AN ABSTRACT OF THE THESIS OF

Brett L. Mellbye for the degree of Honors Baccalaureate of Science in Microbiology presented on June 2, 2006. Title: A Genomic and Proteomic Characterization of the First Cultured Oligotrophic Marine *Gammaproteobacterium* from the SAR92 Clade.

Abstract approved: _____

Stephen J. Giovannoni

High-throughput culturing (HTC) consisting of extinction culturing in autoclaved seawater has led to the isolation and characterization of many novel strains of oligotrophic marine bacteria. Strain HTCC 2207 was isolated from the Oregon coast by the HTC method. Phylogenetic analysis based on 16S rRNA gene sequence showed that this strain fell into the SAR92 clade in the oligotrophic marine *Gammaproteobacteria* (OMG) group. The OMG group is distantly related to previously cultivated genera of *Gammaproteobacteria*. Initial phylogenetic characterization was followed by genome sequencing and interpretation, proteomic analysis by liquid chromatography/tandem mass spectrometry, and determination of the fatty acid profile.

Culture experiments, microscopic observations, and the genome sequence indicate that HTCC 2207 cells are motile, aerobic, heterotrophic, Gram-negative, short rods of approximately 0.148 μm^3. Growth characteristics were observed at six different carbon concentrations and five different temperatures. Optimal growth rate (3.15 d^{-1}) occurred at 16 ^oC in natural seawater amended with nitrogen, phosphorus, vitamins, and a mixture of organic carbon compounds yielding a maximum cell density of 1.85×10^7 cells per ml. In contrast, the maximum cell density in seawater without addition organic carbon was 1.01×10^6 cells per ml. This strain has been described previously to form small colonies on 1/10 R2A agar media, but did not growth in any other artificial media. These growth

characteristics showed that HTCC 2207 is a slow growing, oligotrophic, psychro-mesophilic bacterium.

Initial sequencing has so far revealed an unclosed genome of 2,619,777 base pairs coding for 2390 open reading frames. The G+C content is 49.10 mol %. The bacterium possesses all major metabolic pathways, but is requires some vitamins. Proteomic analyses identified 146 expressed proteins including a biopolymer transporter, nitrate transporter, flagellin modification proteins, urease, and a pilus assembly protein.

HTCC 2207 predominantly contained the unsaturated fatty acids 18:1 ω7c and 16:1 ω7c + 16:1 ω6c. The fatty acids 16:0, 16:1, and 18:1 were commonly found in previously cultivated genera of *Gammaproteobacteria*. This strain also contained significant amounts of 3-OH 10:0, 3-OH 12:0, 17:1 ω8c, 14:0, and 10:0 fatty acids.

From the phenotypic, genotypic, and genomic evidence, it is proposed that HTCC 2207 should be established as a new genus and species.

A Genomic and Proteomic Characterization of the First Cultured Oligotrophic Marine *Gammaproteobacterium* from the SAR92 Clade

by

Brett L. Mellbye

A THESIS

submitted to

Oregon State University

University Honors College

in partial fulfillment of

the requirements for the

degree of

Honors Baccalaureate of Science in Microbiology (Honors Associate)

Presented June 2, 2006

Commencement June 2006

Honors Baccalaureate of Science in Microbiology thesis of Brett L. Mellbye presented on

June 2, 2006

APPROVED:

_____

Mentor, representing Microbiology

_____

Committee Member, representing Chemistry

_____

Committee Member, representing Microbiology

_____

Committee Member, representing Microbiology

_____

Chair, Department of Microbiology

_____

Dean, University Honors College

I understand that my thesis will become part of the permanent collection of Oregon State University, University Honors College. My signature below authorizes release of my project to any reader upon request.

_____

Brett L. Mellbye, Author

AKNOWLEDGEMENTS

I would like to express my gratitude to the entire Giovannoni Lab group for their support and expert advice throughout this entire thesis project. Special thanks to Kevin Vergin, Sarah Sowell, Dr. Ulrich Stingl, and Jim Tripp for devoting so much of their time to helping me with this project. I would like to acknowledge my mentor, Dr. Steve Giovannoni, for providing such a great research opportunity for undergraduates at Oregon State University. I would also like to acknowledge my committee members, Dr. Douglas Barofsky, Kevin Vergin, and Dr. Ulrich Stingl for their comments on my thesis and for providing me with practice for graduate school. Special thanks to Dr. Al Soeldner and Dr. Michael Nesson of the OSU Electron Microscope Facility for electron microscope expertise, and the crew of the RV *Elakha* for assistance with sample and water collection. Finally, I would like to thank all my friends and family who have supported me and my education. This research was supported by the Gordon and Betty Moore Foundation and American Society for Microbiology Undergraduate Research Fellowship.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

A Genomic and Proteomic Characterization of the First Cultured Oligotrophic Marine *Gammaproteobacterium* from the SAR92 Clade

INTRODUCTION

The application of molecular techniques using 16S rRNA gene sequence data has revealed a vast phylogenetic diversity in a variety of marine environments (Béjà et al, 2000; Delong, 1992; Giovannoni et al, 1990a; Suzuki et al, 1997). However, this diversity is not reflected by the cultured representatives from these habitats (Rappé & Giovannoni, 2003). Epifluorescence and direct viable counting methods suggest that between 0.01 to 0.1% of all microbial cells from marine environments will form colonies on standard agar plates typically used for isolating bacteria (Amann et al, 1995; Porter & Feig, 2005; Kogure et al, 1979; Ferguson et al, 1984). Clearly, there is a need to attempt to cultivate the uncultivated marine species by new techniques and characterize them by genomic studies. The abundance of these microorganisms suggests their importance and their physiology may reveal their role in biogeochemical cycles.

Attempts to characterize organisms by DNA sequence started with the use of DNA-DNA hybridizations to compare genome sequences (Gould, 1985; Olsen, 1993; Tourova, 2000). DNA-DNA hybridization was successful in comparing similarity in base pairing between different genomic DNAs to determine evolutionary relationships. However, the technique is time consuming and not suited for rapid

identification of species (Gevers et al, 2005). It also requires the isolation and cultivation of the strains to be compared (Gevers et al., 2005).

In 1986, Carl Woese and his colleagues suggested the use of 16S rRNA sequences as a reliable evolutionary clock (Pace et al., 1986). Since these genes are highly conserved and vary slowly over time, Woese reasoned that a high degree of similarity between two bacterial RNAs meant a close kinship between them (Pace et al., 1986, Woese, 1987). Woese and his colleagues found that bacteria could be characterized phylogenetically by comparison of only 16S rRNA genes and they were able to distinguish the archaebacteria domain by using this technique (Woese, 1987). Molecular techniques allowed the use of 16S rRNA gene amplification and sequencing to study the phylogeny of cultivated and uncultivated strains in marine environments for the first time (Giovannoni et al., 1990a). Still, characterization by 16S rRNA indicates only phylogenetic closeness and not phenotypic characteristics (Gevers et al., 2005). It also may lack resolution and occasionally disagrees with DNA-DNA hybridizations (Gevers et al., 2005).

Recent approaches to characterizing marine bacteria have taken a polyphasic approach that combines aspects of biochemical, phenotypic, 16S rRNA sequence, and DNA-DNA hybridization testing (Cho & Giovannoni, 2003a, 2003b; Cho & Giovannoni, 2004a, 2004b). This approach is successful in combining a variety of techniques to provide information about new strains of bacteria, but it is still flawed due to the inability to cultivate most marine bacteria on conventional media required for biochemical and phenotypic testing.

