Annual Report for Period: 05/2007 - 04/2008
Submitted on: 02/01/2008
Principal Investigator: Stegeman, John J.
Award ID: 0430724
Organization: Woods Hole Ocean Inst
Title:
The Woods Hole Center for Oceans and Human Health

Project Participants

Senior Personnel

Name: Stegeman, John
Worked for more than 160 Hours: Yes
Contribution to Project:
Dr. Stegeman is responsible for the oversight of all of the activities of the Center, and for facilitating communication within and among the Centers.

Name: McGillicuddy, Dennis
Worked for more than 160 Hours: Yes
Contribution to Project:
Dr. McGillicuddy's responsibilities include overall project oversight for Project 2. Design of population dynamics models for the various A. fundyense genotypes, and execution of the physical oceanographic component of large-scale survey operations.

Name: Hahn, Mark
Worked for more than 160 Hours: No
Contribution to Project:
Senior Scientist and Director of the Pilot Project Program.

Name: Anderson, Donald
Worked for more than 160 Hours: Yes
Contribution to Project:
Senior Scientist and PI, is responsible for project oversight, and management.

Name: Erdner, Deana
Worked for more than 160 Hours: Yes
Contribution to Project:
Research Associate and Co-PI, is responsible for technical oversight of experimental work and sampling efforts, project management, and
data review and interpretation.

Name: Gast, Rebecca
Worked for more than 160 Hours: Yes
Contribution to Project:
Co-PI, supervised the pathogen detection,cultured amoebas from sediment samples.

Name: Amaral Zettler, Linda
Worked for more than 160 Hours: Yes
Contribution to Project:
Co-PI, supervised the microbial community surveys, participated in sample collection and processing.

Name: Polz, Martin
Worked for more than 160 Hours: No
Contribution to Project:
PI supervises and coordinates the project and participates in data analysis as well as dissemination of results.

Name: Lerczak, James
Worked for more than 160 Hours: No

Contribution to Project:
Co-PI is overseeing the collection and analysis of the physical measurements within Plum Island Estuary and is responsible for providing the physical estuarine context for the interpretation of Vibrio population dynamics.

Name: Sogin, Mitchell

Worked for more than 160 Hours: Yes

Contribution to Project:
Dr. Sogin directs activities in the Genomics Core including phylogenetic and associated bioinformatics activities. Directed the development of a new massively high-throughput tag sequencing approach for characterizing microbial populations in marine environments.

Name: Morrison, Hilary

Worked for more than 160 Hours: Yes

Contribution to Project:
Co-director of the Core Facility, Morrison has directed the sequencing service activities of the COHH genome core, participated in the tag sequencing project, and developed genomic sequencing protocols for the GS-FLX instrument. Dr. Morrison is also working with Dr. Sogin on a method of sampling two discontinuous variable regions from microbial rRNA genes; an expansion of the tag sequencing approach.

Name: Mark Welch, David

Worked for more than 160 Hours: Yes

Contribution to Project:
(Ellison GID program-no charge to COHH) has continued to participate in development of the bioinformatics analysis of the high-throughput tag sequencing project

Name: Huber, Julie

Worked for more than 160 Hours: Yes

Contribution to Project:
(Microbial Evolution program-no charge to COHH) has provided samples from Axial Seamount and has participated in the bioinformatics analysis of the high-throughput tag sequencing project. Dr. Huber is also working with Dr. Sogin on a pilot project that seeks to use comprehensive microbial population surveys afforded by tag sequencing protocols to examine anthropogenic impacts on a saltwater marsh that experience seasonal elevated levels of coliforms.

Post-doc

Name: Kirkup, Benjamin

Worked for more than 160 Hours: No

Contribution to Project:
Responsible for strain characterization, metabolic and physiological diversity, and pathogenicity determinant gene distribution.

Name: Hackett, Jeremiah

Worked for more than 160 Hours: No

Contribution to Project:
Worked on Alexandrium genomics, with a focus on the identification of genes involved in saxitoxin biosynthesis. Hackett was supported by a WHOI Postdoctoral Scholar award, but supplies and other support were provided by the COHH grant.

Name: Ralston, David

Worked for more than 160 Hours: No

Contribution to Project:
Investigation of the influence of vertical migration on the vertical distribution of A. fundyense.

Name: Carr, Jennifer

Worked for more than 160 Hours: No

Contribution to Project:
Responsible for sampling and development of genetic system for model organisms.

Name: Wildschutte, Hans
Worked for more than 160 Hours: No
Contribution to Project:
Responsible for genomics and O-antigen characterization of the vibrios.

Name: Aretxabaleta, Alfredo
Worked for more than 160 Hours: Yes
Contribution to Project:
Investigation of the mechanisms controlling the Bay of Fundy gyre system, the retentive characteristics of which are key to regional population dynamics of A. fundyense.

Graduate Student

Name: Brosnahan, Michael
Worked for more than 160 Hours: No
Contribution to Project:
Graduate Student, responsible for characterization and application of microsatellite markers used for population studies.

Name: Koid, Amy
Worked for more than 160 Hours: No
Contribution to Project:
REU student

Name: Osborn, Deborah
Worked for more than 160 Hours: No
Contribution to Project:
Hollings Fellow

Name: Willard, Eric
Worked for more than 160 Hours: No
Contribution to Project:
Northeastern University co-op student

Name: Blossom, Hannah
Worked for more than 160 Hours: No
Contribution to Project:
Northeastern University co-op student

Name: Tully, Benjamin
Worked for more than 160 Hours: No
Contribution to Project:
REU student, summer 2006.

Name: Hunt, Dana
Worked for more than 160 Hours: Yes
Contribution to Project:
Examines environmental distribution of vibrios in the water column.

Name: Preheim, Sarah
Worked for more than 160 Hours: Yes
Contribution to Project:
Researching the environmental distribution of vibrios among animal hosts.

Name: Xue, Xong
Worked for more than 160 Hours: No

Contribution to Project:
Identification of isolates, growth dynamics of vibrios over tidal cycles.

Name: Hasegawa, Yuko
Worked for more than 160 Hours: No
Contribution to Project:
Working on the development of PCR primers for the V-tag sequencing studies.

Name: Halliday, Elizabeth
Worked for more than 160 Hours: No
Contribution to Project:
Graduate Student working in PI Gast's lab.

Undergraduate Student

Name: Bobb-Semple, Aisha
Worked for more than 160 Hours: No
Contribution to Project:
Sophomore, biology. Participated under the auspices of the MIT UROP program; involved with strain characterization, environmental sampling and sequence analysis.

Name: Han, Jing
Worked for more than 160 Hours: No
Contribution to Project:
Sophomore, EECS. Participated under the auspices of the MIT UROP program; involved with strain characterization, environmental sampling and sequence analysis.

Name: Buchwald, Carolyn
Worked for more than 160 Hours: No
Contribution to Project:
Senior, EAPS. Participated under the auspices of the MIT UROP program; involved with strain characterization, environmental sampling and sequence analysis.

Name: Ngo, Lynn Ly
Worked for more than 160 Hours: No
Contribution to Project:
Graduated 2006, CE. Participated under the auspices of the MIT UROP program; involved with strain characterization, environmental sampling and sequence analysis.

Name: Chang, Sarah
Worked for more than 160 Hours: No
Contribution to Project:
Wellesley junior, history. Participated under the auspices of the MIT UROP program; involved with strain characterization, environmental sampling and sequence analysis.

Name: Proehl, Sarah
Worked for more than 160 Hours: No
Contribution to Project:
Sophomore, chemistry. Participated under the auspices of the MIT UROP program; involved with strain characterization, environmental sampling and sequence analysis.

Name: Smith, Sarah

Worked for more than 160 Hours: No

Contribution to Project:

Sophomore, chemistry. Participated under the auspices of the MIT UROP program; involved with strain characterization, environmental sampling and sequence analysis.

Name: Schieffer, Stella

Worked for more than 160 Hours: No

Contribution to Project:

Sophormore. Participated under the auspices of the MIT UROP program; involved with strain characterization, environmental sampling and sequence analysis.

Name: Yuen, Grace

Worked for more than 160 Hours: No

Contribution to Project:

Participated under the auspices of the MIT UROP program; they were involved with strain characterization, environmental sampling and sequence analysis.

Technician, Programmer

Name: Keafer, Bruce

Worked for more than 160 Hours: No

Contribution to Project:

Research Associate, is responsible for field work planning and execution, as well as with sampling and sample analysis.

Name: McCauley, Linda

Worked for more than 160 Hours: Yes

Contribution to Project:

Research Assistant, is responsible for testing and application of molecular methods for the population genetic studies, phytoplankton culture and characterization, and data review and analysis.

Name: Kosnyrev, Valery

Worked for more than 160 Hours: Yes

Contribution to Project:

Research Associate and is assisting Dr. McGillicuddy in processing and analysis of the various data sets, seagoing operations, running the coupled models, and visualizing the results.

Name: Anderson, Laurence

Worked for more than 160 Hours: No

Contribution to Project:

Research Associate and has significant seagoing experience, and will participate in the large-scale survey cruises and will support post-cruise processing of the data.

Name: Dennett, Mark

Worked for more than 160 Hours: Yes

Contribution to Project:

Assisted with sample collection and processing.

Name: Moran, Dawn

Worked for more than 160 Hours: Yes

Contribution to Project:

Processed samples and ameba cultures, pathogen detection, cloning, and sequencing.

Name: Rocca, Jennifer

Worked for more than 160 Hours: Yes

Contribution to Project:
Works in the core facility preparing DNA templates and sequencing reactions for members of the COHH team.

Name: Huse, Susan

Worked for more than 160 Hours: Yes

Contribution to Project:
Susan, Research Associate, has been developing methods to evaluate error rates associated with pyrosequencing technology and has been assessing the correlation of variable region distances with full-length 16s distances.

Name: Libera, Katherine

Worked for more than 160 Hours: No

Contribution to Project:
Responsible for testing and application of molecular methods for the population genetic studies, phytoplankton culture and characterization, and data review and analysis.

Name: del Castillo, Erika

Worked for more than 160 Hours: No

Contribution to Project:
Responsible for microbial community survey library construction and analysis.

Other Participant

Name: Dupuis, Erin

Worked for more than 160 Hours: Yes

Contribution to Project:
Helps with grant management, report and proposal preparation, organization of meetings, record-keeping, and other administrative duties. Support is from NSF and NIH as well as institutional funds from the Woods Hole Oceanographic Institution.

Name: Cormier, Bonnie

Worked for more than 160 Hours: Yes

Contribution to Project:
Helps with grant management, report and proposal preparation, organization of meetings, record-keeping, and other administrative duties. Support is from NSF and NIH as well as institutional funds from the Woods Hole Oceanographic Institution.

Research Experience for Undergraduates

Organizational Partners

Massachusetts Institute of Technology

Marine Biological Laboratory

Other Collaborators or Contacts

University of Washington

Pacific Research Center for Marine Biomedicine at the University of Hawaii

University of Miami

Harvard Medical School

Fisheries Research Agency of Japan, National Research Institute of Fisheries and Environment of Inland Sea

Woods Hole United States Geological Survey

Drs. Uwe John and Allan Cembella, Alfred Wegener Institute, Bremerhaven Germany

Dr. Satoshi Nagai, National Research Institute of Fisheries and Environment of the Inland Sea, Japan

Plum Island Sound Estuary Long Term Ecological Research (PIE-LTER) program in collaboration with Charles Hopkinson and Hap Garritt (MBL)

Dr. Wayne Wurtsbaugh, Utah State University

Dr. Philip Roberts, Georgia Tech

Dr. Sandra McLellan at the Great Lakes WATER Institute in Milwaukee,WI

Activities and Findings

Research and Education Activities: (See PDF version submitted by PI at the end of the report)
See attached file.

Findings: (See PDF version submitted by PI at the end of the report)
See attached file.

Training and Development:
See attached file.

Outreach Activities:
See attached file.

Journal Publications

Erdner, D.L. and D.M. Anderson, "Global transcriptional profiling of the toxic dinoflagellate Alexandrium fundyense using Massively Parallel Signature Sequencing.", BMC Genomics, p. 1, vol. 7:88, (2006). Published,

Nagai, S., L. McCauley, N. Yasuda, D.L. Erdner, D.M. Kulis, Y. Matsuyama, S. Itakura and D.M. Anderson., "Development of microsatellite markers in the toxic dinoflagellate Alexandrium minutum (Dinophyceae)", Molecular Ecology Notes, p. 756, vol. 6, (2006). Published,

Anderson, D.M., Keafer, B.A., McGillicuddy, D.J., Mickelson, M.J., Keay, K.E., Libby, P.S., Manning, J.P., Mayo, C.A., Whittaker, D.K., Hickey, J.M., He, R., Lynch, D.R., Smith, K.W., "Initial observations of the 2005 Alexandrium fundyense bloom in southern New England: General patterns and mechanisms", Deep Sea Research II, p. 2856, vol. 52, (2005). Published,

Martins, C.A., "Functional genomics of a non-toxic culture of Alexandrium lusitanicum.", Ph.D. Thesis, MIT/WHOI Joint Program in Oceanography, p. , vol. , (2006). Ph.D. Thesis,

Ralston D.K., McGillicuddy, D.J. and D.W. Townsend, "Asynchronous vertical migration and bimodal distribution of motile phytoplankton",

Journal of Plankton Research

, p. , vol. , ( ). Submitted,

Acinas, S.G, Sarma-Rupavtarm, R., Klepac-Ceraj, V., Polz, M.F., "PCR induced sequence artifacts and bias: insights from two 16S rRNA clone libraries constructed from the same sample. Appl. Environ. Microbiol.", Appl. Environ. Microbiol., p. 8966, vol. 71(12), (2005). Published,

Klepac-Ceraj, V., Ceraj, I., Polz, M.F., "CLUSTERER: extendable java application for sequence grouping and cluster analyses", Online J. Bioinf., p. 15, vol. 7(1, (2006). Published,

Hunt, D.E., Klepac-Ceraj, V., Acinas, S.G., Gauthier, C., Bertilsson, S., Polz, M.F., "Evaluation of 23S rRNA PCR Primers for use in Phylogenetic Studies of Bacterial Diversity.", Appl. Environ. Microbiol., p. 2221, vol. 72(3), (2006). Published,

Marcelino, L., Backman, V., Donaldson, A., Steadman, C., Thompson, J.R., Paccocha-Preheim, S., Lien, C., Lim, E., Veneziano, D., Polz, M.F., "Accurate identification of low abundant targets amidst similar sequences by revealing hidden correlations in oligonucleotide microarray data.", Proc. Natl. Acad. Sci. USA

, p. 13629, vol. 103(37), (2006). Published,

Polz, M.F., Hunt, D.E., Preheim, S.P., Weinreich, D.M., "Patterns and mechanisms of genetic and phenotypic differentiation in marine microbes. Phil. Trans. R. Soc. Lond. B.", Phil. Trans. R. Soc. Lond. B., p. 2009, vol. 361, (2006). Published,

