaproic Acid	G10 D,L-a-Amino-Caprylic Acid	G11 d-Amino-NValeric Acid	G12 aAmino NValeric Acid
H1 Ala-Asp	H2 Ala-Gln	H3 Ala-Glu	H4 Ala-Gly	H5 Ala-His	H6 Ala-Leu	H7 Ala-Thr	H8 Gly-Asn	H9 Gly-Gln	H10 Gly-Glu	H11 Gly-Met	H12 Met-Ala

Figure 15: The differences in Nitrogen metabolism between the isolates is shown with blue indicating similarities where all isolates grew, orange indicating at least one or more isolate's metabolism was different, and white indicating no metabolism for any isolate.

Osmolyte PM9 Plate

A1 NaCl 1%	A2 NaCl 2%	A3 NaCl 3%	A4 NaCl 4%	A5 NaCl 5%	A6 NaCl 5.5%	A7 NaCl 6%	A8 NaCl 6.5%	A9 NaCl 7%	A10 NaCl 8%	A11 NaCl 9%	A12 NaCl 10%
B1 NaCl 6%	B2 NaCl 6% + Betaine	B3 NaCl 6% + N-N Dimethyl Glycine	B4 NaCl 6% + Sarcosine	B5 NaCl 6% +Dimethylsulphonylpropionate	B6 NaCl 6% + MOPS	B7 NaCl 6% + Ectoine	B8 NaCl 6% + Choline	B9 NaCl 6% + Phosphoryl Choline	B10 NaCl 6% + Creatine	B11 NaCl 6% + Creatinine	B12 NaCl 6% + L-Carnitine
C1 NaCl 6% + KCl	C2 NaCl 6% + L-Proline	C3 NaCl 6% +NAcetyl LGlutamin	C4 NaC1 6% +β-Glutamic Acid	C5 NaC1 6% +γ– Amino - NButyric Acid	C6 NaC1 6% + Glutathione	C7 NaCl 6% + Glycerol	C8 NaC1 6% + Trehalose	C9 NaC1 6% +Trimethylamine Noxide	C10 NaC1 6% + Trimethylamine	C11 NaCl 6% + Octopin	C12 NaC1 6% + Trigonelline
D1 Potassium chloride 3%	D2 Potassium chloride 4%	D3 Potassium chloride 5%	D4 Potassium chloride 6%	D5 Sodium sulfate 2%	D6 Sodium sulfate 3%	D7 Sodium sulfate 4%	D8 Sodium sulfate 5%	D9 Ethylene glycol 5%	D10 Ethylene glycol 10%	D11 Ethylene glycol 15%	D12 Ethylene glycol 20%
E1 Sodium formate 1%	E2 Sodium formate 2%	E3 Sodium formate 3%	E4 Sodium formate 4%	E5 Sodium formate 5%	E6 Sodium formate 6%	E7 Urea 2%	E8 Urea 3%	E9 Urea 4%	E10 Urea 5%	E11 Urea 6%	E12 Urea 7%
F1 Sodium Lactate 1%	F2 Sodium Lactate 2%	F3 Sodium Lactate 3%	F4 Sodium Lactate 4%	F5 Sodium Lactate 5%	F6 Sodium Lactate 6%	F7 Sodium Lactate 7%	F8 Sodium Lactate 8%	F9 Sodium Lactate 9%	F10 Sodium Lactate 10%	F11 Sodium Lactate 11%	F12 Sodium Lactate 12%
G1 Sodium Phosphate pH 7 20mM	G2 Sodium Phosphate pH 7 50mM	G3 Sodium Phosphate pH 7 100mM	G4 Sodium Phosphate pH 7 200mM	G5 Sodium Benzoate pH 5.2 20mM	G6 Sodium Benzoate pH 5.2 50mM	G7 Sodium Benzoate pH 5.2 100mM	G8 Sodium Benzoate pH 5.2 200mM	G9 Ammonium sulfate pH8 10mM	G10 Ammonium sulfate pH 8 20mM	G11 Ammonium sulfate pH 8 50mM	G12 Ammonium sulfate pH8 100mM
H1 Sodium Nitrate 10mM	H2 Sodium Nitrate 20mM	H3 Sodium Nitrate 40mM	H4 Sodium Nitrate 60mM	H5 Sodium Nitrate 80mM	H6 Sodium Nitrate 100mM	H7 Sodium Nitrite 10mM	H8 Sodium Nitrite 20mM	H9 Sodium Nitrite 40mM	H10 Sodium Nitrite 60mM	H11 Sodium Nitrite 80mM	H12 Sodium Nitrite 100mM

Figure 16: The differences in Osmolyte metabolism between the isolates is shown with blue indicating similarities where all isolates grew, orange indicating at least one or more isolate's metabolism was different, and white indicating no metabolism for any isolate.

Conclusions

The conflicting results from the 16S and Whole Genome Phylogenetic trees reveal that there is a higher amount of genetic diversity that exists within very closely related species. In the 16S rRNA tree, only the non-*Pseudoalteromonas* organisms showed genetic variance. However, the Whole Genome tree shows that all of the *P. atlantica* isolates from the three basins have large genetic differences. The Phenotype Microarray results confirm our hypothesis of great genetic diversity between isolates by showing the many metabolic differences of Carbon, Nitrogen, and Osmolyte sources. In addition, our hypothesis of different physical and chemical environmental factors influencing the genomic differences of *P. atlantica* isolates is confirmed. The metabolic differences can make up a large percentage of the whole genomic differences seen in the Pan-Genome tree because each metabolic source has a specific gene sequence that encodes its function. Since the *P. atlantica* isolates were found in different basins with varying environmental parameters and depths, it is likely that the isolates acquired different metabolic abilities through lateral gene transfer from DNA in their environment to increase their chances of survival (2, 13). These results indicate that 16S rRNA based approaches are insufficient to distinguish between closely related taxa; thus revealing the importance of whole genome sequencing in identifying unknown bacteria.