New techniques have been developed recently to cultivate previously uncultured microorganisms. These novel approaches include high-throughput culturing (HTC) using dilution-to-extinction (Connon & Giovannoni, 2002; Rappé et al, 2002; Wang et al. 1996). The HTC technique has been a great success in cultivating new strains. The application of this technique has lead to the first cultivation of members of the SAR11 clade, but also many novel strains among the *Proteobacteria*, *Bacteroidetes*, *Planctomycetes*, *Lentisphaera*, and other orders of marine bacteria (Connon & Giovannoni, 2002; Rappé et al, 2002; Cho & Giovannoni, 2003a, 2003b; Cho et al, 2004; Cho & Giovannoni, 2004a, 2004b, 2004c). The success of this approach is thought to be due to the use of growth conditions that closely mimic those of natural environments (Connon & Giovannoni, 2002; Kaeberlein et al, 2002; Zengler et al, 2002).

Since these growth conditions mimic natural environments and are carried out in undefined media, they are still inappropriate for most biochemical and phenotypic testing. In addition, most biochemical tests are designed for identification of enteric bacteria and do not reveal useful information about free-living bacteria in their natural environment (O'Hara et al., 1992; Torsvik et al., 1996; Hall & Brazier, 1997; Hall et al., 1999). For example, marine isolates are interesting because they may be photosynthetic, fix nitrogen, or have specific metabolic cycles. Traditional characterization reveals if an isolate hydrolyzes urea, ferments glucose, produces oxidase, breaks down tryptophan, or is susceptible to different antibiotics. A new approach is needed to characterize the novel oligotrophic strains being isolated from

marine and other natural environments that can reveal interesting aspects of metabolism and biochemistry important in biogeochemical cycles.

Here we report on a new approach to strain characterization that utilizes complete genome sequencing and genomic analysis, proteomic analysis, and analysis of growth conditions. This new technique is essentially an extension of the polyphasic approach, but has a greater application to isolates from marine and other natural environments. These organisms are commonly referred to as oligotrophs.

Oligotrophs are defined as heterotrophic bacteria with the ability to grow at minimal organic carbon substrate concentration (1 to 15 mg C/L) even though they may have the ability to grow in richer media (Morita, 1997). Many oligotrophic marine bacteria appear to be restricted to growing only at low nutrient concentrations (Morita, 1997). However, this second definition is arbitrary and it is difficult to test all possible growth conditions available for the organisms (Morita, 1997). The use of high-throughput culturing is essential for this proposed method of characterization since the only requirement to perform these analyses is the ability to grow a sufficient amount of biomass for genome sequencing, proteomic analysis, and fatty acid analysis. A novel genus and species of the oligotrophic marine *Gammaproteobacteria* (OMG) group, HTC collection number 2207 (HTCC 2207) was chosen to test this new method of characterization (Cho & Giovannoni, 2004c).

This new polyphasic approach replaces techniques that are no longer useful for characterization of environmental isolates. Genome and proteomic analysis takes the place of traditional biochemical and genetic tests. Through the use of the new autoannotation techniques described, we can analyze all known metabolic pathways.

Genomics and proteomics also allow us to quickly search for flagellar genes that suggest motility, photosyntheic genes, and other genes of interest. Other interesting genetic characteristics of the genome such as mol % G+C content, DNA-DNA hybridization *in silico*, and operon structure can also be elucidated. Proteomic analysis allows us to look at genes of interest as well as the gene expression of the organism of interest when grown in amended seawater media. This technique has the potential to look at the levels of gene expression under different conditions. Keeping with tradition, we also analyze the growth rate of the organism under different conditions. Analysis of fatty acids produced by the organism of interest is also carried out. This new method will potentially more fully characterize new species and reveal more useful information about bacteria in natural environments.

MATERIALS AND METHODS

Isolation

Seawater samples were collected from the surface water of the Pacific Ocean at the southern jetty in Newport, Oregon (Cho & Giovannoni, 2004c). Seawater samples were immediately stored in the dark at ambient seawater temperatures until further processing. Several novel strains were isolated using the high throughput culturing method (Connon & Giovannoni, 2002) including HTCC2207, a member of the SAR92 clade and the oligotrophic marine *Gammaproteobacteria* (OMG) group (Cho & Giovannoni, 2004c). Strains were purified by dilution-to-extinction technique in low-nutrient heterotrophic medium (LNHM) (0.2-µm-pore-size-filtered and autoclaved seawater amended 1.0 µM NH NH_4Cl, 0.1µM KH_2PO_4, and vitamin mix at a 10^{-4} dilution stock (VT) [Davis & Guillard, 1958]) and 1X mixed carbons (MC) (LNHM plus 1X mixed carbons; 1X concentration of carbon mixtures composed of 0.001% (w/v) of D-glucose, D-ribose, succinic acid, pyruvic acid, glycerol, N-acetyl D-glucosamine, and 0.002% (v/v) of ethanol).

Microscopy

Micrographs of cells growing in stationary and exponential phase were taken to observe morphology and determine cell growth by counting the number of bacterial cells. 200 µl of samples were filtered on a 0.2 µm-pore-size polycarbonate membrane (Nucleopore; Whatman, Clifton, N.J.) in a 48-well-array. The cells in the arrays were stained with 4',6-diamidino-2-phenylindole (DAPI) (Porter & Feig,

1980) and observed using an epifluoroscence microscope (DMRB, Leica, Germany) equipped with a cooled charge-coupled device (CCD) camera (ORCA-ER, Hamamatsu, Japan) and IPLab v3.5 scientific imaging software (Scanalytics, Fairfax, VA).

40 ml of HTCC 2207 culture, grown in LNHM +0.1X MC +1X VT at 16ºC until late exponential phase, was collected by centrifugation at 20,000 RPM for 1 hour (Beckman J2-21). The cells were resuspended in 1 ml of 10% LNHM and collected by centrifugation at 20,000 RPM for 1 hour (Beckman TL-100 Ultracentrifuge). The cells were resuspended in and fixed with 1 ml of 2% gluteraldehyde in PBS (pH 7.2), collected by centrifugation at 20,000 RPM for 1 hour (Beckman TL-100 Ultracentrifuge), and resuspended in 10% LNHM. Cells were examined using a transmission electron microscope (Phillips CM12 transmission electron microscope operated at 60 kV in transmission mode). The cell size and volume was calculated from measurements made from TEM using the equation, biovolume = $(\pi/4)w^2(l\text{-}w/3)$, where w is cell width and l is cell length.

Growth Conditions

The growth characteristics of HTCC 2207 were examined under various temperatures and carbon concentrations in polycarbonate flasks with 50 ml of LNHM collected in May 2005. For determining the optimum growth temperature, the cell density was diluted to approximately 1000 cells per ml. The experiment was carried out in triplicate with three 50 ml cultures incubated at each temperature. The cultures were incubated at 4, 10, 16, 23, and 30ºC. For determining the optimum carbon

concentration, the cell density was diluted to approximately 3000 cells per ml. This experiment was carried out in triplicate with three 50 ml cultures at each carbon concentration. Cultures in exponential growth phase were inoculated in LNHM + 1X VT amended with different concentrations of MC, including 0X MC, 0.1X MC, 0.5X MC, 1X MC, 5X MC, and 10X MC. The calculated dissolved organic carbon (DOC) concentration of 1X MC was 35 mg per liter (Cho & Giovannoni, 2004c), and the average DOC concentration of the LNHM media (Oregon coast seawater media unamended with carbon) measured by a TOC analyzer was 1.0 mg per liter (Connon & Giovannoni, 2002).