Veneziano, D., Klepac-Ceraj, V., Polz, M.F., "Likelihood estimation of richness and species abundance distribution in microbial communities.", Journal of Theoretical Biology, p. , vol. , ( ). Submitted,

Sogin, M.L., H.G. Morrison, J.A. Huber, D.Mark Welch, S.M. Huse, P.R. Neal, J.M. Arrieta, and G.J. Herndl., "Microbial diversity in the deep sea and the under-explored "rare biosphere.", Proc. Natl. Acad. Sci. USA, p. 12115, vol. 103(32), (2006). Published,

Messerli, S.M. and Greenberg, R.M., "Cnidarian toxins acting on voltage-gated ion channels.
", Marine Drugs, p. 70, vol. 4, (2006). Published,

Ehrenreich, I.M., J.B. Waterbury, and E.A. Webb, "The Distribution and Diversity of Natural Product Genes in Marine and Freshwater Cyanobacterial Cultures and Genomes.", Appl. Environ. Microbiol., p. 7401, vol. 71, (2005). Published,

Anderson, D.M., J.M. Burkholder, W.P. Cochlan, P.M. Glibert, C.J. Gobler, C.A. Heil, R. Kudela, M.L. Parsons, J.E. Rensel, D.W. Townsend, V.L. Trainer, and G.A. Vargo (in review)., " Harmful algal blooms and eutrophication: Examples of linkages from selected coastal regions of the United States.", Harmful Algae., p. , vol. , (2007). Submitted,

Erdner, D.L. et al. (in review), "Centers for Oceans and Human Health: A unified approach to the challenge of harmful algal blooms.", Environmental Health., p. , vol. , (2007). Submitted,

Heisler, J., P. Glibert, J. Burkholder, D. Anderson, W. Cochlan, W. Dennison, C. Gobler, Q. Dortch, C. Heil, E.Humphries, A. Lewitus, R. Magnien, H. Marshall, K. Sellner, D. Stockwell, D. Stoecker, and M. Suddleson (in review)., "Eutrophication and harmful algal blooms: A scientific consensus.", Harmful Algae., p. , vol. , (2007). In press,

Ho, A.Y.T., J. Xu, X.C. Yuan, K. Yin, L. He, Y.L. Jiang, D.M. Anderson, and P.J. Harrison (in review)., "Seasonal and spatial dynamics of nutrients and phytoplankton biomass in Victoria Harbour and its vicinity before and after sewage abatement.", Mar. Poll. Bull., p. , vol. , (2007). Submitted,

Lilly, E.L., K. Halanych, and D.M. Anderson (in press)., "Species boundaries and global biogeography of the dinoflagellate ?A. tamarense? complex of the dinoflagellate genus Alexandrium (Dinophyceae).", J. Phycol., p. , vol. , (2007). In press,

McCauley, L.A.R., D.L. Erdner, S. Nagai, and D.M. Anderson (in review), "Bioegeographic analysis of the globally distributed harmful algal bloom species Alexandrium minutum (Dinophyceae) based on LSU rDNA and ITS sequences, and microsatellite markers.", Journal of Phycology., p. , vol. , (2007). Submitted,

Xu, Jie, A.Y.T. Ho, K. Yin, X. Yuan, D.M. Anderson, and P.J. Harrison (in review)., "Temporal and spatial variations in nutrient stoichiometry and regulation of phytoplankton biomass in Hong Kong waters: Influence of the Pearl River outflow and sewage inputs.", Mar. Poll. Bull., p. , vol. , (2007). Submitted,

Backer, L.C. and McGillicuddy, D.J.,, "Harmful Algal Blooms: At the interface between coastal oceanography and human health.", Oceanography, p. 94, vol. 19, (2006). Published,

Ralston D.K., McGillicuddy, D.J. and D.W. Townsend, "Asynchronous vertical migration and bimodal distribution of motile phytoplankton.", Journal of Plankton Research, p. , vol. , (2007). Published,

He, R. and D.J. McGillicuddy., "Gulf of Maine Harmful Algal Bloom in summer 2005 - Part 1: In situ Observations of Coastal Hydrography and Circulation.", Journal of Geophysical Research., p. , vol. , (2007). Submitted,

He, R., McGillicuddy, D.J., Keafer, B.A. and D.M. Anderson., "Gulf of Maine Harmful Algal Bloom in summer 2005 - Part 2: Coupled Bio-physical Numerical Modeling.", Journal of Geophysical Research, p. , vol. , (2007). Submitted,

Smith, K.W., McGillicuddy, D.J., and D.R. Lynch., "Parameter estimation using an ensemble smoother: the effect of the circulation in biological estimation.", Journal of Marine Systems, p. , vol. , (2007). Submitted,

Anderson, D.M., Libby, P.S., Mickelson, M.J., Borkman, D.G., He, R., McGillicuddy, D.J., "The 2005 New England red tide of Alexandrium fundyense: observations, causes, and potential outfall linkages.", Boston: MWRA Report 2007-10, p. , vol. , (2007). Published,

Aretxabaleta, A.L., McGillicuddy, D.J. , Smith, K.W., and D.R. Lynch, "Model Simulations of the Bay of Fundy Gyre: 1. Climatological Results", Journal of Geophysical Research, p. , vol. , (2007). Submitted,

Dyble, J., Bienfang, P., Dusek, E., Griffiths, W., Hitchcock, G., Holland, F., Laws, E., Lerczak, J., McGillicuddy, D.J., Minnett, P., Moore, S., O?Kelly, C., Solo-Gabriel, H., and J. Wang., "Environmental controls, oceanography and population dynamics of pathogens and harmful algal blooms: Connecting sources to human exposure.", Environmental Health, p. , vol. , (2007). Published,

Jill R. Stewart, Rebecca J. Gast, Roger S. Fujioka, Helena M. Solo-Gabriele, J. Scott Meschke, Linda A. Amaral-Zettler, Ericka Del Castillo, Martin F. Polz, Tracy K. Collier, Mark S. Strom, Christopher D. Sinigalliano, Peter D. R. Moeller, and A. Fredrick, "The coastal environment and human health: Microbial indicators, pathogens, sentinels and reservoirs", Environmental Health, p. , vol. , (2007). Submitted,

Shimeta, J., R.J. Gast and J.M. Rose, "Community structure of marine sedimentary protists in relation to flow and grain size", Aquatic Microbial Ecology, p. , vol. , (2007). Published,

Sinigalliano, CD, Gidley, ML, Shibata, T, Whitman, D, Dixon, TH, Laws, E, Hou, A, Bachoon, D, Brand, L, Amaral-Zettler, L, Gast, R, Steward, GF, Nigro, OD, Fujioka, R, Betancourt, WQ, Vithanage, G, Mathews, J, Fleming, LE and HM Solo-Gabriele., "Impact of Hurricanes Katrina and Rita on the microbial landscape of the New Orleans area", PNAS, p. , vol. , (2007). Published,

Bogomolni A, Ellis J, Gast B, Harris R, Pokras M, Touhey K, Moore M, "Emerging Zoonoses in Marine Mammals and Seabirds of the Northeast U.S. Oceans '06", MTS/IEEE-Boston, Massachusetts September 18-21, 2006, p. , vol. , (2006). Published,

Sabehi, G. Kirkup, B. C., Rozenberg, M., Stambler, N., Polz, M.F., Beja, O., "Niche adaptation and spectral tuning in marine proteorhodopsins", ISME Journal, p. , vol. , (2007). Published,

Hunt, D. E., Gevers, D., Vahora, N. M., Polz, M. F., "Conservation of the chitin utilization pathway in the Vibrionaceae.", Appl. Environ. Microbiol., p. , vol. , (2008). Published,

Stocker, R, Seymour, J. R., Samadani, A., Hunt, D. E., Polz, M. F., "Rapid chemotactic response enables marine bacteria to exploit ephemeral microscale nutrient patches.", Proc. Natl. Acad. Sci., p. , vol. , (2007). In press,

Thompson, J.T., Marcelino, L., Polz, M.F., "Diversity and sources of human bacterial pathogens and overview of methods of their detection and quantification. In: Shimshon Belkin and Rita Colwell (Eds.), Ocean and Health: Pathogens in the Marine Environment.", Textbook: 'Oceans Health', p. , vol. , (2007). Published,

Huse, S.M., J.A. Huber , H.G. Morrison, M.L. Sogin and D. Mark Welch, "Accuracy and quality of massively parallel DNA pyrosequencing.", Genome Biology, p. , vol. , (2007). Published,

Pope, W. H., Weigele, P. R., Chang, J., Pedulla, M. L., Ford, M. E., Houtz, J. M., Jiang, W., Chiu, W., Hatfull, G. F., Hendrix, R. W., and King, J., " Genome sequence, structural proteins, and capsid organization of the cyanophage Syn5: a "horned" bacteriophage of marine Synechococcus.", Journal of molecular biology, p. , vol. , (2007). Published,

Weigele, P. R., Pope, W. H., Pedulla, M. L., Houtz, J. M., Smith, A. L., Conway, J. F., King, J., Hatfull, G. F., Lawrence, J. G., and Hendrix, R. W., "Genomic and structural analysis of Syn9, a cyanophage infecting marine Prochlorococcus and Synechococcus.", Environmental microbiology, p. , vol. , (2007). Published,

H.L. Kite-Powell, L.E. Fleming, L.C. Backer, E. Faustman, P. Hoagland, A. Tsuchiya, L. Younglove, B.A. Wilcox, and R. Gast., "Linking the oceans to public health: Where is the ?human health? in ?oceans and human health??", Environmental Health, p. , vol. , (2007). Submitted,

Books or Other One-time Publications

Thompson, J.T., Marcelino, L., Polz, M.F., "Diversity and sources of human bacterial pathogens and overview of methods of their detection and quantification.", (2006). Book, Published
Editor(s): Shimshon Belkin and Rita Colwell
Bibliography: Ocean and Health: Pathogens in the Marine Environment. Springer. p. 464.

Martins, C.A., "Functional genomics of a non-toxic culture of Alexandrium lusitanicum.", (2006). Thesis, Published
Collection: MIT/WHOI Joint Program in Oceanography
Bibliography: N/A

Web/Internet Site

URL(s):
http://www.whoi.edu/science/cohh/whcohh/index.htm
Description:
Additional Web Sites:
http://www.whoi.edu/sbl/liteSite.do?litesiteid=3230&articleId=13371
http://science.whoi.edu/users/ruoying/Redtide_05/movie.html
http://science.whoi.edu/users/ruoying/Redtide_06/
http://science.whoi.edu/users/mcgillic/cohh/oc412/data/
www.whoi.edu/people/rgast
www.whoi.edu/sites/zoonoses
http://science.whoi.edu/users/mcgillic/cohh/oc412/data/
http://science.whoi.edu/users/mcgillic/en435/
http://science.whoi.edu/users/mcgillic/en437/
http://science.whoi.edu/users/ruoying/Redtide_05/movie.html
http://omgrhe.meas.ncsu.edu/Redtide/Redtide_06/
http://omgrhe.meas.ncsu.edu/Redtide/Redtide_07/

Future web site:
Genomics Facility Core: We are working on a web site principally with funding from the Sloan Foundation to make available the 454 sequence tags including taxonomic breakdown. This web site (MICROBIS) will be released sometime during the first or second quarter of 2007.

Other Specific Products

Product Type:

Poster presentation

Product Description:

04/06: Initial Observations of the 2005 Alexandrium fundyense Bloom in southern New England: General Patterns and Mechanisms. Symposium on Boston Harbor and adjacent coastal waters. New England Estuarine Research Society. Nantasket, MA.

11/06: D.M. Anderson, B.A. Keafer, D.J. McGillicuddy, and R. He. The 2005 and 2006 Alexandrium red tides: A tale of two blooms. Wellfleet Harbor Conference.

Sharing Information:

Presented at meetings.

Product Type:

Software (or netware)

Product Description:

We are currently collaborating with colleagues in Austria and Germany to develop widely applicable microarray software based on our paper Marcelino et al. (2006).

Sharing Information:

Will be available to interested researchers.

Product Type:

Software (or netware)

Product Description:

Env454 database schema; in-house scripts to assemble reads generated from the same PCR clone with different sequencing primers (?clonebyclone?) and to estimate the taxonomic distribution of clones amplified from a particular sample (?clone2taxbreak?).

Sharing Information:

Service using database will be available to interested researchers.

Product Type:

Software (or netware)

Product Description:

We are currently collaborating with colleagues in Austria and Germany to develop widely applicable microarray software based on our paper Marcelino et al. (2006).

Further, we have developed the AdaptML software, which can be used to detect ecologically cohesive bacterial populations among strains sampled from varied environmental samples.

Sharing Information:

See above.

Contributions

Contributions within Discipline:

Anderson û Project 1 & Project 2:
Our microsatellite dataset encompasses three years of toxic Alexandrium blooms and is the first multi-year dataset on the genetic composition of and changes in a toxic bloom population. The data support our primary hypothesis that bloom populations are heterogeneous, and have led to further hypotheses about how blooms form and persist, and about the genetic heterogeneity of the different, widely spatially separated source

populations (cyst seedbeds) and how they are formed.

The auxiliary data collected during our field efforts (cell counts, cyst maps, hydrographic information) have proved extremely useful in our continued collaboration with Dr. Dennis McGillicuddy of the Woods Hole COHH. They have provided valuable data for use in numerical model runs focusing on the 2005, 2006, and 2007 bloom seasons. The results of these 'hindcast' analyses are of great importance for understanding the factors that affect the magnitude of toxic blooms.

Gast û Project 3:
We have been able to provide standardized methods for the collection and rapid processing of numerous environmental water samples. These methods were used by multiple SGER projects collaboratively examining the effects of Hurricane Katrina on the microbes of Lake Pontchartrain.

Polz û Project 4:
We have obtained funding from the Moore Foundation to sequence the genome of several of the strains isolated in this project. This is a significant step in providing marine microbiology with the necessary core dataset for future research. Annotation and correlation with data from the other ongoing sequencing efforts will provide additional value to the users of these genomes.

Sogin û Genomics Core Facility:
The tag sequencing technology has the potential to impact all of the Centers of Oceans and Human Health. It provides a new tool for detailed monitoring of microbial populations in marine environments. This will allow detection of a wide range of human pathogens even before they become significant fractions of the analyzed populations. We have submitted a proposal to the Sloan Foundation that will characterize entire microbial populations in an investigation of the Falmouth MA, Boulder CO, and Centennial CO water distribution systems ('The Rare Biosphere and the Human Habitat'). The proposed research will identify new indicator organisms and differential persistence of microorganisms associated with sewage pollution as a function of environmental conditions including physical forcings.

Hahn û Pilot Project Program:
The funded pilot projects address important questions regarding the interactions between the oceans and human health. Many of the projects represent new directions for the Center, expanding the scope of our research efforts into, for example, marine natural products (e.g. BMAA) and employing new technologies (e.g. signature tagged mutagenesis). Many of the funded or proposed pilot projects involve postdoctoral researchers or graduate students, and will thus contribute to the training of the next generation of researchers in this area.