The optimal temperature range for all 10 isolates was determined to be between 12°C and 21°C, making the *P. atlantica* isolates psychrophiles. Even though their optimal growth temperature is between 12°C and 21°C, most of the isolates were pulled from much colder temperatures such as 2°C and 4°C. The isolates were

probably able to thrive in a less desirable environment because of an abundance of carbon sources and nutrients available in the deeper and colder parts of their basin. The optimal growth curve results reveal some phenotypic differences between the isolates for temperatures outside of their optimal growth range. For many isolates, growth in the colder temperatures of 4°C and 8°C was very slow, and the final optical density readings were not high, while other isolates were able to thrive much better in the colder temperatures. Genes that impact their membrane composition or metabolic rate may affect their ability to survive in colder temperatures. For example, a cold-active and salt-tolerant α-amylase gene has been found in marine organisms that are able to survive in temperatures as low as 0°C (12). Differences in these cold genes might also account for some of the genomic diversity between the isolates. Their ability to grow significantly at 4°C is another important adaption in these environments since oil has been shown to degrade faster aerobically in colder waters (9).

Future Directions

Further research should be done to determine which cold genes are impacting the isolates' ability to thrive better in colder environments, which may help explain some of the genomic diversity between the closely related species. More research should also be done to investigate the metabolic differences of the *P. atlantica* isolates. Specifically, we should determine the exact metabolic pathways that are accounting for the genomic differences between the isolates. Once the metabolic pathways are determined, we should identify the identical pathways of isolates from the same basin to determine if there are geographical patterns. These

patterns could further explain the differences of metabolic pathways between the isolates, since isolates from the same basin will be more likely to laterally transfer the same genes (13).

References

1. **Stackebrandt E, Goebel B.** 1994. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. International Journal of Systematic Bacteriology **44:**846-849.
2. **Konstantinidis KT TJ.** 2005. Genomic insights that advance the species definition for prokaryotes. . P Natl Acad Sci USA 2567-2572.
3. **Shapiro BJ FJ, Cordero OX, Preheim SP, Timberlake SC, Szabo G** 2012. Population Genomics of Early Events in the Ecological Differentiation of Bacteria. . Science**:**48-51.
4. **King GM, J. E. Kostka, T. C. Hazen and P. A. Sobecky.** 2015. Microbial Responses to the Deepwater Horizon Oil Spill: From Coastal Wetlands to the Deep Sea. Annual Review of Marine Science **7:**377-401.
5. **Abbott JMJSL.** 2007. 16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls. Journal of Clinical Microbiology **45**.
6. **Techtmann SM, Kathleen S Fitzgerald, Savannah C Stelling, Dominique C Joyner, Sagar M Uttukar, Austin P Harris, Noor K Alshibli, Steven D. Brown and Terry C Hazen.** 2016. Colwellia psychrerythraea strains from distant deep sea basins show adaptation to local conditions. Frontiers in Environmental Science
7. **Scientific TF.** 2009. Nanodrop Spectrophotometer.
8. **Tamura K SG, Peterson D, Filipski A, Kumar S** 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. . Mol Biol Evol 2725-2729.
9. **Hazen TC, E. A. Dubinsky, T. Z. DeSantis, G. L. Andersen, Y. M. Piceno, N. Singh, J. R. Jansson, A. Probst, S. E. Borglin, J. L. Fortney, W. T. Stringfellow, M. Bill, M. S. Conrad, L. M. Tom, K. L. Chavarria, T. R. Alusi, R. Lamendella, D. C. Joyner, C. Spier, M. Auer, M. L. Zemla, R. Chakraborty, E. L. Sonnenthal, P. D'haeseleer, H.-Y. N. Holman, S. Osman, Z. Lu, J. D. Van Nostrand, Y. Deng, J. Zhou, and O. U. Mason.** 2010. Deep-sea oil plume enriches psychrophilic oil-degrading bacteria. Science **330:204-208**
10. **Biolog.** 2013. GEN III MicroPlate Instructions for Use.

11. **G. Gauthier MG, R. Christen.** 1995. Phylogenetic Analysis of the Genera Alteromonas, Shewanella, and Moritella Using Genes Coding for Small-Subunit rRNA Sequences and Division of the Genus Alteromonas into Two Genera, Alteromonas (Emended) and Pseudoalteromonas gen. nov., and Proposal of Twelve New Species Combinations. International Journal of Systematic and Evolutionary Microbiology.
12. **Qin Y HZ, Liu Z.** 2014. A novel cold-active and salt-tolerant α-amylase from marine bacterium Zunongwangia profunda: molecular cloning, heterologous expression and biochemical characterization. PubMed.
13. **Howard Ochman JGLEAG.** 2000. Lateral gene transfer and the nature of bacterial innovation. Nature**:**299-304.
14. **Nurk S BA, Antipov D, Gurevich A, Korobeynikov A, Lapidus A.** 2013. Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads. . Research in Computational Molecular Biology 158-170.