Cell densities were examined at 1 day intervals until cultures reached stationary phase, except cultures grown at 4ºC, which were examined weekly.. Epifluorescence microscopy was used to measure the growth of cells in the MC concentration experiment. Cell densities in the temperature conditions growth experiment were measured using flow cytometery (Guava EasyCyte Flow Cytometer, Guava Technologies, Inc.) and Cytosoft v3.4 Data Analysis and Aquisition software (Guava Technologies, Inc., 2002). 200 µl of each sample was pipetted into a 90-well polycarbonate plate for analysis and stained with SYBR Green I nucleic acid gel stain (Invitrogen Molecular Probes) for fifteen minutes.

DNA Preparation and Phylogenetic Analysis

Genomic DNA of HTCC 2207 was extracted for phylogenetic analysis of 16S rRNA as previously described (Cho & Giovannoni, 2004c). Additional genomic

DNA was isolated for genome sequencing as previously described (Giovannoni et al, 2005).

Proteomic Analysis

HTCC 2207 was grown in 20 liters of LNHM + 1X MC + 1X VT at 16 ºC until late exponential phase. The cells were concentrated by tangential flow filtration with a Millipore Pellicon II Mini system equipped with a 30-kDa regenerated cellulose filter. Concentrated cells were pelleted by centrifugation in a Beckman J2-21 centrifuge using a JA-20 rotor at 20,000 RPM for one hour at 4 ºC. Cells were resuspended in a minimal volume of LNHM and stored at -20 ºC until analysis.

Membrane material was obtained from cell pellets by cell lysis in 20mM Tris buffer (pH 7.4), sonication, and centrifugation. The lysate was removed and the pelleted membrane material was solubilized in 0.1% dodecylmaltoside in Tris buffer. Sodium Dodecyl Sulfate Polyacrilamide Gel Electrophoresis (SDS-PAGE) was used to chromatographically separate the solubilized membrane proteins approximately 2 cm down the gel. The gel lane was then divided into 8 sections and in-gel reduction, alkylation, digestion, and extraction was performed on each section (Shevchenko et al., 1996). Specifically, gel sections were washed with water and 50/50 (v/v) acetonitrile/50 mM ammonium bicarbonate. Disulfide bonds were reduced using 1mg/ml dithiothreitol and alkylated with 10 mg/ml iodoacetamide and the gel sections were dehydrated using acetonitrile and vacuum centrifugation. Each section was then rehydrated in 12.5 ng/μl trypsin in 25 mM ammonium bicarbonate and digestion occurred at 37 ºC overnight. Peptides were extracted from the gel slices

twice using 50/50 (v/v) acetonitrile/50 mM ammonium bicarbonate and once using acetonitrile. All peptides extracted from an individual gel slice were pooled, evaporated to 5 μl in a vacuum centrifuge, redissolved in 20 μl 1% acetonitrile, 0.1% TFA, and acidified to pH < 3 with 10% TFA if necessary.

Offline high performance liquid chromatography (HPLC) followed by tandem MALDI mass spectrometry was performed as described by Stapels et al (2004).

In-line HPLC was performed on a Waters NanoAcquity Ultraperformance liquid chromatography instrument with a Symmetry C18 trap (Waters, Milford, MA) coupled with an Atlantis dC18 75 μm column (Waters). The solvent system consisted of 0.1% TFA in water or acetonitrile for solvent A and B respectively. The eluent from the UPLC was introduced directly into a quadrupole-time-of-flight (QTOF) Global Ultima tandem mass spectrometer (Micromass, Manchester, UK) with a spray voltage of 3.5 kV. MS/MS spectra were collected in data-dependent mode using a collision-induced dissociation (CID) energy between 25 and 65 eV depending on the mass of the precursor ion.

Mascot from Matrix Science (London, UK) was used to search all tandem mass spectra. Data from the MALDI mass spectrometer were analyzed with Mascot using GPS explorer software from Applied Biosystems, while Masslynx software (Waters) was used to analyze data obtained on the Q-TOF mass spectrometer.

Genomic Analysis

Genomic DNA of HTCC 2207 was sequenced using shotgun sequencing by J. Craig Venter Institute. A draft, unclosed genome consisting of six contigs was

obtained and the contigs were loaded into the GenDB annotation application program. An automated pipeline runs within GenDB to produce an annotation. The first step in the pipeline is to run Glimmer 2.0 to predict ORFs by length, ribosome binding site (RBS), TATA boxes, and check codon adaptive index. The second step is to run protein-protein Basic Local Alignment Search Tool (BLASTP) of the predicted open reading frames (ORFs) against the Kyoto Encyclopedia of Genes and Genomes (KEGG), SwissProt, and Clusters of Orthologous Groups (COG) databases of known proteins. The third step in the pipeline is to use the Hidden Markov Model (HMM) motif search using the Pfam and Interpro databases to search for protein motifs to assign putative function. The fourth step is to use a transmembrane helices HMM to search for putative membrane proteins. In the fifth and last step of the pipeline, GenDB attempts to annotate the enzyme commission (EC) number, make a gene call, and a gene description. This annotation is used to predict major metabolic pathways and biosynthesis of amino acids, vitamins, and growth factors.

G+C mol % measurements were computed using the genome sequence. DNA base composition was calculated from the six contigs using Practical Extraction and Reporting Language (PERL).

Cellular Fatty Acid Analysis

HTTC 2207 was grown in 4 liters of LNHM + 0.1X MC + 1X VT at 16 ºC until late exponential phase. The cells were concentrated by tangential flow filtration with a Millipore Pellicon II Mini system equipped with a 30-kDa regenerated cellulose filter. Concentrated cells were pelleted by centrifugation in a Beckman J2-21

centrifuge using a JA-20 rotor at 20,000 RPM for one hour at 4 ºC. Cells were resuspended in a minimal volume of LNHM and stored at -80 ºC until they are shipped for analysis by MIDI (Microbial ID, Inc., Newark, DE).

RESULTS AND DISCUSSION

Morphology

Cell morphology and size of HTCC 2207 was determined using an epifluorescence microscope (after staining with DAPI) and a transmission electron microscope (TEM). Figure 1 shows images of cells stained with DAPI and Figure 2 shows micrographs taken by TEM. HTCC 2207 appear to be short rods that divide by binary fission. The approximate cell size was 1.08 μm by 0.45 μm. The cell volume was calculated to be 0.148 μm^3.

Figure 1. Epifluorescence microscopy image of HTCC 2207 cells stained with DAPI.