Contributions to Other Disciplines:

Stegeman û Administrative Core:
To improve public health through an enhanced understanding of how oceanic processes affect the distribution and persistence of human pathogens and toxin producing organisms.

Sogin û Genomics Core Facility:
The tag sequencing technology can be applied to any microbial population structure study including marine, terrestrial and even the human biome.

Contributions to Human Resource Development:

Stegeman û Administrative Core:
The Center supported an undergraduate Summer Fellow in 2007, through the NSF-REU grant. The fellow was Lara Polansky, University of Miami, who worked with pilot project recipient Porter Hoagland on assessing the role of K. brevis blooms on emergency department respiratory diagnoses in Florida. Another student was supported through the MBL to work in Becky Gast's lab. Other students worked on COHH projects in the lab of Center Investigator Martin Polz at MIT. The Woods Hole Center also provided support to two students to attend the fall Harmful Algal Bloom conference in Woods Hole.

Polz û Project 4:
Undergraduate student research and training: The Polz lab has, over the past 5 years, hosted 14 undergraduate researchers who all stayed for more than one semester and have been working on their own projects. These included several students from other institutions, two underrepresented minorities and 12 women. This has resulted in 4 publications with 5 undergraduate co-authors.
Polz also participated in the INGEB Integrative Biology Workshop in Pßtzcuaro, Mexico (October 2006), which trains students from Latin

American countries.
High school student research and training: We are assisting students from the Thomas Jefferson High School for Science and Technology, Alexandria, in their project on detection and remediation of Vibrio contamination in seafood. Polz has met with a representative from the school and the lab is currently transferring protocols to the school for bacterial isolation and characterization. Moreover, the Polz lab has hosted high school students in research internships. Most recently, a student came from the University School, Milwaukee who worked on characterization of antagonistic bacterial interactions.
Members of the Polz lab actively participated in setting up a high school student laboratory experience centered on bacteriophage characterization by purification, electron microscopy and protein analysis [http://web.mit.edu/pweigele/www/phiG01/Site/Welcome.html]. This activity will be continued and has also led to sequencing of 5 phages isolated by the Polz lab. These will serve to train the next batch of students in basic bioinformatics tools.
Popular science: The Museum of Science has a formal outreach activity with the MIT Earth Systems Initiative. The Polz lab actively participates in this activity, which involves: (1) Yearly PI visit to the Museum with podcast, New England Cable News spot, and presentation; (2) Quarterly researcher 'appearance' at the museum to be available for informal conversation with museum guests following a museum staff presentation in which their research is referenced; (3) Weekly staff presentation in which the ESI research efforts are highlighted as part of a general presentation on 'Global Warming,' 'Earth Dynamics' (or similar topics); (4) Daily availability of research highlights within a touchstory attached to the Earth Vitrine.

Special activity in response to hurricane Katrina: Polz took a group of undergraduate students to Lake Pontchartrain to study potential effects of Hurricane Katrina. In a week-long activity, the students carried out independent projects ranging from assessment of populations abundance of potential pathogens and indicators (vibrios and coliforms) to determination of heavy metal contamination in sediments. The group stayed at Southeastern Louisiana University in Hammond and interacted with local scientists, students and engineers involved in post-hurricane cleanup and lake studies.

Contributions to Resources for Research and Education:

Anderson û Project 1:
With regard to human resource development, we have trained numerous undergraduate, graduate, and post-graduate students during this project.

Gast û Project 3:
Amaral-Zettler has been very active in teacher workshops and classroom interactions over the history of this project. Recently she participated as a science advisor in a webquest exercise. A webpage of the results from that workshop is available at http://serc.carleton.edu/microbelife/topics/pontchartrain/index.html

Amaral-Zettler gave a lecture entitled 'Oceans and Human Health' for a course featuring the application of molecular techniques in environmental science being offered by Falmouth Hospital. About 30 people attended including 12 or so physicians.

Sogin û Genomics Core Facility:
We have developed high capacity molecular sequence analysis software and curated sequence tag databases that will be valuable to the entire microbiology community.

Contributions Beyond Science and Engineering:

Anderson û Project 1:
During field projects, we maintained a web site providing detailed information about Alexandrium status: shellfishing closures, cell concentrations, drifter tracks, etc. (see http://www.whoi.edu/sbl/liteSite.do?litesiteid=3230). We also support a regional listserve that facilitated communication among managers and scientists. Both resource proved very useful to managers, scientists, and the general public, and will be maintained throughout this project.

Gast û Project 3:
This work is generating a baseline of microbial diversity for Mt Hope Bay that includes sequences of potential human pathogens that has not been described before. This information may be valuable to current discussions regarding potential dredging of a channel at the mouth of the Taunton River (nearby our sewage outfall and the power plant sites) and pending rulings against the power plant and its thermal discharge.

Gast has been contacted by several individuals in the past year who represent health and safety advisors as a result of several general articles on the presence of legionellae in saltwater. There are questions about whether biofilms within pipes and hoses used for saltwater distribution may be a reservoir for pathogenic legionellae.

Polz û Project 4:

The creation and intense characterization of this strain collection will permit it to be used in testing technologies meant to identify and enumerate pathogens from among environmental isolates, allowing the precise cause of false positives and negatives to be discerned by associating them with well known environmental strains.

Sogin û Genomics Core Facility:
The tag sequencing strategy has the potential to identify new indicator organisms and develop targets for DNA microarrays that will allow for more comprehensive tracking of microbes introduced into the environment through anthropogenic activity.
Using the GS20 (and later, the GSFLX) sequencer the core facility continued development of a new, massively parallel DNA sequencing approach to characterize microbial populations with unprecedented level of sensitivity for detecting low abundance organisms. The International Census of Marine Microbes (ICOMM) provided the test samples and funds for reagents and the WH-COHH provided computational support for analysis of the data. We described in our last report the use of this approach to assess the species richness and evenness in water samples from the North Atlantic Deep Water flow and from Axial seamount (Sogin et al., 2006). During 2007, we demonstrated the feasibility of using 5-nt 'keys' to multiplex reactions on the same region of a plate, deconvoluting the reads into sample bins informatically. This approach has been used on hundreds of samples from seawater, freshwater, sediment, soil, sewage, and human feces. We analyzed just over 1.3 million V6 tag sequences from human stool samples (500,000 from David Relman's laboratory and 800,000 generated in our laboratory as part of a collaboration with Jeffrey Gordon). This will provide a reference data for interpreting V6 tag sequences studies from human-impacted environments that we are studying in the context of a WH-COHH pilot project. Our data processing pipeline is well established, and datasets from tag and full length studies are displayed using a new web tool, Visualization of Microbial Population Structures (VAMPS), developed with support from the Sloan Foundation (awarded to M.L. Sogin and David Mark Welch).

Special Requirements

Special reporting requirements: None
Change in Objectives or Scope: None
Animal, Human Subjects, Biohazards: None

Categories for which nothing is reported:

Research and Education Activities, Findings, Training and Development, Outreach, and Products

Stegeman – Administrative Core

Research and Education Activities:

1. Foster the communication, planning, integration and interaction among Center members, thereby assuring that the goals developed in the Center vision are achieved.
2. Provide and maintain an administrative structure to oversee and monitor the financial aspects of the Center, including grants management.
3. Establish a structure that will support effective communication and planning with contribution from all Center Investigators.
4. Establish an External Advisory committee, which will review the Center programs and advise the Director and Deputy Director.
5. Oversee the operation of and work toward a successful Pilot Project small grant program.
6. Establish and maintain public and private web pages for the dissemination of information and for data transfer and communication.
7. Pursue interactions between the COHH and other Centers within the three Institutions, as well as with other COHH and EHS Centers, in order to expand the reach and impact of the center, and to expand the research base on issues relevant to the Center mission and themes.
8. Encourage the involvement of center investigators in community outreach and education efforts.

Findings:

1. As before, the Director and Deputy Director have continued to oversee the Center Office and the activities of each of the component projects and cores in the Center. We continue to build visibility in and impact through all three of the component institutions, the Woods Hole Oceanographic Institution (WHOI), the Marine Biological Laboratory (MBL) and the Massachusetts Institute of Technology (MIT).
2. The Center Director and staff (administrative professionals), and the respective Grants Management offices at the Woods Hole Oceanographic Institution and other institutions continue to monitor the accounts for each of the Research Projects, the Pilot Projects, and the Advanced Genomics Core.
3. Center Investigator meetings have been held monthly, on the 1st Friday of each month. In addition, the Director of the Administrator informs the Center members of activities and opportunities pertinent to Oceans and Human Health via frequent email. The meetings and notes address all aspects of the Center interactions and communications, internally as well as externally. The Director and Deputy worked to increase visibility of the Center.
4. The Administrative Core organized a day and a half retreat for the investigators in the WH-COHH in January 2008, at the nearby National Academy Study Center. The retreat focused overall progress in the research and possible new directions. There will be a similar meeting of all the Pilot Project awardees, to be held in February.
5. The Director met during the summer with the Chair of the External Advisory Committee (Dr. Michael Gallo, University of Medicine and Dentistry of New Jersey), to discuss the center's activities.

6. The Director and Deputy continued to work together with the director of the Pilot Project program (Dr. Mark Hahn), to coordinate the operation of the successful Pilot Project program. Four pilot projects were funded with 2006 funds, including one involving a joint project with the Miami Center for Oceans and Human Health. In excess of $40,000 was leveraged to augment these projects. The fourth call was issued in September of 2007 by email to all faculty and research staff at WHOI, MBL, and MIT. On November 1, 2007, nine proposals were received and subsequently were reviewed and scored by members of the Internal Advisory Committee and external reviewers, including several members of the other COHH Centers and other Ad Hoc reviewers. The proposals requested approximately $308,000 in total costs. All applicants were provided with written reviews. Three projects were selected for funding: Three were recommended for funding, two from WHOI and one from MIT. Additional funds in the program may be used to fund partially two other highly meritorious proposals, one from MIT and one from MBL. The distribution of awards to all three of the constituent institutions is a strength of the program.
7. The COHH website developed at WHOI has continued to serve the needs of the four primary OHH Centers, Woods Hole, Hawaii, Miami and Washington (http://www.whoi.edu/science/cohh/). COHH links continue to be added and an ftp site (internal) is in use.
8. Interactions with other groups continue to grow.
    a. Interactions between the WH-COHH and the other COHH. The Director and Deputy Director, with the help of the whole Woods Hole Center, hosted the third joint Center Directors' and Investigators meeting, in Woods Hole, in April, 2007. This meeting was highly successful and resulted in 5 manuscripts focused on progress in different areas of OHH, which have been completed and are under review for publication in the BiomedCentral journal *Environmental Health*, a journal with a very good impact factor. The meeting also resulted in a collation of outlines of several academic courses on Oceans and Human Health that have been developed over the past years by investigators in Woods Hole Center and in other OHH centers.
    b. We have continued to participate in monthly conference calls among the leadership of the four NSF-NIH Centers, and a second monthly call that includes the NOAA OHHI leaders, to discuss points in collaboration and interaction in all aspects of the Centers' activities.
    c. Interactions between the WH-COHH and other NIEHS Centers: We are continuing to explore opportunities for jointly sponsored enrichment activities, as an outgrowth of the Center Directors meeting in April. These interactions are being facilitated by the Supplemental funds obtained from NSF in 2007.

Outreach and Impact:

The Center has been involved in informing the wider scientific community. We also have continued to encourage the involvement of center investigators in community outreach and education efforts. There have been several major activities this year.

a. Drs. Stegeman and Gast made presentations to the Massachusetts Shellfish Officers Association quarterly meeting, describing COHH activities and discussing possible interactions.

b. Dr. Stegeman and Dr. Lora Fleming of the Miami COHH have worked to develop an exciting agenda for the new Gordon Research Conference supported meeting on Oceans and Human Health, scheduled for 2008. We have enlisted and confirmed participation of the full slate of speakers and discussions leaders and are seeking funding to support this exciting new conference.
c. Dr. Amaral-Zettler gave a lecture on Oceans and Human Health in a course for physicians and health care professionals, at Falmouth Hospital. The Course, entitled "Molecular Biology's role in Modern Medicine, was organized by Genomics Core Director Mitch Sogin.
d. The Administrative Core and the particularly the Director helped Center Investigators with the preparation of the SGER grant application submitted to the NSF.
e. The Center also sponsored several seminars by researchers from other institutions, and provided support fore a visit by Dr. Debashish Battacharya (University of Iowa) to the Center.

Training and Development:
The Center supported an undergraduate Summer Fellow in 2007, through the NSF-REU grant. The fellow was Lara Polansky, University of Miami, who worked with pilot project recipient Porter Hoagland on assessing the role of *K. brevis* blooms on emergency department respiratory diagnoses in Florida. Another student was supported through the MBL to work in Becky Gast's lab. Other students worked on COHH projects in the lab of Center Investigator Martin Polz at MIT. The Woods Hole Center also provided support to two students to attend the fall Harmful Algal Bloom conference in Woods Hole.

Significance:
As before, the Administrative Core has overseen growth of the Center and facilitated the activities of all of the units, through communication and support. The Center has had important successes in the research projects and the pilot project program. Inter-center communication is robust. Funds continue to be leveraged to support and expand the activities of the Center. Center activities include research that has had direct consequences for the public health. The Administrative Core is the focal point for all Center activities.

Plans:
As in the past, during the next project period we will continue to oversee the management of the grant and the Center, to foster intra- and inter-center communication and collaboration, to increase the activities in out-reach and enrichment, and to seek additional funding sources that might be leveraged to maintain a robust pilot project program. During the next year the WH-COHH also will continue recruiting undergraduate student researchers through the NSF-funded REU program. We will participate in the fourth COHH meeting in Hawaii, and work to the success of the new Gordon Conference on OHH.

Research and Education Activities, Findings, Training and Development, Outreach, and Products

Anderson – Project 1

Activities:
Work in Year 4 has focused on Specific Aims 3 and 5 of the project:
3) Characterize the relationships between toxicity, physiological variability and genotype in *Alexandrium* spp. from the Gulf of Maine; and
5) track changes in the genotypic diversity of *Alexandrium* populations through time throughout the Gulf of Maine.

Findings:
In this project year, we continued our studies of *Alexandrium* bloom populations to include a third year of data from summer of 2007. In 2005, we collected and analyzed samples from a massive toxic *Alexandrium* bloom that occurred that summer. Results of the microsatellite genotyping showed that the late-bloom and early-bloom populations were significantly different from each other (Fisher's combined test, $p<0.05$), indicating that the genetic composition of the bloom population changed on the order of about 3 weeks. The observed change could be due to natural succession of the bloom community, or by the addition of new genotypes from a separated population, or both.