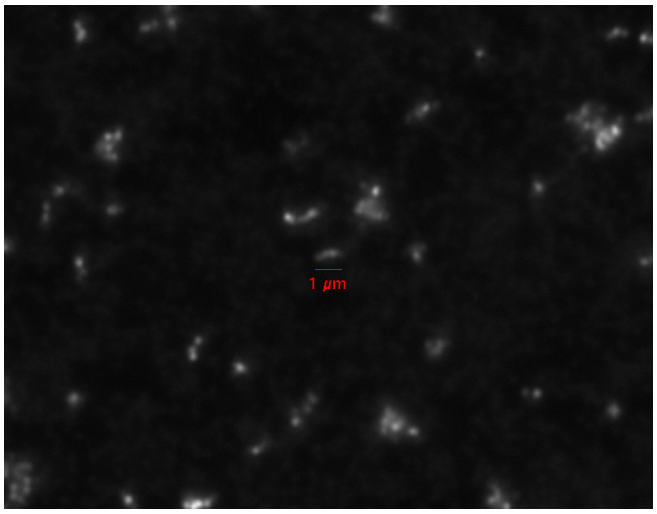
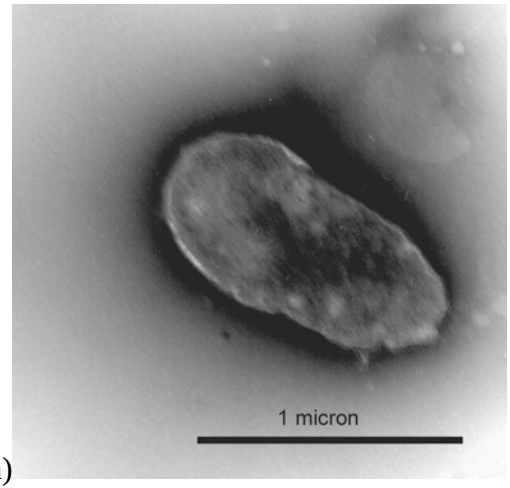
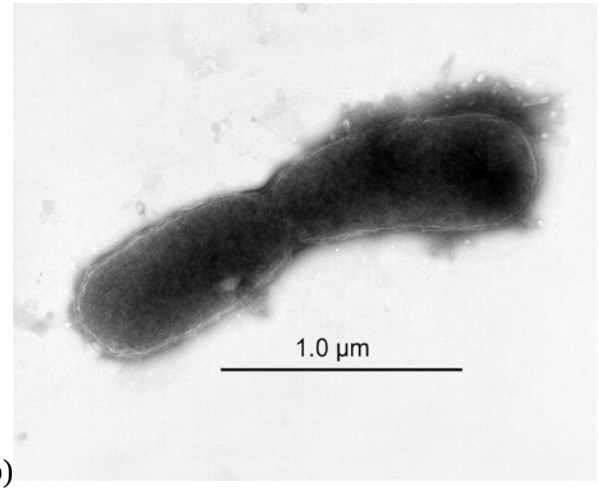


Figure 2. Transmission electron micrographs of negatively stained cells of HTCC 2207. (a) Single cell, (b) Dividing cells.

a)



b)



Growth Characteristics

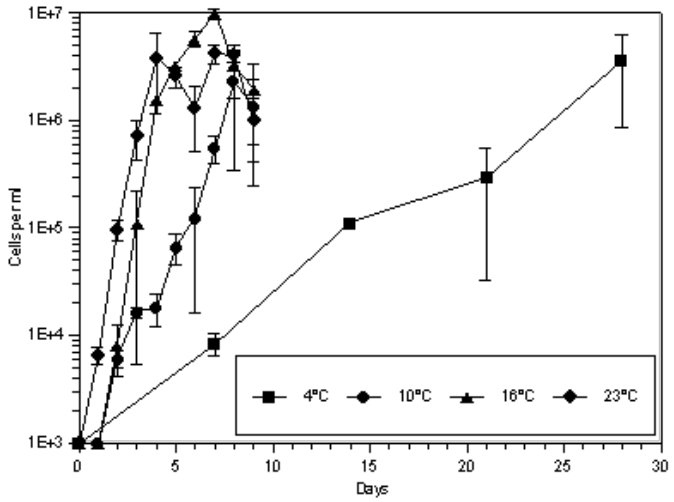
The growth characteristics of HTCC 2207 were investigated at different temperatures and mixed carbon concentrations (Figure 3). The specific growth rate at each temperature and carbon concentration was calculated from the bacterium's growth during exponential phase and can be visualized in Figure 4 and Table 1. Specific growth rate is compared to the maximum cell density reached in Table 1.

The strain was psychro-mesophilic since it showed growth from 4 ºC-23 ºC. No growth was observed at 30 ºC. Cultures incubated at 16 ºC reached the highest density, but grew at a similar rate compared to cultures at 23 ºC.

HTCC 2207 is considered an oligotrophic strain because it grows in low nutrient heterotrophic media (LNHM) that contains 1 mg C/L (Morita, 1997). However, the cultures reached the highest cell densities when grown in LNHM and 1X vitamin mix with 0.1X to 5X mixed carbon (MC). The bacterium showed the highest specific growth rate when incubated in LNHM + 1X vitamin mix + 0.1X MC. There was a distinct difference in growth rate and maximum cell density in cultures with added carbon versus cultures with no added carbon. However, growth rate and maximum cell density began to decrease when more than 5X MC was added.

Figure 3. Growth characteristics of HTCC 2207 at different temperatures (a) and dissolved organic carbon concentrations (b). (LNHM medium contains approximately 1.0 mg/L of dissolved organic carbon and 1X MC contains 3.5 mg/L of dissolved organic carbon.)

a)



b)

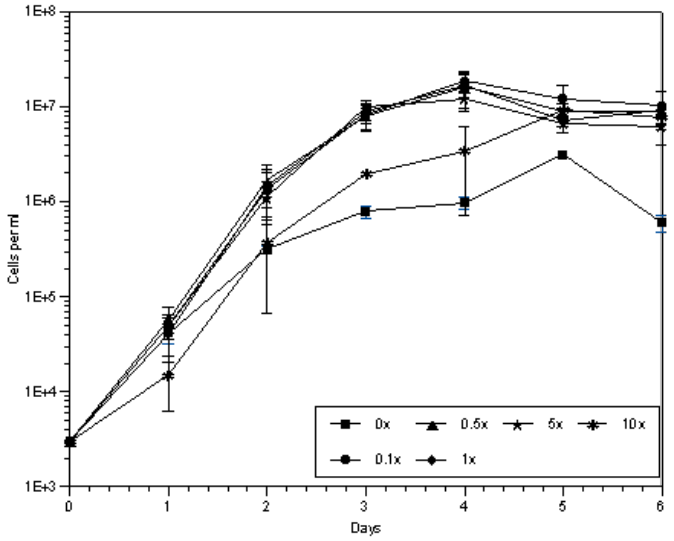
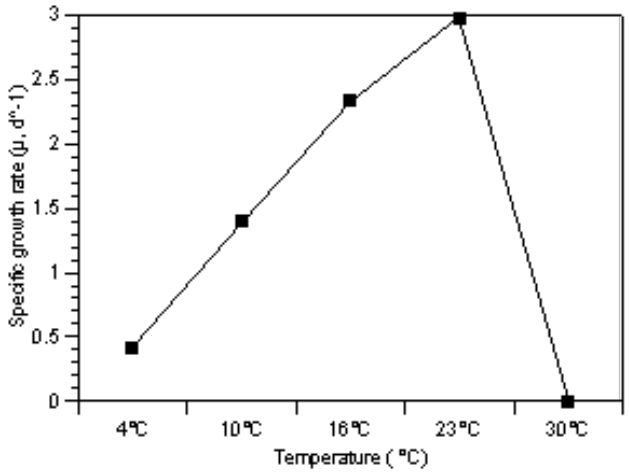


Figure 4. Specific growth rates (μ) of HTCC 2207 at different temperatures (a) and dissolved organic carbon concentrations (b).

a)



b)

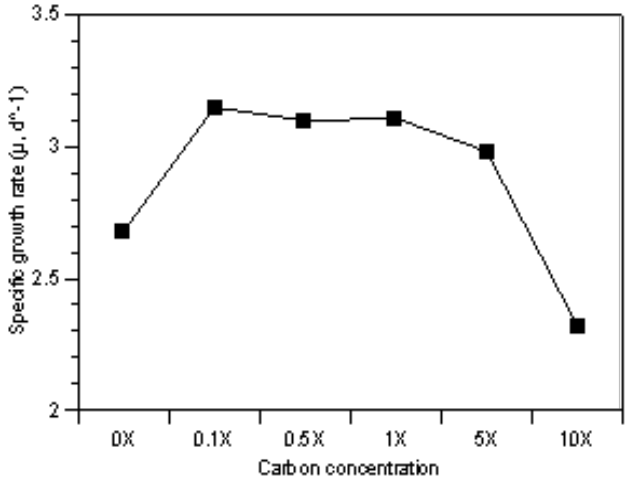


Table 1. Comparisons of specific growth rates and maximum cell densities of HTCC 2207 under different growth conditions. (All temperature treatments carried out in LNHM + 1X MC + 1X VT. All carbon concentration treatments carried out at 16°C in LNHM + 1X VT.)

Treatment	Specific growth rate (day^{-1})	Maximum cell density (X 10^6 cells/ml)
4°C	0.42	3.60
10°C	1.40	2.31
16°C	2.34	9.94
23°C	2.98	4.22
30°C	0	0
0X MC	2.68	3.14
0.1X MC	3.15	18.5
0.5X MC	3.10	16.0
1X MC	3.11	16.6
5X MC	2.98	11.9
10X MC	2.32	9.38

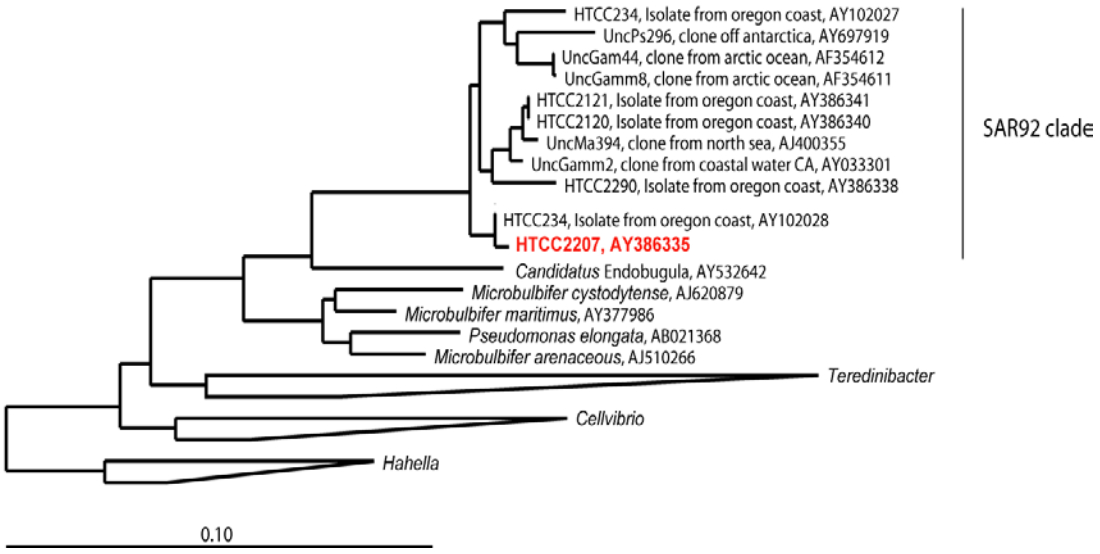
Medium tests

The ability of HTCC 2207 to grow on six different agars has been previously described (Cho and Giovannoni, 2004c). HTCC 2207 was not able to grow in any liquid artificial seawater tested, but it does grow slowly on 1/10 R2A agar plates (Cho and Giovannoni, 2004c). The inability of the bacterium to grow in liquid ASW makes it difficult to impossible to investigate ranges and optima of pH and salinity, or utilization of sole carbon sources.

Phylogenetic Analysis based on 16S rRNA sequence

The phylogenetic tree based on the 16S rRNA gene sequences of HTCC 2207, environmental clones, and characterized *Gammaproteobacteria* shows the position of HTCC 2207 within the SAR92 clade (Figure 5). The 16S rRNA gene sequence of HTCC 2207 was 95.9-99.8% similar to environmental clones in the SAR92 clade. This clade was previously shown to have sequence similarities between members to form a distinct lineage that may share a common evolutionary origin with *Teredinibacter* and *Microbulbifer* (Cho and Giovannoni, 2004c). The most closely related clones outside of the SAR92 clade were uncultured *Pseudomonas* that showed 95.1% similarity and *Candidatus* Endobugula that showed 89.7% similarity. The previously cultivated genera *Microbulbifer* species showed 90.1 – 91.0% 16S rRNA gene sequence similarity to HTCC 2207. As a result, the SAR92 clade was placed distantly from many previously cultivated *Gammaproteobacteria* such as *Microbulbifer*, *Oceanospirillum*, *Pseudomonas*, *Vibrio*, and *Marinobacter*.

Figure 5. Phylogenetic tree based on 16S rRNA gene sequences showing the relationship of HTCC 2207 with environmental clones and previously cultured genera in the *Gammaproteobacteria*



Genomic, Proteomic, and Fatty Acid Analysis

The draft, unclosed genome of HTCC 2207 was obtained as six contigs that were loaded into the GenDB autoannotation application program for analysis. The genome size was 2,619,777 base pairs with 2390 predicted open reading frames (ORFs). DNA base composition of the genome was calculated from the genome sequence using Practical Extraction and Reporting Language (PERL). The DNA G+C content of HTCC 2207 was 49.10%. Unclosed genome structure is listed in Table 2. The automated pipeline that runs within GenDB compares predicted ORFs from the genome with KEGG to generate metabolic wiring diagrams that predict the metabolic capabilties of the bacterium (Table 3). HTCC 2207 possesses all major metabolic pathways including glycolysis, tricarboxylic acid cycle, glyoxylate shunt, pentose phosphate, and amino acid biosynthesis. However, the bacterium is deficient in the biosynthesis of some vitamins, including thiamine and vitamin B12. Analysis of the genome to reconstruct metabolic pathways has the potential to replace traditional tests that identify enzymatic activity, sole carbon utilization, and oxygen requirements. If genome sequencing becomes a characterization tool, DNA-DNA hybridizations may be carried out *in silico* to analyze genomic DNA relatedness.

Proteomic analysis of HTCC 2207 identified 146 expressed proteins (Table 3). Proteins of interest included a biopolymer transporter, nitrate transporter, flagellin modification proteins, urease, and a pilus assembly protein. Proteomic analysis may help replace traditional tests if coverage of the proteome can be increased. We utilized this test to identify interesting proteins being expressed in late exponential phase.