These two hypotheses were addressed during fieldwork performed during 2006, when we collected samples from across the Gulf of Maine region during a second *Alexandrium* bloom (summer of 2006) as well as from a toxic bloom in an isolated embayment, Salt Pond, MA. In the wider Gulf of Maine, populations were collected from across the region including offshore on Georges Bank, and none were significantly different from one another. A comparison of the 2005 and 2006 bloom samples showed that, in general, populations from the Gulf of Maine blooms in the two different years were not genetically distinct. Populations collected from Salt Pond, however, were genetically distinct from those in the wider Gulf of Maine, and they changed over the course of the 3-week bloom. From these two years of data, it appears that overall genetic composition of *Alexandrium* blooms in the Gulf of Maine is not significantly different from year to year. Within a year, however, we did observe changes in bloom populations on the timescale of approximately one month. This could result from the natural progression or 'turnover' of genotypes during a bloom, or from the mixing of genetically distinct cells from other (unknown) sources. Results of the 2006 analysis of the Salt Pond bloom provides support for the former hypothesis, although the mixing of different source populations cannot be discounted.

We continued our population genetic analysis in 2007, sampling repeatedly in 9 locations across the Gulf of Maine. We are in the process of genotyping over 1000 isolates from 2007, including biweekly samples from the offshore Georges Bank region and the most northerly "upstream" population in the Bay of Fundy. Although *Alexandrium* cells were not as numerous in 2007 as they were in the previous years, the monitoring cruises did reveal high numbers of cells in the offshore region of Georges Bank. The discovery of a large offshore population of cells when inshore areas had much lower numbers could have significant implications for *Alexandrium*

bloom ecology and shellfish toxicity in the region. The Georges Bank region is the focus of activity for a separate project headed by PI Anderson. Supplementary funding from NSF for COHH activities will enable us to send personnel on these cruises to collect samples to continue our population genetic analysis, with the aim of understanding the connection between inshore and offshore *Alexandrium* populations.

In this project year we have also initiated studies of the link between toxicity and physiology in *Alexandrium*. This includes paired growth rate and toxicity measurements on a large number of cultured *Alexandrium* collected from different years. The first experiment used 45 cultures isolated during 2001, and determined growth rates and toxicity under constant conditions. A 1.75-fold difference in maximum growth rates was observed amongst the cultures, ranging from 0.41 to 0.72 divisions/day. A second growth rate experiment used a total of 44 clonal isolates grown at 15°C and 6°C at a light level of either ~400 and 100 $\mu E \cdot m^{-2} \cdot s^{-1}$. Growth rates were determined for all isolates, under each of the four conditions. Toxicity measurements are currently underway. This dataset of matching information on genotype, growth rate, and toxicity will be used to examine the relationships between these three characteristics of *Alexandrium* strains. Ultimately, the information will be used in a modeling effort by WH COHH PI McGillicuddy to examine the role of genetic heterogeneity in the abundance and overall toxicity of *Alexandrium* populations in the region.

During the course of our work, we have involved a number of graduate and undergraduate students, as listed above. These students have worked on several aspects of this project, including the cruise sample collection and isolation of the *Alexandrium* cultures for microsatellite analysis.

Outreach and Impact:

Presentations:

Anderson, D.M. Don't eat the clams: Managing the threat from the New England red tide. Frontiers in Environmental Sciences, Weekly Series on Emerging Issues in Environmental Health Sciences. (Invited presentation at NIEHS headquarters in the *Frontiers in Environmental Science* seminar series. The talk was also webcast, and can be viewed at: http://www.niehs.nih.gov/news/video/science/frontiers/)

Anderson, D.M., B.A. Keafer, K. Norton, D.J. McGillicuddy, R. He, C.H. Pilskaln, and d. Couture. *Alexandrium fundyense* cyst dynamics in the Gulf of Maine. Fourth Symposium on Harmful Algae in the U.S. October 2007, Woods Hole, MA.

Anderson, D.M., B.A. Keafer, D.J. McGillicuddy, and R. He. The 2005 and 2006 New England red tides: mechanisms, management challenges, and implications for future forecasting capabilities", Estuarine Research Foundation meeting, November 2007, Providence, RI.

Erdner, D.L., L.A.R. McCauley, K. Libera, and D.M. Anderson. 2007. Population genetics of toxic *Alexandrium* blooms in the Gulf of Maine. Fourth Symposium on Harmful Algae in the U.S. October 2007, Woods Hole, MA.

Erdner, D.L., L.A.R. McCauley, K. Libera, and D.M. Anderson. 2007. A real-time PCR assay for the toxic dinoflagellate *Alexandrium fundyense*: laboratory studies and field validation. American Society of Limnology and Oceanography international meeting, February 2007, Santa Fe, NM.

Posters:

Brosnahan, M.L., R.J. Olson, D.L. Erdner, and D.M. Anderson. Evidence of a self-recognition system in the sexual life-cycle of *Alexandrium tamarense*. Fourth Symposium on Harmful Algae in the U.S. October 2007, Woods Hole, MA.

Products:

Websites Created

The Northeast PSP site (http://www.whoi.edu/sbl/liteSite.do?litesiteid=3230&articleId=13371)

Research and Education Activities, Findings, Training and Development, Outreach, and Products

McGillicuddy – Project 2

Research and Education:

1. Formulate a suite of population dynamics models for the various genotypes of *A. fundyense*. Model formulation will be guided by existing observations, as well as laboratory experiments to be conducted in Project 1 (Anderson).
2. Incorporate the ensemble of population models into existing models of Gulf of Maine coastal hydrodynamics.
3. Use the coupled physical-biological models to construct hindcast simulations of *A. fundyense* survey observations to be collected jointly with Project 1 (Anderson).
4. Diagnose the simulations to determine the processes regulating the space/time expression of the different genotypes in terms of *A. fundyense* abundance.
5. Utilize toxigenicity data for the various genotypes (provided by Project 1 (Anderson)) together with the coupled physical-biological models (Aim 3) to make predictions of shellfish toxicity along the coast. Toxicity predictions will be tested with observations from ongoing shellfish monitoring programs.

Findings:

Our effort this year was focused on execution of two research cruises, voyages 435 and 437 of R/V *Endeavor*. The first survey (Figure 1, left panel) revealed surprisingly low cell concentrations in the Gulf of Maine, in contrast to the severe blooms that took place in 2005 (Anderson et al., 2005) and 2006. For example, we can compare the southwestern block of our survey plan, from Cape Cod to Cape Ann, with a nearly identical survey conducted on R/V *Tioga* during exactly the same time period in 2006 (Figure 2). Surface live counts in 2007 were all zeros except for one station in which a single cell was observed (corresponding to a concentration of 14 cells per liter). This constitutes a dramatic change from this same time period in 2006, when most of the area was covered by several hundred to several thousands of cells per liter. In interpreting this comparison, keep in mind that the 2007 observations were collected during and in the aftermath of significant wind forcing, resulting in surface mixed layers of up to 40m in some places. Therefore the possibility of subsurface populations cannot be discounted.

Another major finding of our 2007 field work was a large bloom of *A. fundyense* on Georges Bank (Figure 1, left panel). Highest abundances occurred along the Southern Flank, with peak concentrations just over 13,000 cells per liter. The population extended southwest of Georges Bank along the outer continental shelf, consistent with the southwestward exit pathway from the bank at the southern end of the Great South Channel.

Our second survey (EN437) revealed persistence of the bloom on Georges Bank (Figure 1, right panel). Surface live counts indicate that cell concentrations on Georges Bank dropped somewhat since EN435, although there are several stations with thousands of cells per liter and a peak concentration near 10,000 cells per liter. The bank-wide pattern also changed, insofar as the central crest of the bank became a local minimum in concentration. Cell concentrations on the

southern New England shelf rose from near zero during EN435 to several hundred cells per liter during EN437.

Whereas the EN435 survey revealed near absence of *A. fundyense* in the coastal Gulf of Maine, cell concentrations in that area during EN437 were fairly similar to what has been observed in the past this time of year (Figure 1, right panel). A key question is how did the system "catch up" to "normal" conditions given such a late start? This question is even more puzzling in light of the severe blooms that took place in 2005 and 2006.

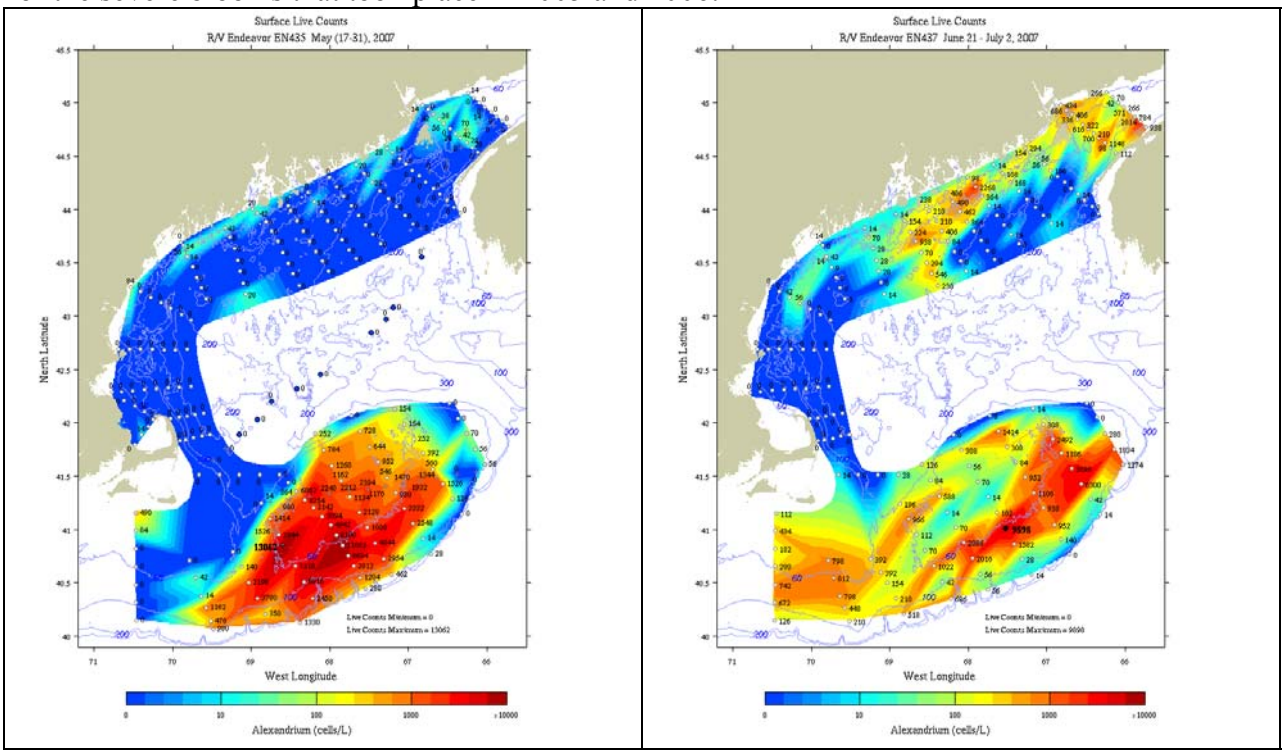


Figure 1. *Alexandrium* concentration (cells l^{-1}) from surface live counts on EN435 (left) and EN437 (right).

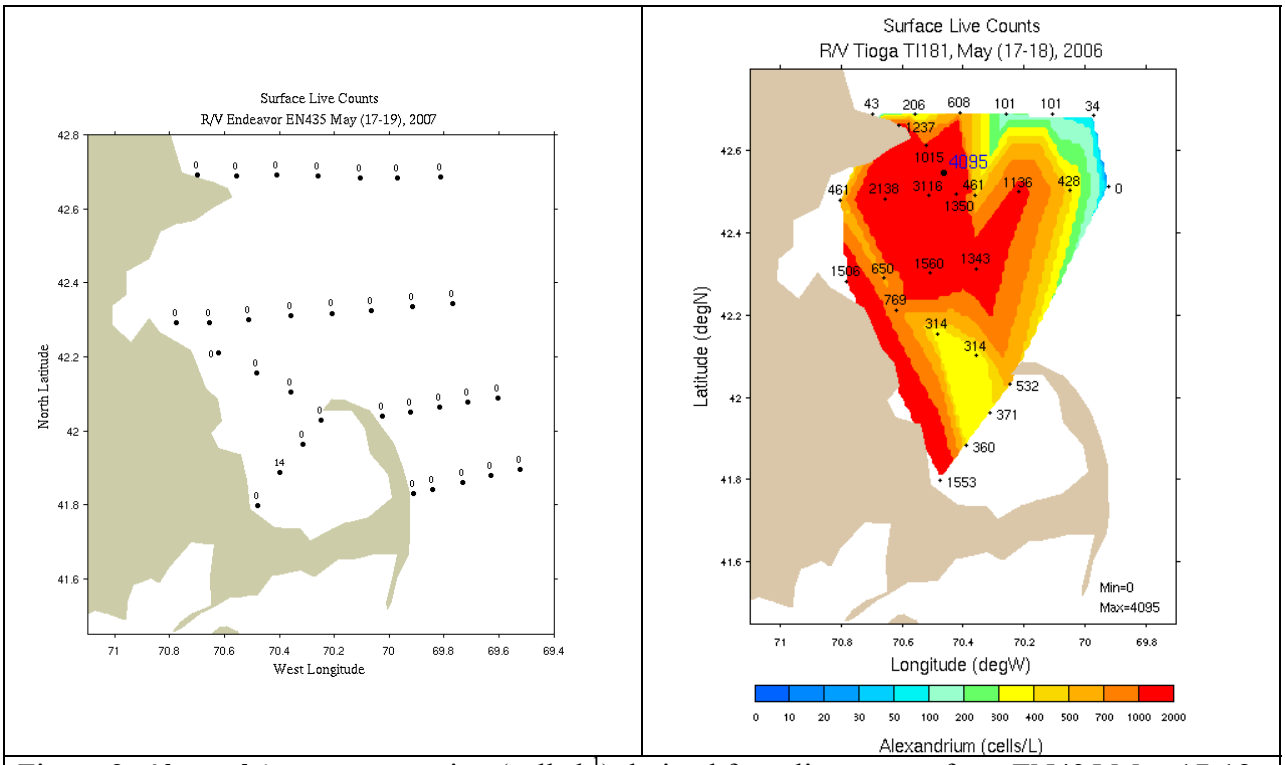


Figure 2: *Alexandrium* concentration (cells l^{-1}) derived from live counts from EN435 May 17-18 2007 (left) versus TI181 May 17-18 2006 (right).