The fatty acid profile of HTCC 2207 was analyzed by gas chromatography (Table 5). Unsaturated fatty acids were dominant in this bacterium with 18:1 ω7c accounting for 12.07% and 16:1 ω7c + 16:1 ω6c accounting for 38.01%. The total percentage of unsaturated fatty acids was 64.08%. Commonly detected major fatty acids in many previously cultivated *Gammaproteobacteria* have been 16:0, 16:1, and 18:1 (Oliver and Colwell, 1973; Lambert et al., 1983; Bertone et al., 1996). HTCC 2207 shows these characteristics of the *Gammaproteobacteria* with 6.69% for 16:0 fatty acids, 43.97% for 16:1 fatty acids, and 16.25% 18:1 fatty acids. This strain can be differentiated by the presence of 5.71% 3-OH 10:0, 6.00% 3-OH 12:0, 3.33% 17:1 ω8c, 4.63% 14:0, and 4.51% 10:0. Fatty acid composition analysis has been previously shown to give results in agreement with DNA-DNA hybridization data when determining phylogenetic relatedness between strains (Bertone et al., 1996).

Table 2. Unclosed genome structure of HTCC 2207

Contigs	6
Genome Size	Approx. 2,619,777 base pairs
Open Reading Frames	Approx. 2390 ORFs
Mol G+C %	49.10

Table 3. Metabolic pathways of HTCC 2207 predicted by genomic comparison of ORFs to KEGG

Pathway	Prediction
Glycolysis	+
TCA cycle	+
Glyoxylate shunt	+
Respiration	+
Pentose phosphate cycle	+
Fatty Acid Biosynthesis	+
Cell Wall Biosynthesis	+
Amino Acid Biosynthesis (20)	+
Heme Biosynthesis	+
Ubiquinone	+
Nicotinate and nicotinamide	+
Folate	+
Riboflavin	+
Pantothenate	+
B6	+
Thiamine	-
Biotin	?
B12	-

Table 4. Expressed proteins identified in proteomic analysis.

ORF Name	Predicted Function
C7_0328	short chain dehydrogenase
C3_0039	uncharacterized conserved motif
C3_0081	TolR, protein transport
C3_0105	sulfate adenylyltransferase
C3_0107	short-chain dehydrogenase/reductase SDR
C3_0121	enoyl-CoA hydratase/isomerase family protein
C3_0131	no hits
C3_0138	ATP-dependent hsl protease
C3_0142	probable enoyl-CoA hydratase/isomerase
C3_0155	no hits
C3_0185	no hits
C3_0209	acetyl-CoA acetyltransferase
C3_0222	uridylyltransferase
C3_0227	ribosome recycling factor
C3_0233	no hits
C3_0241	acetyl-CoA carboxylase carboxyl transferase
C3_0253	TolQ, Biopolymer transport proteins
C3_0270	no hits
C3_0295	acyl carrier protein
C3_0306	Peroxiredoxin
C3_0319	electron transfer flavoprotein-ubiquinone oxidoreductase
C3_0345	thioredoxin reductase
C3_0361	malate synthase G
C3_0409	uncharacterized paraquat-inducible protein B
C3_0418	no hits
C3_0430	no hits
C3_0462	no hits
C3_0468	aspartate-semialdehyde dehydrogenase
C3_0469	isocitrate/isopropylmalate dehydrogenase
C3_0484	hydrolase, alpha/beta fold family
C3_0498	succinate dehydrogenase (C subunit)
C3_0502	hypothetical protein PP2378
C4_0003	ClpP protease
C4_0069	glycine dehydrogenase
C4_0096	Mfd, Transcription-repair coupling factor
C4_0161	molybdenum cofactor biosynthesis protein B
C4_0208	AhpF, Alkyl hydroperoxide reductase
C4_0220	SecD, secretion
C4_0236	no hits
C4_0239	TonB dependent receptor
C4_0253	ExbB, biopolymer transport
C4_0262	related to tryptophan synthase
C4_0264	ribonucleoside-diphosphate reductase
C4_0265	hypothetical protein
C4_0268	sensory box histidine kinase
C4_0276	dethiobiotin synthetase

C5_0039	hypothetical protein
C5_0045	MucC, positive transcriptional regulator
C5_0063	aspartyl aminopeptidase
C5_0080	uroporphyrinogen decarboxylase
C5_0186	leuA; 2-isopropylmalate synthase
C5_0191	no hits
C5_0234	putative ATP-dependent RNA helicase RhlE
C5_0266	clpB, chaperone
C5_0268	no hits
C5_0311	GlmS, glucosamine 6-phosphate synthetase
C5_0322	no hits
C5_0329	no hits
C5_0342	MviM, Predicted dehydrogenases
C5_0353	CysC, Adenylylsulfate kinase
C6_0008	50S ribosomal subunit protein
C6_0017	50S ribosomal protein L30
C6_0032	rotamase
C6_0037	no hits
C6_0050	gspE, general secretion
C6_0059	no hits
C6_0099	no hits
C6_0122	glucose/galactose transporter family protein
C6_0145	ATP-dependent RNA helicase
C6_0164	DnaK suppressor protein
C6_0167	htpX, probable protease htpX
C6_0177	EFP, Elongation factor P
C6_0190	Thiolase
C6_0203	ArgG, Argininosuccinate synthase
C6_0204	small-conductance mechanosensitive channel
C6_0217	RnfG, Predicted NADH:ubiquinone oxidoreductase
C6_0231	deoxycytidine triphosphate deaminase
C6_0270	hypothetical protein
C7_0015	secretion protein SecE
C7_0031	N-acetyl-gamma-glutamyl-phosphate reductase
C7_0034	delta-aminolevulinic acid dehydratase
C7_0043	5-formyltetrahydrofolate cyclo-ligase
C7_0044	pyruvate kinase
C7_0046	ribose 5-phosphate isomerase
C7_0047	RbsK, Sugar kinases, ribokinase family
C7_0059	AICAR transformylase/IMP cyclohydrolase PurH
C7_0064	ubiquinone biosynthesis protein
C7_0071	trpE; anthranilate synthase component I
C7_0073	no hits
C7_0085	PLP dependent enzymes class III
C7_0087	homoserine acetyltransferase
C7_0094	thioredoxin
C7_0121	FtsE, Predicted ATPase involved in cell division
C7_0124	phosphopantetheine adenylyltransferase
C7_0125	mutM; formamidopyrimidine-DNA glycosylase
C7_0197	Ggycyl-tRNA synthetase alpha chain