Also in the past year, we developed a simple theory for the observed bimodal vertical distribution of motile phytoplankton (Ralston et al., 2007). Some motile phytoplankton have the capability to exploit deep sources of nutrients in a vertical migration cycle: photosynthesis in the near-surface layer, transit to depth, uptake of the limiting nutrient, and transit back to the surface layer. If all four steps can be completed within 24 hours, then migrations can be synchronized to the day/night cycle to maximize photosynthetic efficiency. Alternatively, if physiological, behavioral, or environmental factors make it impossible for the cycle to be completed in 24 hours, then migration may be asynchronous. Many observations of phytoplankton reveal bimodal vertical distributions of organisms, with maxima near the surface and the nutricline. We demonstrated how bimodal vertical distributions of phytoplankton may be symptomatic of asynchronous vertical migration using a Lagrangian Ensemble numerical model. We simulated vertical migration of the dinoflagellate *A. fundyense* in conditions similar to those in the Gulf of Maine, where bimodal distributions of *A. fundyense* have been observed. Migration is regulated by internal nutritional state – organisms swim down toward the nitracline when depleted of nitrogen, and return to the surface after nutrient uptake. We tested the sensitivity of the results to growth rate, nitrogen uptake rate, and swimming speed, and found that organism distributions can be bimodal or unimodal depending on conditions. Finally, we developed an analytical estimate for population distribution based on organism characteristics and nutricline depth.

Outreach and Impact:

10/06 Toxic algae in the Gulf of Maine: observations and models. WHOI Topics in Oceanography workshop for middle-school and high-school science teachers, WHOI Exhibit Center.

4/07 Modeling Blooms of *Alexandrium fundyense* in the Gulf of Maine: From Climatology to Forecasting. Environmental fluid mechanics seminar, Department of Civil and Environmental Engineering, MIT, Cambridge, MA (Invited).

4/07 Skill Assessment for Coupled Biological/Physical Models of Marine Systems. Ocean Color Research Team Meeting, Seattle, WA.

10/07 Observations and models of *Alexandrium fundyense* blooms in the Gulf of Maine and Georges Bank: From Climatology to Forecasting. Fourth Symposium on Harmful Algae in the U.S., Woods Hole, MA.

Products:

Websites Created:

1. Data from R/V *Oceanus* voyage 412, and R/V *Endeavor* cruises EN435 and EN437:
   http://science.whoi.edu/users/mcgillic/cohh/oc412/data/
   http://science.whoi.edu/users/mcgillic/en435/
   http://science.whoi.edu/users/mcgillic/en437/

2. Near-real-time nowcasting and forecasting of the 2005, 2006, and 2007 blooms:
   http://science.whoi.edu/users/ruoying/Redtide_05/movie.html
   http://omgrhe.meas.ncsu.edu/Redtide/Redtide_06/
   http://omgrhe.meas.ncsu.edu/Redtide/Redtide_07/

Research and Education Activities, Findings, Training and Development, Outreach, and Products

Gast – Project 3

Research and Education Activities:

This proposal examines the diversity, distribution and persistence of human pathogens in the coastal marine environment of Mt. Hope Bay, Massachusetts. This estuarine environment has been heavily impacted by human activity, including significant thermal pollution, several sewage outfalls and the effects of past industrial activities. Monitoring for bacterial contamination occurs in the shellfish areas, but outside of these regions, very little is known about the occurrence of either bacterial pathogens or protistan parasites that are likely introduced into the Bay through sewage or runoff. The potential impact of the thermal output of the Brayton Point Power Plant on these microbial populations has not been examined at all. We hypothesize that both bacterial and protistan human pathogens are present in Mt. Hope Bay and nearby rivers, and that their distribution and persistence are affected by physico-chemical condition in the Bay. To examine this hypothesis we propose to accomplish the following specific aims:

Aim 1: Determine the distribution and persistence of human protistan pathogens in Mt. Hope Bay. We will conduct a molecular survey of human pathogens within the Mt. Hope Bay estuarine system, including sewage outfalls near and removed from thermal point sources, contaminated shellfish beds and shorebird nesting grounds, to better characterize the types of protistan pathogens present in this coastal environment. Our project will specifically target *Giardia*, *Cryptosporidium*, *Naegleria*, and *Acanthamoeba*, but will also perform general microbial eukaryotic and prokaryotic diversity assessments to determine whether there are novel sequence types present that are related to known pathogens (might represent unidentified human pathogens).

Aim 2: Determine whether naturally occurring marine amoebae can serve as reservoirs for pathogenic bacteria, and look for evidence of the natural occurrence of these associations in Mt. Hope Bay. We will test whether locally collected marine amoebae are capable of harboring the pathogenic bacteria *Vibrio parahaemolyticus*, *Vibrio vulnificus*, and *Legionella pneumophila*. The research will test for the induced presence and replication of bacteria in actively growing and dormant phases of the amoebae, and will utilize both PCR-based and in situ hybridization based methods to detect naturally occurring associations.

Aim 3: Establish physico-chemical parameters associated with pathogen presence in Mt. Hope Bay and test predictive capability of water quality models regarding these correlations.
We will correlate measurements of water and sediment physico-chemical parameters with microbial community structure, including the presence of human pathogens. These measurements and our biological data will be used to assist in establishing and validating a water quality model for Mount Hope Bay that may also be used to understand and predict the potential presence of pathogens.

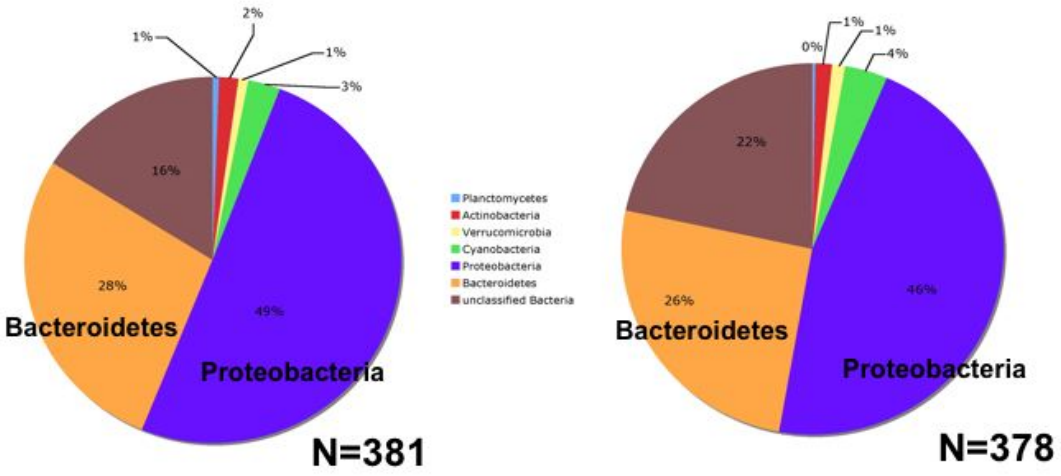
Findings:

Fieldwork. Four sites in Mt. Hope Bay were selected for our survey of human parasites and the effect of physico-chemical parameters on their distribution and persistence. One is located within the thermal plume of the Brayton Point power plant outlet. The second site (Braga Bridge) is located at an underwater sewage outfall that releases secondarily treated sewage into the Taunton River just a bit north of the Bay proper. A third site is farther up the Taunton River (Dighton Bridge) at a brackish water/marsh site, and the fourth site (Common Fence) is near the bottom of the Bay at an area not directly impacted by the thermal plume or the sewage outfall. Sediment and water samples were collected in triplicate at all four sites. Our seasonal sampling officially began in November 2004, and continued through April, June, July, August, November 2005 and February 2006. An additional sampling was carried out in June 2006 in response to large amounts of precipitation and increased runoff. In addition to our four sites we also conducted two transect sampling trips in August of 2005 and February of 2006. Each time our transect began at our Brayton Point site and included 7 stations along a direct line to our control site Common Fence. Transect sampling included water samples (collected at the surface and at depth just above the sediment) and sediment samples. Sediment and water samples were again also obtained and processed from five sites in the Great Salt Lake, Utah.

Results:

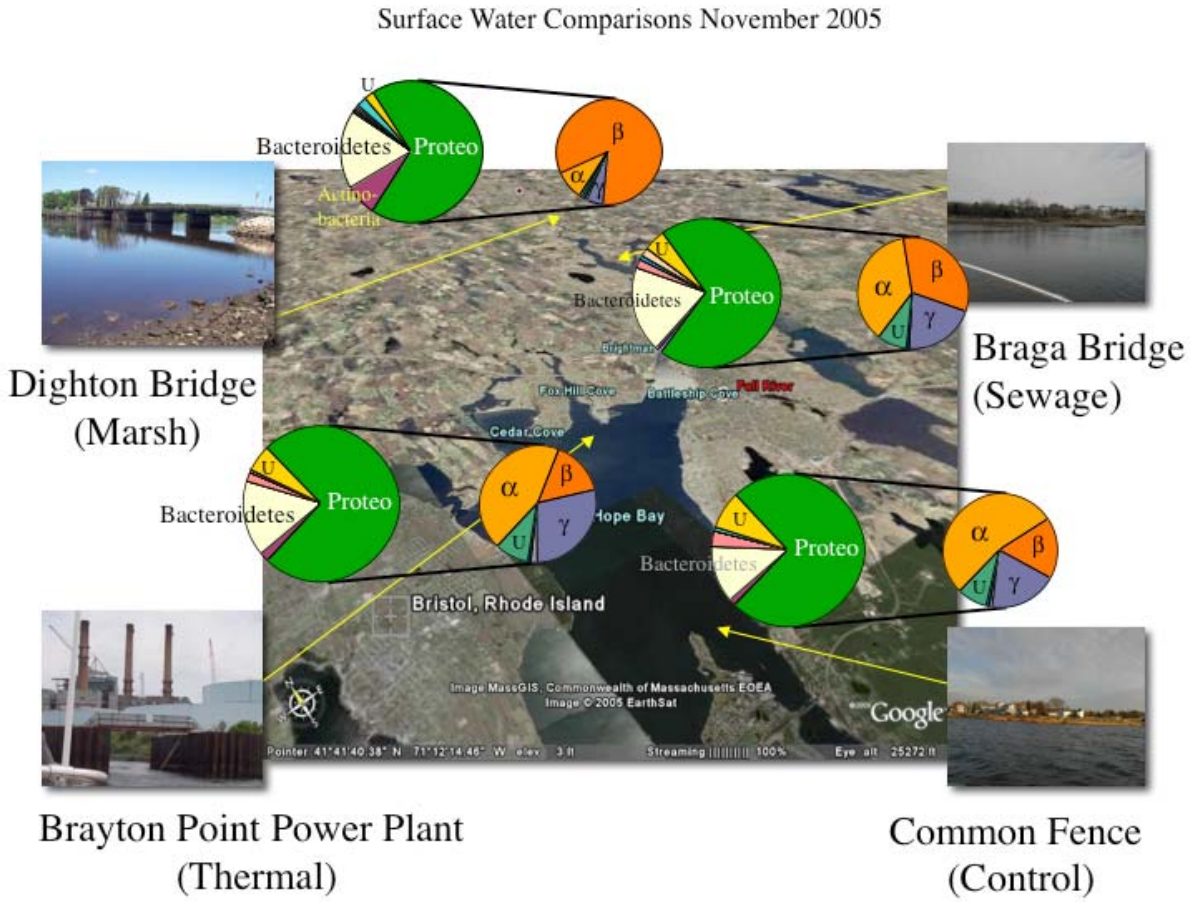
Microbial Diversity Surveys.

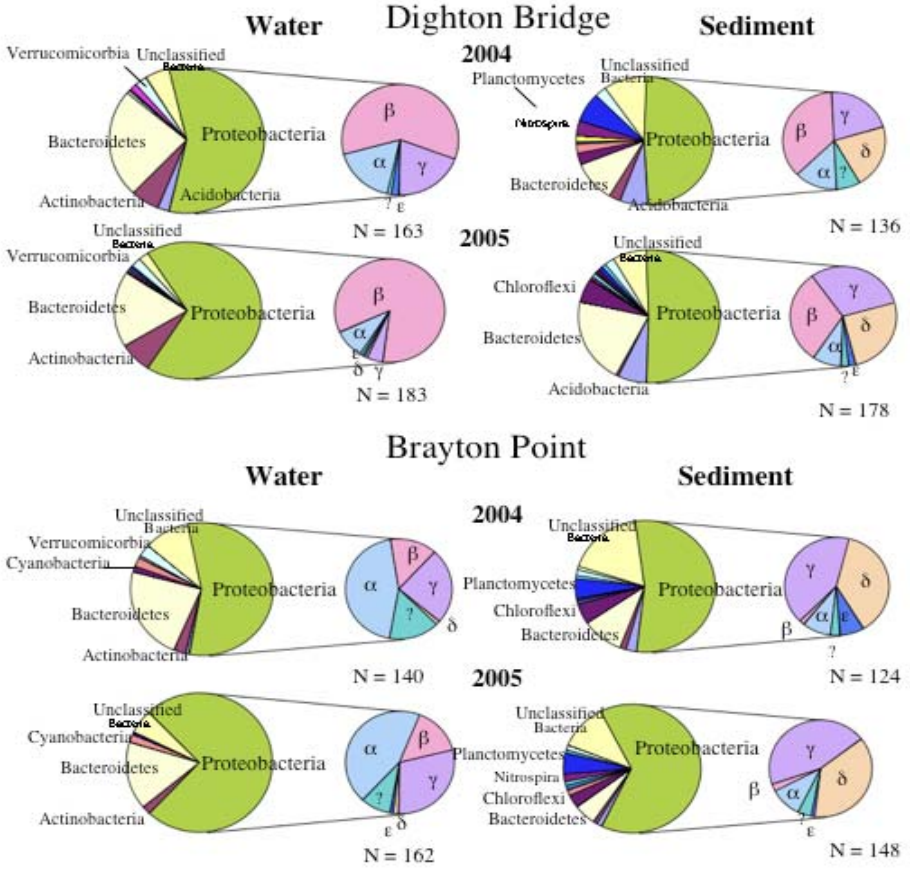
Separate Domain-specific Libraries. We have compiled the first comprehensive (eukaryal, bacterial, archaeal) data from small-subunit ribosomal RNA (SSU rRNA) gene clone libraries for water and sediment samples collected near the thermal plume and underlying sediments of the Brayton Point Power Plant (abbreviated BP in figures below) during November 2004. We have partial sequences of nearly 4,000 clones from 2 different sites and have further sequenced 1,000 unique clones from these to full-length. Not surprising, our findings reveal a highly diverse consortium of the three domains including relatives of sludge bacteria, polyaromatic hydrocarbon-degrading bacteria, and representatives related to the genera *Staphylococcus, Streptococcus, and Clostridium*. Comparisons between overall diversity in water samples for two replicate samples collected at the Brayton Point Power Plant site yielded very similar phylum-distributions. Our phylum-level comparisons between sediment samples, however, revealed significant differences in the microbial assemblages recovered. Preliminary phylogenetic placement of these environmental sequences revealed that the largest number of potential human pathogens were detected in sediment samples and not in the water-column samples. The two subsequent diagrams contrast water versus sediment replicates, asterisks indicate significant differences between groups based on the Ribosomal Database Project II Library Compare Program.



Serial Analysis of Ribosomal Sequences Tag (SARST) analyses: Our clone library constructs provide us with a reference library of full-length sequences for our Brayton Point Power Plant site. This serves as a baseline dataset with which to compare sequences obtained from higher-throughput methods that target short variable regions of small-subunit rRNA genes. Given that our initial phylogenetic analyses identified the largest number of potential human pathogens in sediment samples, we sought to further explore the microbial diversity in these samples using methods that would allow us to sample more deeply. One such method that we have begun applying to our sediment samples is called SARST (Serial Analysis of Ribosomal Sequence Tags). We currently have a total of 2,365 tags from the V6 region of SSU rRNA from one of our Brayton Point sediment samples. 513 of these possessed 100% matches to known organisms in GenBank. Of these we further identified potential pathogens belonging to the following bacterial genera that possess pathogenic representatives: *Shigella, Shewanella, Pseudomonas and Francisella*. The occurrence of *Francisella* in estuarine environments is not well-documented. As a result of this finding, we will begin to screen our samples for this pathogen in both environmental samples and amoeba cultures.

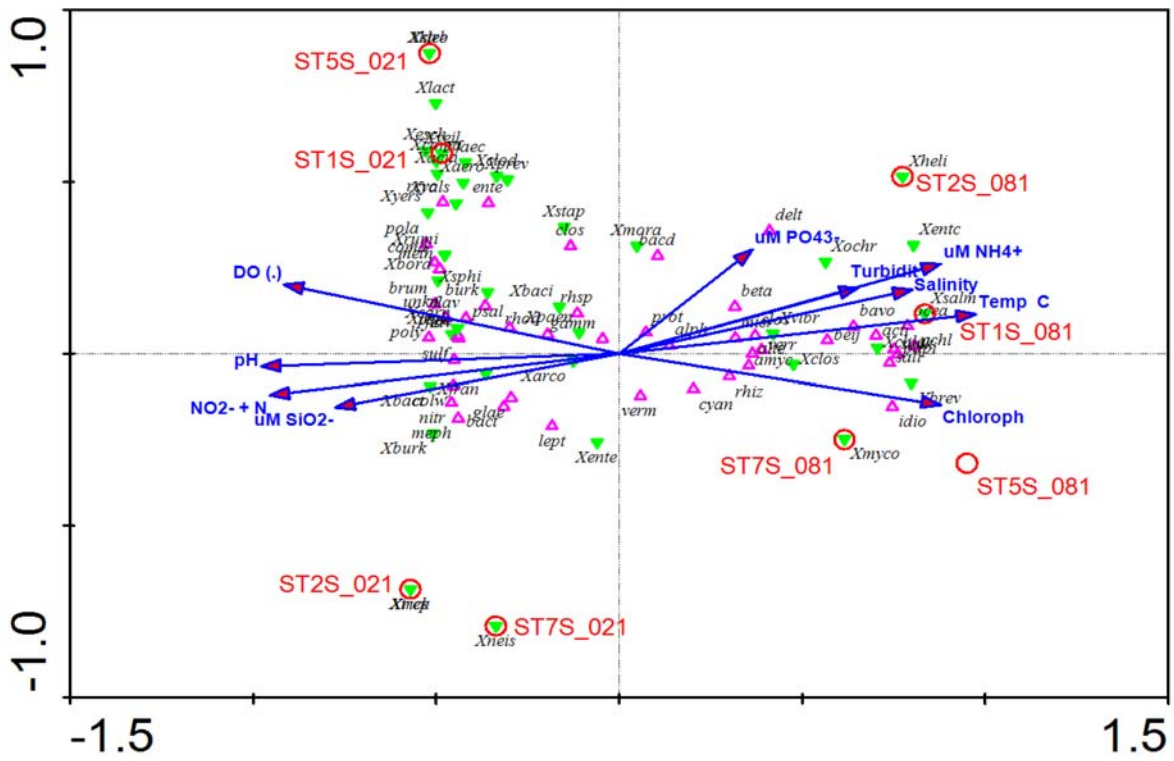
Comparison of 3-domain targeted Clone Libraries: To explore the seasonality and distribution of microbial populations and their corresponding potentially pathogenic members, we constructed clone libraries using primers that target all three domains of life simultaneously and capture the region in between the V4 through V8 hypervariable region of SSU rRNA genes. We constructed libraries from both water and sediment samples from November 2004 and November 2005 sampling seasons for the Brayton Point, Braga Bridge, Dighton Bridge and Common Fence locations. For each library we selected 192 clones for DNA sequencing in the forward and reverse orientations. Clone library composition was dominated by bacteria although some sediment samples yielded archaeal hits. The data have passed through initial quality control screening but still need to be screened for possible chimeric sequences. The pie charts below show preliminary data at the phylum level that were generated using the Ribosomal Database Project II Classifier program that assigns taxonomy to rRNA gene sequences as an example of the kind of data that are forthcoming from our analyses.





Seasonal Extreme Transect Study applying 454 Tag Sequencing:

Our final and most powerful approach employs a method developed by Mitchell Sogin and colleagues as part of the WHCOHH Genomics facility. This approach targets the same hypervariable region as the SARST-V6 method but circumvents the cloning step by proceeding directly to tag sequencing via pyrosequencing chemistry and 454 sequencing technology. As in SARST-V6, microbial tag taxonomic assignments are based on similarity to a reference SSU rRNA V6 library. Both SARST-V6 and 454 V6 tag sequencing allow for the recovery of both kinds of microbes present in a sample and their relative abundances. Our first run of the 454 machine has resulted in the recovery of 113,893 tags and included 4 summer and 4 winter surface water samples along a transect away from the power plant. The Canonical Correspondence Analysis triplot (CANOCO 4.5) below shows the relationship between Operational Taxonomic Units (OTUs = bacterial "species"), winter (STXS_021) and summer stations (STX_081) and environmental parameters. The plot includes the overall top 50 most abundant OTUs and those of any potential pathogens (indicated by an inverted green triangle and containing an "X" prefix). This analysis is preliminary but is compelling in that it reveals an abundance of potential pathogens falling in the area near Station 1 (the same as Bp 1), winter (ST1S_021). A second 454 run has just been completed with water column samples from the surface and at depth, as well as sediment samples. This second run was conducted with redesigned primers to target a broader range of bacteria. Analyses of these data are planned for the next several months.



Detection of specific pathogens. We continue in the process of setting up amplification arrays with our extracted samples for use with organism specific primers. These include primers for *Giardia, Cryptosporidium, Acanthamoeba, Naegleria* and *Legionella*. The *Legionella* primers were used in nested amplifications of our seawater, sediment and ameba culture samples. Sediment and ameba cultures from Mt. Hope Bay showed positive amplification results, and there appears to be a seasonal pattern to the prevalence of *Legionella*-like species in amebas. All of the water and sediment samples from the Great Salt Lake were positive using the *Legionella* amplification method. Both MHB and GSL samples are not often positive for *L. pneumophila*, but it has been detected. *Giardia* is also being detected in sediment and water samples of MHB, primarily in the summer months.

Legionella and amoeba sequences. Samples that were positive for the presence of legionellae were amplified, cloned and sequenced. In this work we have now newly described both the significant diversity of legionellae in the marine environment and have documented the seasonal pattern of their abundance. During the past year we have focussed on analyzing the *Legionella* sequence types present within amoeba cultures positive for the genus. Overall, 41% of MHB amoeba cultures were positive for the *Legionella* genus, while 57% of the GSL were positive. Of the positive cultures, 10% of the MHB amoeba cultures and 20% of the GSL amoeba cultures

were positive for *L. pneumophila*. It is of note that all of the amoebae that were positive for *L. pneumophila* were growing on either marine or brackish water media. This indicates that marine amoebae are very capable of supporting the growth of *Legionella pneumophila* in the marine environment. When we examined sequence types present in the non-pneumophila positive amoebae, we found that 31% were related to legionella-like amoebal pathogens, while 16% were related to other *Legionella* species that are also human pathogens. Seventy percent of the sequence types recovered from amoebae were also recovered directly from the environment, suggesting that they comprise a significant proportion of the *Legionella* population. We also investigated what types of amoebae were able to harbor *Legionella* species. Based upon small subunit ribosomal sequences, *Acanthamoeba* and *Hartmannella* were identified, along with several other genera of amoebae that have not previously been identified harboring these bacteria. This suggests that most species of amoebae have the potential to harbor bacterial pathogens in the natural environment (including the marine environment).

Doheny Beach/Avalon Beach Epidemiology Study. We have become involved in the Southern California Coastal Water Research Project (SCCWRP) beach epidemiology study. While the primary objective of the study is to correlate new indicator studies with human disease incidence, we requested to be included in the project to examine the incidence of *L. pneumophila* with respect to respiratory symptoms. We have also acquired beach sediment samples for analysis of the presence of human pathogens, including *Brucella* and *Campylobacter*. All of the water samples have been processed for Legionella detection, and those results have been submitted to SCCWRP for inclusion in the epi analysis. The sediment work was started in the fall by Elizabeth Halliday, a graduate student supported by the grant. She has developed quantitative PCR methodology that allows the identification of percent recovery and PCR inhibition so that more accurate values can be obtained, and has begun analyzing the beach sand samples from Avalon. We will also participate in the Miami COHH group beach epi study currently underway.

Significance. We have completed over a year's worth of sampling within the Mt. Hope Bay system and have been processing these samples for pathogen presence and microbial community structure. Although the detection of legionellae in marine environments was documented, the extensive presence of Legionella-like sequences in Mt Hope Bay was unexpected. This is also true for the Great Salt Lake samples. These organisms appear to be fairly prevalent in amoebae recovered from the sediments of these environments, and the number of sites that they were recovered from increased as the year progressed. This suggests that legionellae can persist and spread within the marine environment.

Our clone library analyses are still in progress, so we cannot make any conclusions at this time, however, the baseline data recovered from our study will provide a valuable comparative dataset for other studies of thermally and sewage impacted environments. Our microbial diversity data suggest that there are a diversity of potential human pathogens that can be detected using our employed approaches. Whether or not these potential pathogens can further become associated with free-living amoebae is a question we still need to explore. Of particular interest are

sequence tags recovered from both SARST and 454 tag sequencing that have a 100% match to sequences in the genus *Francisella*. *Francisella* is an obligate endosymbiont and *F. tularensis* a known human pathogen and causative agent of Tularemia. Tularemia is a localized problem on the nearby island of Martha's Vineyard, so ascertaining whether the human disease-causing strain is harbored in association with estuarine environments will be an important goal for the project over the next year.

Plans:
For the next 12 months we will be analyzing samples from the Doheny/Avalon study and summarizing the microbial diversity data from our combined approaches. *Giardia, Cryptosporidium, Naegleria and Acanthamoeba* amplifications will be completed for water, sediment and guano samples. We will also explore the relationship between our microbial diversity data and corresponding environmental data, and will begin working with the MHB FvCOM model to examine the impact of hydrodynamics on microbial populations and distribution. Manuscripts will be prepared that summarize the results of our Mt. Hope Bay study.

Training:
Support of a graduate student (Halliday), training of a postdoc (Del Castillo), mentoring of undergraduate students through REU programs, informal epidemiology training, and participation in SCCWRP California beach study and Miami beach study.

Gast has constructed and taught a graduate level course on Oceans and Human Health. Lectures are available to other COHH researchers.

Gast was a speaker at the WHOI HARP undergraduate career symposium "Global Environmental Challenges in Oceanography: Climate Change and Oceans and Human Health" June 27, 2007

Outreach and Impact:
Gast is a new Subject Editor (zoonotic diseases, protists) for Diseases of Aquatic Organisms, and is a co-editor for a special issue of Diseases of Aquatic Organisms on Marine Vertebrate Zoonoses, and will serve as a co-editor for a special issue on Marine Vertebrate Zoonoses.

Gast participated in a Massachusetts shellfish officers meeting with Stegeman during the summer to discuss potential interactions and impacts of COHH projects.

Products:
Websites Created:
1. COHH website (www.whoi.edu/science/cohh/whcohh/)
2. Gast homepage (www.whoi.edu/people/rgast/index.html)
3. Zoonotic disease website (www.whoi.edu/sites/zoonoses)

Research and Education Activities, Findings, Training and Development, and Outreach

Polz – Project 4

Research and Education Activities:
Our research pursues two overarching goals. The first is to explore environmental forcings on multiple spatial and temporal scales leading to abundance of pathogenic vibrios (Aim 1 and 2); the second is to understand the nature of genetic differences and gene flow between co-existing pathogenic and non-pathogenic variants of vibrios (Aim 3 and 4). All our aims are pursued within the Plum Island estuary and Parker River where we determine whether changes in estuarine physics, water chemistry and biology correlate with transitions in diversity and abundance of different *Vibrio* species and to determine estuarine conditions that might trigger the emergence of pathogenic strains of *Vibrio*. Overall, we are nearing completion of the tasks within AIM 1-3, and report progress on AIM 4 since last year's report.

Findings:

AIM 1 and 2. To characterize and model dynamics and reservoirs of *V. vulnificus* and *V. parahaemolyticus* populations over seasonal cycles, and to test the link between estuarine physics, nutrient and particle abundance and growth patterns of *Vibrio* species over tidal cycles
We summarize here progress on both aims. The sampling for both aims is complete and was spelled out in the last report. The analysis of population structure by culture-independent assays is currently in progress and will be finished by early 2008.

To determine ecologically coherent groups among co-existing vibrios in the water column, we examined the temporal and spatial distribution of *Vibrionaceae* genotypes. We sampled the water column to differentiate the free-living and attached (to both particles and zooplankton) compartments of the planktonic community under different macroecological conditions (spring and fall). Particle-attached and free-living cells were separated (in 4 replicates each) into a total of four consecutive size

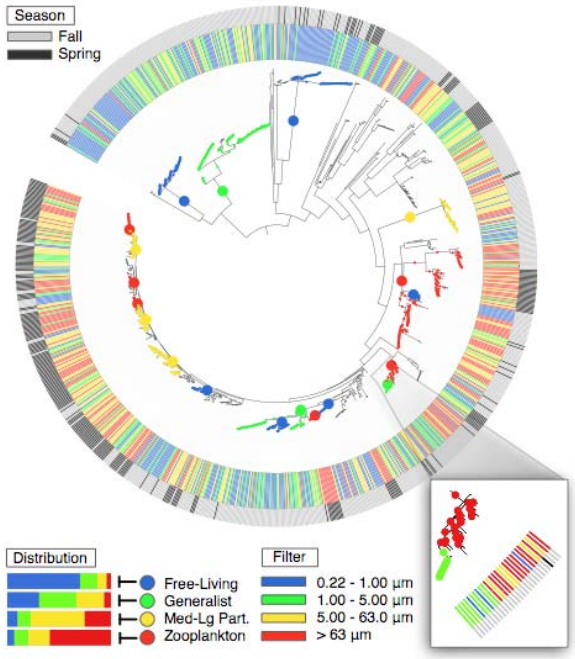


Fig. 1 Mapping genotypes across environments. Shown is a phylogeny of *Vibrio* isolates based on the *hsp60* gene. The inner ring depicts by color the size fraction from which each strain was isolated (see legend in figure), and the outer ring shows the season. The *AdaptML* software package was used to infer four commonly occurring distributions for phylogenetic groups (see legend in figure), and ecological populations demonstrating one of these characteristic distributions are indicated by colored dots (large dots indicate the ancestor of the group). The inset figure shows two closely reated groups (nested clades) that have distinct ecological distributions.

fractions, which are enriched in zooplankton (≥63 μm), large (63-5 μm) and small (5-1 μm) particles, and free-living cells (1-0.22 μm). The 5-1 μm size fraction is somewhat ambiguous, likely containing cells attached to small particles, as well as large or dividing cells; however, it provides a firm buffer between obviously particle-attached (>5 μm) and free-living (<1μm) cells.