C7_0205	oxaA, inner membrane protein
C7_0210	parA; ATPase involved in chromosome partitioning
C7_0233	peptide deformylase
C7_0240	aroE; shikimate 5-dehydrogenase
C7_0258	predicted GTPase
C7_0270	probable transcriptional regulator
C7_0273	pgk; 3-phosphoglycerate kinase
C7_0286	putative nitrate transporter
C7_0302	ribonuclease PH
C7_0303	PyrE, Orotate phosphoribosyltransferase
C7_0323	phosphoenolpyruvate-protein phosphotransferase PtsP
C7_0343	3-deoxy-D-manno-octulosonic-acid (KDO) transferase
C7_0373	two-component sensor NtrB
C7_0387	FKBP-type peptidyl-prolyl cis-trans isomerase
C7_0396	phosphate ABC transporter, permease protein
C7_0397	Phosphate import ATP-binding protein pstB
C7_0398	Phosphate transport system protein phoU
C7_0448	tryptophan halogenase, putative
C7_0452	putative RNA polymerase sigma factor
C7_0469	putative permease protein
C7_0478	ureC; urease, alpha subunit
C7_0480	urease gamma subunit
C7_0488	no hits
C7_0494	no hits
C7_0575	hypothetical UPF0042 protein PA4465
C7_0585	toluene tolerance
C7_0597	secG, auxillary membrane component
C7_0604	ribosome-binding factor A
C7_0609	grpE, HSP-70 cofactor
C7_0611	dnaJ, chaperone
C7_0613	carbamoyl-phosphate synthase small chain
C7_0618	acetylornithine and succinylornithine aminotransferase
C7_0622	monothiol glutaredoxin
C7_0624	hypthetical protein
C7_0637	protein of unknown function, DUF484 superfamily
C7_0648	HemY, vitamin & cofactor metabolism
C7_0654	ACT domain protein
C7_0670	efflux transporter
C7_0698	probable carbonic anhydrase
C7_0758	TonB-dependent receptor
C7_0773	no hits
C7_0792	periplasmic binding transglycosylase
C7_0795	MraZ, probable transcription factor
C7_0804	murG, cell wall formation
C7_0812	SecA, secretion
C7_0819	pilB, Type 4 fimbrial assembly/pilus assembly
C7_0830	hypotheical protein
C7_0832	fumarate hydratase
C7_0860	conserved hypothetical protein
C7_0862	threonine synthase

Table 5. Cellular Fatty Acid Composition (%) of HTCC 2207

Fatty Acids	Composition (%)
Straight-chain:	
9:0	0.91
10:0	4.51
11:0	0.18
12:0	1.21
13:0	0.15
14:0	4.63
15:0	0.90
16:0	6.38
17:0	0.12
18:0	0.44
Branched:	
iso14:1	0.18
iso15:1	0.13
iso15:0	0.14
anteiso15:0	0.53
iso16:0	0.31
anteiso17:0	0.12
Unsaturated:	
15:1 ω8c	0.39
16:1 ω9c	5.96
17:1 ω8c	3.33
17:1 ω6c	0.14
18:1 ω9c	4.18
18:1 ω7c	12.07
Hydroxy:	
3-OH 10:0	5.71
2-OH 11:0	0.16
3-OH 11:0	1.29
3-OH iso12:0	0.14
3-OH 12:0	6.00
3-OH 12:1	0.46
Other acids	
unknown 11.825	0.24
3-OH 13:0 + iso15:1	0.27
N alcohol 16:0	0.12
16:1 ω7c + 16:1 ω6c	38.01
10 methyl + 16:0	0.12
iso17:1 + anteiso17:1	0.46
unknown 18.821	0.09

Conclusions

Phylogenetic analysis based on 16S rRNA sequences identified HTCC 2207 as a candidate novel genus and species. HTCC 2207 exhibited a low 16S rRNA gene sequence similarity of 90.1-91.0% with *Microbulbifer*, the most closely related previously cultivated genus in the *Gammaproteobacteria*. Experimental results from analysis of growth conditions, genome structure, expressed proteins, metabolism, and fatty acid profile serves to support this evidence by emphasizing significant differences in phenotypic and genotypic traits between these two genera (Yoon et al, 2004).

The analysis of growth conditions of HTCC 2207 identifies this strain as a unique oligotrophic model. This bacterium shows the ability to grow in LNHM, which contains 1 mg C/L of dissolved organic carbon. However, it responds to additional carbon with a significant increase in growth rate and maximum cell density. This is in contrast to '*Candidatus* Pelagibacter ubique' of the SAR11 clade of *Alphaproteobacteria* and many other oligotrophic bacteria that are sensitive to increases in dissolved organic carbon concentration. The ability of this marine oligotroph to reach a higher maximum cell density in liquid culture over a few days offers many future opportunities for analyses that require biomass.

Description of HTCC 2207

Cells are Gram-negative, flagellated, short rods that divide by binary fission, and are approximately 1.08 µm by 0.45 µm. Growth occurs at 4ºC - 23 ºC and is considered psychro-to-mesophilic. Growth in LNHM amended in 0X -10X MC (1 mg to 350 mg of dissolved organic carbon/L). Small colony formation on 1/10 R2A agar.

Chemoheterotrophic metabolism capable of carrying out glycolysis, tricarboxylic acid cycle, glyoxylate shunt, respiration, pentose phosphate cycle. Biosynthetic metabolic pathways for biosynthesis of fatty acids, cell wall, amino acids, heme, ubiquinone, nicotinate and nicotinamide, folate, riboflavin, pantothenate, and vitamin B6 are present. Biosynthesis of thiamine, biotin, and B12 may be absent. Predominant fatty acids are 18:1 ω7c (12.07%) and 16:1 ω7c + 16:1 ω6c (38.01%). Other minor fatty acids are16:0 (6.69%), 3-OH 10:0 (5.71%), 3-OH 12:0 (6.00%), 17:1 ω8c (3.33%), 14:0 (4.63%), and 10:0 (4.51%). No 16S sequences of bacterial species with validly published names show more then 91.0% similarity. HTCC 2207 was isolated from the surface at the southern jetty in Newport, Oregon, a depth of 10 m at a station (NH5) 27.6 km off of the coast of Oregon (44º 4.82´ N, 124º 24.7´ W).

BIBLIOGRAPHY

Amann, R.I., W. Ludwig, and K.H.Schleifer. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev* **59**: 143 169.

Béjà, O., M.T. Suzuki, E.V. Koonin, L. Aravind, A. Hadd, L.P. Nyugen, R. Villacorta, M. Amjadi, C. Garrigues, S.B. Jovanovich, R.A. Feldman, and E.F. Delong. 2000. Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage. *Environ Microbiol* **68**: 505-518.

Bertone, S., Giacomini, M., Ruggiero, C., Piccarolo, C., and Calegari, L. 1996. Automated systems for identification of heterotrophic marine bacteria on the basis of their fatty acid composition. *Appl Environ Microbiol* **62**: 2122-2132.

Cho, J.C. and Giovannoni S.J. 2003a. *Parvularcula bermudensis* gen. nov., sp. nov., a marine bacterium that forms a deep branch in the alpha-proteobacteria. *Int J Syst Bacteriol* **53**: 1031-1036.

Cho, J.C. and Giovannoni, S.J. 2003b. *Fulvimarina pelagi* gen. Nov., sp. Nov., a marine bacterium that forms a deep evolutionary lineage of descent in the order '*Rhizobiales*'. *Int J Syst And Evol Microbiol* **53**: 1853-1859.

Cho, J.C. and Giovannoni, S.J. 2004a. *Oceanicola granulosus* gen. nov., sp. nov. and *Oceanicola batsensis* sp. nov., poly-ß-hydroxybutyrate-producing marine bacteria in the order '*Rhodobacterales*'. *Int J Syst Bacteriol* **54**: 1129-1136.

Cho, J.C. and Giovannoni, S.J. 2004b. *Robiginitalea biformata* gen. nov., sp. nov., a novel marine bacterium in the family *Flavobacteriaceae* with a higher G+C content. *Int J Syst Bacteriol* **54**: 1101-1106.