Roughly 1,000 isolates were characterized by sequencing of a protein-coding gene (*hsp60*). To confirm relationships, between 1 and 3 additional gene fragments (mdh, adk and pgi) were sequenced for all *V. splendidus*, the dominant taxon during warm water conditions (Thompson et al., 2005b), and several other microdiverse groups. These data allow conservative estimation of ecological differentiation because inadvertent mixing of strains among microhabitats and homologous recombination among strains homogenize rather than create associations.

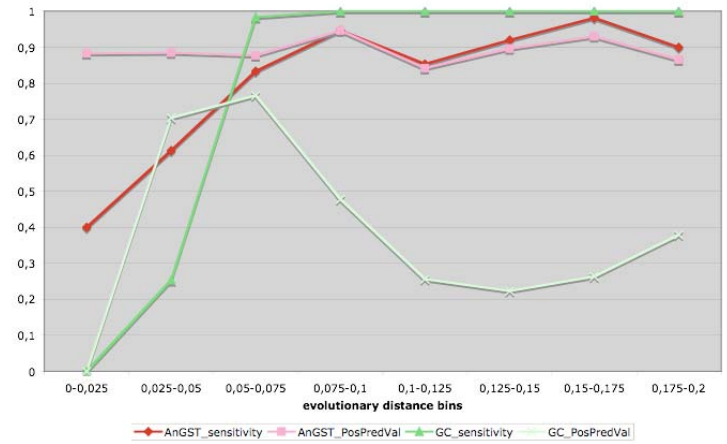


Fig 2. *AnGST-HR* performance on simulated vibrio data compared with pairwise approach. Simulated gene sequences were simulated according to an actual vibrio phylogeny inferred from 100 MLSA samples, and recombination events between leafs were randomly inserted into the tree among taxa at different evolutionary distances. *GENECONV* (GC) and *AnGST-HR* were evaluated based on their sensitivity (1- fraction of false negatives) and positive predictive value (PPV = fraction of true predictions that are correct). Both methods show a 'blind spot' for detecting recombination between very similar sequences, but *AnGST-HR* shows a high PPV across the entire range of distances, while *GENECONV* tends to make false positive predictions when highly divergent sequences are considered.

In collaboration with Eric Alm of the MIT Systems Microbiology group, we developed a software package (*AdaptML*) to identify associations between phylogenetically-defined groups and microhabitats that combines a 'model-free' empirical approach and an explicit evolutionary model of habitat association. Briefly our empirical framework establishes whether or not each clade in our strain phylogeny shows a non-random distribution across environments, while the explicit evolutionary model estimates population boundaries, rates of habitat 'switching', and the distribution of habitats across a set of measured environmental parameters.

Our data reveal important aspects of vibrio ecology in the water column. First, phylogenetic groups with distinct ecological preferences are clearly identifiable among all clades with sufficient representation in the dataset (Fig. 1). There is strong temporal structure evident where populations occurring in the spring and fall are frequently distinct at deep phylogenetic levels (Fig. 1). However, spatial partitioning is also evident for many groups, including *V. splendidus*, which contain many shallow clades with displaying distinct distributions (*V. splendidus* occupies the lower section in the tree spanning ~4-10 o'clock, Fig. 1). We were even able to identify nested clades where a single nucleotide change in several housekeeping genes is correlated to a switch in environmental distribution and thus suggests that this group is diversifying (one of these is highlighted in the inset to Fig. 1).

Second, although the currently crude spatial sampling (by size fraction) does not allow assignment of specific habitats (e.g. specific types of organic particles, zooplankton body sections), the analysis indicates differentiation into populations with broad ('generalist') and more narrow ('specialist') distributions, respectively. For example, *V. ordalii* was identified as specialized to free-living lifestyle, while the potentially pathogenic *V. parahaemolyticus* appears to be a generalist inhabiting particles, zooplankton and the free-living fraction of the water.

AIM 3 and 4. To determine the 'rules' of recombination, gene transfer and clonality in co-occurring *V. vulnificus* and *V. parahaemolyticus* and related vibrios, and to assess the diversity, mobility and evolutionary history of genes implicated in pathogenicity
Completion of this aim required the development of a new tool for recombination determination (Fig. 2). AnGST-HR is a modification of the AnGST software, which can identify gene dynamics within a comparative genomics framework. Fig. 2 shows that the method already performs better than other current methods in simulated datasets. The software is now being adapted to multi-locus sequence data sets to allow determination of recombination rates against sequence distance, which represents a critical parameter in judging the genetic isolation of different populations of microbes.

We have leveraged the work carried out under the auspices of the COHH to obtain funding from the Moore Foundation for sequencing of several of our genomes. We already have one *V. splendidus* and *V. alginolyticus* genome in hand. Technical developments now enable us to sequence 21 additional genomes by a combination of 454 and Solexa sequencing. This will provide an unprecedented dataset for both aims. Moreover, we are in the process of determining correlation of O-antigen diversity to population structure in the populations isolated from diverse habitats (water column, animal associated, etc.).

Significance:
The last few years have seen rising concern about the emergence of new variants of pathogens and spread of existing pathogens due to local or global environmental change. This has focused attention on the ecological context of pathogens in both the human body and the environment. Advances in population biology, aided by genomics, have demonstrated that many closely related (genomic) variants of microbial species exist in the environment. Furthermore, it has been shown that virulent bacteria frequently emerge from non-virulent strains via lateral gene transfer and it has been suggested that bacterial genomes are capable of extensive recombination. This raises the fundamental question to what extent observed genomic variants represent ecological and evolutionary units that can be seen as the bacterial equivalent to the eukaryotic sexual species. Do observed genomic variants occupy different environmental niches or do they represent a common gene pool capable of rapid 'assembly' of new variants in response to environmental challenges? These questions are crucial for interpretation of pathogen biology, risk assessment of emergence, and insights into how representative currently extensively studied strains (e.g., of *E. coli* or of pathogenic species) are for the 'species' they represent.

Products:
We are currently collaborating with colleagues in Austria and Germany to develop widely applicable microarray software based on our paper Marcelino et al. (2006).

Further, we have developed the AdaptML software, which can be used to detect ecologically cohesive bacterial populations among strains sampled from varied environmental samples.

Research and Education Activities, Findings, Training and Development, and Outreach

Sogin – Genomics Core

Research and Education Activities:

a. Facility operations and administration:

The genomics core provides DNA sequencing services and computational support to investigators in the Woods Hole Center for Oceans and Human Health (WH-COHH). Many of the Center's projects take advantage of our strengths in molecular microbial ecology and the bioinformatics. Over the past calendar year, the genomics core facility has provided WH-COHH projects with just over 31,000 reads corresponding to ~23 million base pairs. **Figure 1** Shows the usage pattern for the genomics core in 2007 and **Figure 2** describes the use of the Genome Core Facility by different WH-COHH projects.

Figure 1. WH-COHH sequencing usage pattern.

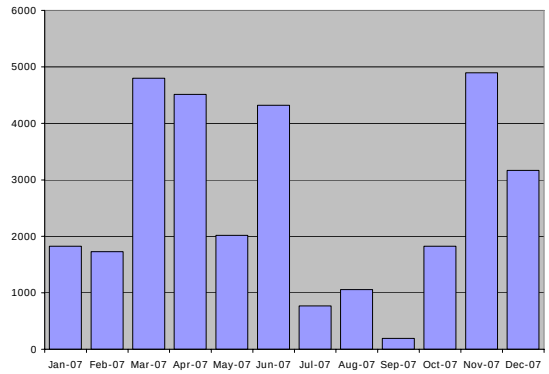
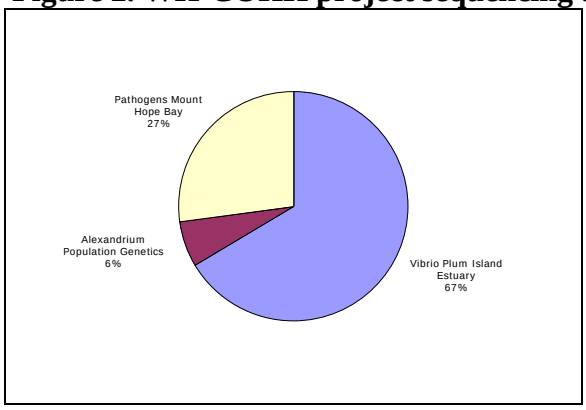


Figure 2. WH-COHH project sequencing activity.



Sequencing costs have changed little over the past year. Each template costs 30 cents using a protocol run on our BiomekFX platform. This is still a relatively inexpensive template production cost. In July, our Roche Genome Systems20 (GS20) pyrosequencing system (funded by a award titled ***Microbial population structure of the world's oceans*** from the W.M. Keck Foundation to the Marine Biological Laboratory at Woods Hole) was upgraded to a GS-FLX, which averages ~400,000 reads of ~238 nucleotides/read in each run. We anticipate that this

system will continue to be a key element of future COHH investigations. With this system, we have generated V6-tag datasets (as described below) for several of the MHB transect samples, providing very deep sampling to complement the traditional clone libraries outlined in the original proposal. We have also sequenced the genomes of six *Vibrio* phages in collaboration with Drs. Jonathan King and Peter Weigele (WH-COHH pilot project) using the GS-FLX.

b. Scientific and technology developments in the genome core.

Using the GS20 (and later, the GSFLX) sequencer the core facility continued development of a new, massively parallel DNA sequencing approach to characterize microbial populations with unprecedented level of sensitivity for detecting low abundance organisms. The International Census of Marine Microbes (ICOMM) provided the test samples and funds for reagents and the WH-COHH provided computational support for analysis of the data. We described in our last report the use of this approach to assess the species richness and eveness in water samples from the North Atlantic Deep Water flow and from Axial seamount (Sogin et al., 2006). During 2007, we demonstrated the feasibility of using 5-nt "keys" to multiplex reactions on the same region of a plate, deconvoluting the reads into sample bins informatically. This approach has been used on hundreds of samples from seawater, freshwater, sediment, soil, sewage, and human feces. We analyzed just over 1.3 million V6 tag sequences from human stool samples (500,000 from David Relman's laboratory and 800,000 generated in our laboratory as part of a collaboration with Jeffrey Gordon). This will provide a reference data for interpreting V6 tag sequences studies from human-impacted environments that we are studying in the context of a WH-COHH pilot project. Our data processing pipeline is well established, and datasets from tag and full length studies are displayed using a new web tool, Visualization of Microbial Population Structures (VAMPS), developed with support from the Sloan Foundation (awarded to M.L. Sogin and David Mark Welch).

Findings:

The ability to accurately characterize microbial population structures through the use of tag sequencing protocols is contingent upon the sequencing accuracy of this techonology. The published accuracy of 454 sequencing technology is only 96%. In genome projects, highly-redundant consensus assemblies can compensate for sequencing errors. In contrast, studies of microbial diversity that catalogue differences between PCR amplicons of ribosomal RNA genes (rDNA) or other conserved gene families cannot take advantage of consensus assemblies to detect and minimize incorrect base calls. To explore error modalities, we generated more than 340,000 reads from a PCR amplicon library that was prepared from a collection of 43 reference templates of known sequence. Each reference template contains a distinct rDNA including the V6 hypervariable region from a collection of 43 divergent bacteria. Differences between pyrosequences and their cognate reference sequences identified signatures of low quality data. We determined that 454 technology has a 99.5% accuracy rate in unassembled sequences, and we identified several factors that can be used to remove a small percentage of low-quality reads, improving the accuracy to 99.75% or better. Table 1 lists the source of those errors.

Table I.

Data Selection	Percent of Reads	Error Rate
All reads	100.0%	0.49%
Reads with no Ns	94.4%	0.24%
Reads with more than one N	5.7%	4.7%
Reads with length >=81 and <=108	98.8%	0.33%
Reads with length <81 or >108	1.2%	18.9%
Reads with no Ns and length >=81and <=108	93.3%	0.20%
Reads with no proximal errors	97.0%	0.45%
Reads with fewer than 3 proximal errors	>99.99%	0.48%
Reads with more than 3 proximal errors	<0.01%	12.2%
Reads with no Ns and length >=81and <=108 and no proximal errors	90.6%	0.16%

The simple removal of sequences that contain any unidentified ambiguities within the sequence of the target amplicon and elimination of sequences shorter than 50 bp or 20% longer than the target amplicon reduces the error rate to less than 0.25%. This is roughly ½ the error rate associated with capillary methods. This work appeared in Genome Biology (Huse, S.M., J.A. Huber , H.G. Morrison, M.L. Sogin and D. Mark Welch. Accuracy and quality of massively parallel DNA pyrosequencing. Genome Biology, 8: R143).

We have also demonstrated that the small amplicon size (~100 bp) more accurately captures the microbial diversity in a sample, most likely because small amplicons contain fewer regions of secondary structure that cause the polymerase to dissociate from the template. Results are shown in Table II. A manuscript describing this result is in progress (Huber, J.A., H.G. Morrison, S.M. Huse, P.R. Neal, M.L. Sogin, and D.B. Mark Welch. Effect of PCR amplicon size on assessments of microbial diversity and community structure, unpublished).

Table II.

Library	No. Clones	Forward Primer	Reverse Primer	% Exact Matches[a]	% Unique Sequences[b]
FS312_100bp	761	967F	1046R	44%	51%
FS312_400bp	860	967F	1391R	74%	22%
FS312_1000bp	381	337F	1391R	72%	12%
FS396_100bp	685	967F	1046R	11%	40%
FS396_400bp	866	967F	1391R	48%	22%
FS396_1000bp	677	337F	1391R	38%	19%

[a] Percent of V6 sequences within library that are exact matches to an existing sequence in the reference database

[b] Perecent of V6 sequences within library not detected by the other two libraries from the same sample

Training:

Through the genome core activities we continue to train a graduate student, Yuko Hasegawa in the use of advanced genomic techniques in the analysis of microbial communities associated with anthropogenic activities.

Outreach and Impact:

Drs. Sogin and Morrison have given lectures as part of a continuing education seminar series for Falmouth physicians and science teachers titled “Medical Genetics”. Analysis of microbial communities using the novel tag approach was included in the material presented.

Proposals Funded:
VAMPS – Visual analysis of Microbial Population Structures
Alfred P. Sloan Foundation. M.L. Sogin PI $300,000 July 1 2007- Jun 30, 2009.

Research and Education Activities, Findings, Training and Development, Outreach, and Products

Hahn – Pilot Project Program

Activities:

The specific objectives of the Pilot Project Program are:

A. To assess the feasibility of new areas of study, especially those that are not currently represented in our Research Project base but would contribute to the overall goals of the center.