Cho, J.C. and Giovannoni, S.J. 2004c. Cultivation and growth characteristics of a diverse group of oligotrophic marine *Gammaproteobacteria*. *Appl. Environ Microbiol* **70** (1): 432-440.

Cho, J.C., K.L. Vergin, R.M. Morris, and S.J. Giovannoni. 2004. *Lentispaera araneosa* gen. Nov., sp. Nov., a transparent exopolymer producing marine bacterium, and the description of a novel bacterial phylum, *Lentisphaerae*. *Environ Microbiol* **6**: 611 621.

Connon, S.A. and Giovannoni, S.J. 2002. High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. *Appl Environ Microbiol* **68**: 3878-3885.

Davis, H.C. and R.R.L. Guillard. 1958. Relative value of ten genera of micro-organisms as food for oyster and clam larvae. *USFWS Fish Bull* **58**: 293-304.

Delong, E.F. 1992. Archaea in coastal marine environments. *Proc Natl Acad Sci USA* **89**: 5685-5689.

Ferguson, R.L., E.N. Buckley, and A.V. Palumbo. 1984. Response of marine bacterioplankton to differential filtration and confinement. *Appl Environ Microbiol* **47**: 49-55.

Gevers, D., Cohan, F.M., Lawrence, J.G., Spratt, B.G., Coenye, T., Feil, E.J., Stackebrandt, E., Van de Peer, Y., Vandamme, P, Thompson, F.L., and Swings, J. 2005. Re-evaluating prokaryotic species. *Nature Reviews Microbiol.* **3**: 733-739.

Giovannoni, S.J., Britshgi, T.B., Moyer, C.L., and Field, K.G. (1990a). Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**: 60-63.

Giovannoni, S.J.,Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., Bibbs, L., Eads, J., Richardson, T.H., Noordewier, M., Rappe, M.S., Short, J., Carrington, J.C., and Mathur, E.J. 2005. Genome Streamlining in a Cosmopolitan Oceanic Bacterium. *Science* **309**:1242-1245.

Gould, S.J. 1985. A Clock of Evolution. *Natural History* **94**: 12-20.

Hall, V., and J. S. Brazier**.** 1997. Identification of actinomyceswhat are the major problems?, p. 187-192. *In* A. R. Eley, and K. W. Bennett (ed.), Anaerobic pathogens. Sheffield Academic Press, Sheffield, United Kingdom.

Hall, V., O'Neill, G.L., Magee, J.T., and Duerden, B.I. 1999. Development of Amplified Ribosomal DNA Restriction Analysis for Identification of *Actinomyces* Species and Comparison with Pyrolysis-Mass Spectrometry and Conventional Biochemical Tests. *J Clinical Microbiol.* **37**: 2255-2261.

Kaeberlein, T., K. Lewis, and S.S. Epstein. 2002. Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* **296**: 1127-1129.

Kogure, Y., Yabuuchi, E., Naka, T., Fujiwara, N., and Taga, N. (1979). A tentative direct microscopic method for counting living marine bacteria. *Can J Microbiol* **25**: 415 420.

Lambert, M.A., Hickman-Brenner, F.W., Farmer III, J.J., and Moss, C.W. 1983. Differentiation of *Vibrionaceae* Species by their cellular fatty acid composition. *Int J Syst Bacteriol* **33**: 777-792.

Morita, R.Y. 1997. Bacteria in Oligotrophic Environments: Starvation-Survival Lifestyle. Chapman and Hall, New York, NY.

O'Hara, C.M., Rhoden, D.L., and Miller, J.M. 1992. Reevaluation of the API 20E identification system versus conventional biochemicals for indentification of members of the family Enterobacteriaceae: a new look at an old product. *J Clin Microbiol* **30**: 123-125.

Oliver, J.D., and Colwell, R.R. 1973. Extractable lipids of Gram-negative marine bacteria: fatty acid compostition. *Int J Syst Bacteriol* **23**: 442-458.

Olsen, I. 1993. Recent approaches to the chemotaxonomy of the *Actinobacillus Haemophilus-Pasteurella* group (family *Pasteurellacease*). *Oral Microbiology & Immunology* **8**: 327-336.

Pace, N.R., Olsen, G.J., and Woese, C.R. 1986. Ribosomal RNA Phylogeny and the primary lines of evolutionary descent. *Cell* **45**: 325-326.

Porter, K.G. and Feig, Y.S. (1980). The use of DAPI for identifying and counting aquatic microflora. *Limnol Oceanogr* **25**: 943-948.

Rappé, M.S., Connon, S.A., Vergin, K.L., and Giovannoni, S.J. (2002). Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* **418**: 630-633.

Rappé, M.S. and Giovannoni, S.J. (2003). The uncultured microbial majority. *Ann Rev Microbiol* **57**: 369-394.

Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. 1996. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. *Anal Chem.* **68**: 850-858.

Stapels, M.D., Cho, J.-C., Giovannoni, S. J., and Barofsky, D.F. 2004. Proteomic analysis of novel marine bacteria using MALDI and ESI mass spectrometry. *J of Biomolec Tech*. **15**:191-8.

Suzuki, M.T., Rappé, M.S., Haimeberger, Z.W., Winfield, H., Adair, N., Ströbel, J., and Giovannoni S.J. (1997). Bacterial diversity among SSU rRNA gene clones and cellular isolates from the same seawater sample. *Appl Environ Microbiol* **63**: 983 989.

Torsvik, V., Sørheim, R., and Goksøyr, J. 1996. Total Bacterial Diversity in Soil and Sediment Communities- A Review. *J of Industrial Microbiology and Biotechnology* **17**: 170-178.

Tourova, T.P. 2000. The Role of DNA-DNA Hybridization and 16S rRNA Gene Sequencing in Solving Taxonomic Problems by the Haloanaerobiales. *Microbiol.* **69**: 623-634.

Wang, Y., Lau, P.C.K., and Button, D.K. 1996. A marine oligobacterium harboring genes known to be part of aromatic hydrocarbon degradation pathways of soil Pseudomonoads. *Appl. Environ. Microbiol.* **62**: 2169-2173.

Woese, Carl R. 1987. Bacterial Evolution. *Microbiol Reviews* **51**: 221-271.

Yoon, J.H., Kim, I.G., Oh, T.K., and Park, Y.H. 2004. *Microbulbifer maritimus* sp. Nov., isolated from intertidal sediment from the Yellow Sea, Korea. *Int J Syst Evol Microbiol* **54**: 1111-1116.

Zengler, K., G. Toledo, M.S. Rappé, J. Elkins, E.J. Mathur, J.M. Short, and M. Keller. 2002. Cultivating the uncultured. *Proc Natl Acad Sci USA* **99**: 15681-15686.

APPENDIX A

The preliminary name proposed for strain HTCC 2207 is *Litoribacter rhodopsinicus* gen. nov., sp. nov., as recommended by Dr. Jean P. Euzeby, an expert in etymology at Ecole Nationale Veterinaire in Toulouse, France.

Litoribacter(L. n. litus -oris, the coast; N.L. masc. n. bacter, a rod; N.L. masc. n. Litoribacter, a rod isolated near the coast of Oregon).

Litoribacter rhodopsinicus (rho.dop.si'ni.cus. N.L. masc. adj. rhodopsinicus, pertaining to rhodopsin, because the organism has a light harvesting proton pump called rhodopsin).