B. To support the collection of preliminary data that can be used to generate full proposals to NSF, NIH, or other agencies or organizations.

C. To recruit scientists not currently involved in research on oceans and human health, but who may have expertise in one of these areas, to become participating members of the center and interact with other center investigators.

D. To stimulate collaborative and interdisciplinary research within the center.

E. To foster the application of new technologies and experimental approaches to questions concerning the impact of oceanic processes on public health.

F. To ensure the ability of the center to respond rapidly to new scientific information and emerging challenges in this field.

G. To contribute to the training of future researchers in the field of Oceans and Human Health through enhancement of graduate and postdoctoral training opportunities.

Findings:

Completed pilot projects (2004-2007) (list of publications and grants at end)

- *Characterization of a cyanobacterial anti-algal compound* (Eric Webb and Chris Reddy, WHOI).
- *Cnidarian toxins against voltage-gated Ca2+ channels* (Robert Greenberg, MBL).
- *Marine phage as vectors of gene transfer between marine bacteria and bacterial pathogens* (Peter Weigele and Jonathan King, MIT).
- *Transcriptome profiling in the harmful alga Aureococcus anophagefferens.* (Sonya Dyhrman, WHOI).
- *Beach Pathogens* (Steve Elgar, Britt Raubenheimer, & Rebecca Gast, WHOI)
- *Names-based cyberinformatics tools for rapid response communications and outreach during event management – a pilot based on harmful algal blooms in NE US coastal waters* (David J. Patterson, MBL, and Don Anderson, WHOI).

<u>Ongoing pilot projects (2006-2008)</u>

• *Anthropogenic impacts and profiling fecal microbial populations at a salt marsh* (Mitch Sogin, MBL)

Fecal coliforms are indicator organisms that warn of possible fecal contamination and its potential impact on human health. Coliform surveillance activities generally rely upon cultivation assays or molecular Microbial Source Tracking (MST) technology to assay for the presence of a particular indicator organism in environmental samples. This project will employ a

massively parallel DNA tag sequencing strategy to profile entire microbial communities in the Little Sippewissett Salt Marsh and the barrier Woodneck Beach. Possible coliform sources include human waste from failed septic systems, bird populations or other animals. The objective is to use the tag sequence data to locate specific sources of fecal contamination and identify suites of genes that could serve as multi-species indicators of human pollution. The experimental strategy takes advantage of rapidly evolving hypervariable regions in ribosomal RNAs and our ability to generate many thousands of short DNA tag sequences using 454 Life Science's pyrosequencing sequencing technology.

• *Transcriptional Markers of Life Cycle Transitions in Harmful Algal Blooms* (Don Anderson, WHOI)
Bloom dynamics of the red tide dinoflagellate, *Alexandrium fundyense* are driven in large part by transitions in its life cycle. While these stages are well documented, the biological and oceanographic forces that trigger transitions between the stages are not. A major obstacle to determining the conditions that trigger these transitions is our inability to rapidly identify sexual stage cells. Here, a novel transcriptome experiment will be used to discover molecules that are uniquely expressed by conjugating gamete cells and by germinating cysts. This experiment will utilize sequencing-by-synthesis technology that is newly available through the Genomics Core.

• *The Economic Effects of Harmful Algal Blooms: A pilot project to estimate the costs of human respiratory ailments associated with aerosolized brevetoxins* (Porter Hoagland and Di Jin, WHOI; Lora Fleming, Miami)
This study will estimate the costs-of-illness associated with human respiratory ailments that arise as the consequence of the aerosolization and coastal to inland transport of brevetoxins from blooms of the marine dinoflagellate, *Karenia brevis*, in the Gulf of Mexico. The research will develop models to link the occurrence of HAB events in the coastal-ocean with exposures to aerosolized brevetoxins. The researchers will compile datasets and develop models of illness rates that would permit historical estimates of these kinds of impacts and the simulation of future potential impacts.

• *The Economics of Human Health Risks from Pathogens and Toxins in the Marine Environment* (Hauke Kite-Powell and Porter Hoagland, WHOI)
These researchers will produce an "order of magnitude" estimate of the annual human health cost imposed on residents of the United States by exposure to pathogens and toxins from the marine environment. The estimate will be derived from a review and synthesis of information in the existing literature on (1) the spatial and temporal prevalence of marine pathogens, (2) the pathways by which they affect humans and the potentially exposed populations, (3) the human health effects of exposure, and (4) the economic cost of resulting medical conditions. The estimate will inform future research on pathogens in the marine environment, allowing scholars and public officials to target pathogens and settings where improved scientific understanding is most likely to produce significant economic benefits, and setting the stage for focused economic analyses. (co-funded with the WHOI Marine Policy Center)

Fourth call for pilot project proposals.
The fourth call was issued in September of 2007 by email to all faculty and research staff at WHOI, MBL, and MIT. On November 1, 2007, nine proposals were received and subsequently

were reviewed and scored by members of the Internal Advisory Committee and external reviewers, including several members of the other COHH Centers. The proposals requested approximately $308,000 in total costs. All applicants were provided with written reviews. Three projects were selected for funding:

Hydrodynamics and Transport Pathways for Fecal Microbial Populations in a Salt Marsh and Barrier Beach System (David Ralston, WHOI, Dept of Applied Ocean Physics and Engineering)

We propose to develop and implement a numerical model of flow and transport in a salt marsh and barrier beach system on Cape Cod, MA. The project will focus on Little Sippewissett Marsh and Wood Neck Beach, a system that during summer months frequently experiences high fecal coliform concentrations that indicate potential for significant impacts on human health and lead to restrictions on recreational use. Potential sources of fecal contamination include failed septic systems, birds, or other animals in the marsh. Recent investigations have found that fecal coliform counts are highly variable temporally (tidally, with precipitation events, and seasonally) and spatially across the marsh and beach. A COHH pilot project currently underway has begun to quantify the spatially heterogeneity of microbial communities in the system, intending to distinguish among different sources of fecal contamination and to develop a broader suite of indicator organisms (Sogin, 2006). The project proposed here will work in conjunction with that microbial community mapping project to quantify transport pathways, residence times, and exchange rates in the marsh. Combining transport mechanisms with the spatially heterogeneous source terms will permit calculation of potential exposure and risk to human health associated with the disparate sources of contamination. The model results will also provide guidance for public health officials to redesign of monitoring efforts to sample at times and locations of maximum potential exposure to elevated coliform concentrations, thereby minimizing human health risks. The tasks in the project include acquiring bathymetric data from existing sources and from new field surveys, constructing a numerical grid from the bathymetry, establishing and collecting data for model boundary conditions, and initial calibration and testing of the model based on available observations in the marsh. The model is intended to serve as a basis for future interdisciplinary studies at the study site – model results will aid in design of field observations, and more extensive field studies will aid in refinement of the model. The salt marsh and barrier beach of the study site represent a common coastal environment, and the impacts on human health of fecal contamination in coastal settings are of concern regionally and nationally. The pilot project will provide a step toward building an integrated program in Little Sippewissett to study linkages among physical, chemical, and biological processes in salt marshes and very shallow estuaries, coastal environments where frequent human interactions make potential health impacts particularly significant.

Using signature tagged mutagenesis (STM) to investigate how pandemic Vibrio parahaemolyticus persists in the bacterioplankton and associates with epithelia in the marine environment (Janelle Thompson, MIT, Dept of Civil and Environmental Engineering)
The overarching goal of this proposal is to investigate the evolution and emergence of pathogens from the marine environment by studying the interactions of the model pathogen pandemic *Vibrio parahaemolyticus* with the marine protist *Cafeteria roenbergensis* as a model for marine grazing pressure, and the starlet sea anemone *Nematostella vectensis* as an evolutionarily basal

host epithelium.

The specific objectives of this proposal will be to:
1. Develop a signature-tagged mutagenesis (STM) system for *Vibrio parahaemolyticus* using mariner-transposition and a novel PCR-tRFLP-based strategy for signature tracking.
2. Generate a library of tagged transposon mutants of *Vibrio parahaemolyticus.*
3. Characterize *Nematostella vectensis* as a model for microbe-epithelial interactions.
4. Screen a subset of the tagged *Vibrio parahaemolyticus* mutants for reduced fitness under grazing by *Cafeteria roenbergensis* and during association with *Nematostella vectensis* to demonstrate proof-of concept for use of STM to investigate the genetic mechanisms of environmental persistence and marine epithelial interactions and to generate preliminary results for future proposals.

BMAA, a cyanobacterial neurotoxin, in marine food webs: a pilot project
(Carl Lamborg, Mak Saito, Paul Drevnick; WHOI, Dept. of Marine Chemistry and Geochemistry)
ß-methylamino-L-alanine (BMAA) is a neurotoxic amino acid produced by cyanobacteria. High concentrations of BMAA in human brain tissue have been linked to neurodegenerative diseases (ALS, Alzheimer's, Parkinson's) in Guam and Canada. The source of BMAA in Guam is cyanobacteria in the roots of cycad plants and biomagnification through a unique food web. The source of BMAA in Canada is unknown. A recent study, however, reported that many marine cyanobacteria also produce BMAA. Cyanobacteria are ubiquitous in the ocean and especially abundant in coastal areas that have experienced harmful algal blooms, representing a potentially significant source of BMAA to marine food webs. Fish or shellfish that eat cyanobacteria or otherwise accumulate BMAA may thus pose a health risk to human consumers of seafood. We propose to address the most fundamental question concerning the distribution of BMAA in the temperate coastal ocean: Are BMAA concentrations in seafood high enough to be of concern for human health? We will examine fish and shellfish of commercial, recreational, and subsistence importance for BMAA concentrations. If we find BMAA concentrations that pose a human health risk, (i) this could form the basis of a human health risk assessment for BMAA and (ii) we will have preliminary data to generate a full proposal for further study.

c. Significance

The funded pilot projects address important questions regarding the interactions between the oceans and human health. Many of the projects represent new directions for the Center, expanding the scope of our research efforts into, for example, marine natural products (e.g. BMAA) and employing new technologies (e.g. signature tagged mutagenesis). Many of the funded or proposed pilot projects involve postdoctoral researchers or graduate students, and will thus contribute to the training of the next generation of researchers in this area.

d. Plans

During the next project period, we will continue the Pilot Project Program, issuing a new call for proposals early in the year. We also will follow-up the previously funded projects, to gauge their progress and success in achieving research goals and obtaining externally funded grants using preliminary data generated in the pilot projects.

e. See Publications List

f. Project-generated resources:

One objective of the pilot project program is to support the collection of preliminary data that can be used to generate full proposals to NSF, NIH, or other agencies. Proposals (funded and pending) arising from our pilot projects are listed below:

Approved for Funding:
NSF: S. Dyhrman, M. Saito - EN-GEN: Transcriptional and Proteomic Analyses of Multiple Environmental Stressors in Marine Diatoms (TP-AMES) $999,500

Florida Fish & Wildlife Conservation Commission. P. Hoagland and others. Proposal to extend the WH-COHH pilot project to the entire west coast of Florida. $84,975.

NSF: H. Kite-Powell: Supplemental COHH funding to study the economic losses associated with beach closures in the Myrtle Beach region due to contamination of marine waters ($20,000).

Pending:
EPA: S. Dyhrman - Linking Biogeochemistry to Harmful Algal Bloom Nutritional Physiology with Gene Expression Analysis: A Case Study with *Aureococcus anophagefferens* $497,821

NIH: PI Sandra McLellan and Co/I Mitchell Sogin. Microbial community profiling of sewage contamination in the Great Lakes: R21 grant submission.

Research and Education Activities, Findings, Training and Development, Outreach, and Products

Stegeman – Administrative Core

Research and Education Activities:

1. Foster the communication, planning, integration and interaction among Center members, thereby assuring that the goals developed in the Center vision are achieved.
2. Provide and maintain an administrative structure to oversee and monitor the financial aspects of the Center, including grants management.
3. Establish a structure that will support effective communication and planning with contribution from all Center Investigators.
4. Establish an External Advisory committee, which will review the Center programs and advise the Director and Deputy Director.
5. Oversee the operation of and work toward a successful Pilot Project small grant program.
6. Establish and maintain public and private web pages for the dissemination of information and for data transfer and communication.
7. Pursue interactions between the COHH and other Centers within the three Institutions, as well as with other COHH and EHS Centers, in order to expand the reach and impact of the center, and to expand the research base on issues relevant to the Center mission and themes.
8. Encourage the involvement of center investigators in community outreach and education efforts.

Findings:

1. As before, the Director and Deputy Director have continued to oversee the Center Office and the activities of each of the component projects and cores in the Center. We continue to build visibility in and impact through all three of the component institutions, the Woods Hole Oceanographic Institution (WHOI), the Marine Biological Laboratory (MBL) and the Massachusetts Institute of Technology (MIT).
2. The Center Director and staff (administrative professionals), and the respective Grants Management offices at the Woods Hole Oceanographic Institution and other institutions continue to monitor the accounts for each of the Research Projects, the Pilot Projects, and the Advanced Genomics Core.
3. Center Investigator meetings have been held monthly, on the 1st Friday of each month. In addition, the Director of the Administrator informs the Center members of activities and opportunities pertinent to Oceans and Human Health via frequent email. The meetings and notes address all aspects of the Center interactions and communications, internally as well as externally. The Director and Deputy worked to increase visibility of the Center.
4. The Administrative Core organized a day and a half retreat for the investigators in the WH-COHH in January 2008, at the nearby National Academy Study Center. The retreat focused overall progress in the research and possible new directions. There will be a similar meeting of all the Pilot Project awardees, to be held in February.
5. The Director met during the summer with the Chair of the External Advisory Committee (Dr. Michael Gallo, University of Medicine and Dentistry of New Jersey), to discuss the center's activities.

6. The Director and Deputy continued to work together with the director of the Pilot Project program (Dr. Mark Hahn), to coordinate the operation of the successful Pilot Project program. Four pilot projects were funded with 2006 funds, including one involving a joint project with the Miami Center for Oceans and Human Health. In excess of $40,000 was leveraged to augment these projects. The fourth call was issued in September of 2007 by email to all faculty and research staff at WHOI, MBL, and MIT. On November 1, 2007, nine proposals were received and subsequently were reviewed and scored by members of the Internal Advisory Committee and external reviewers, including several members of the other COHH Centers and other Ad Hoc reviewers. The proposals requested approximately $308,000 in total costs. All applicants were provided with written reviews. Three projects were selected for funding: Three were recommended for funding, two from WHOI and one from MIT. Additional funds in the program may be used to fund partially two other highly meritorious proposals, one from MIT and one from MBL. The distribution of awards to all three of the constituent institutions is a strength of the program.
7. The COHH website developed at WHOI has continued to serve the needs of the four primary OHH Centers, Woods Hole, Hawaii, Miami and Washington (http://www.whoi.edu/science/cohh/). COHH links continue to be added and an ftp site (internal) is in use.
8. Interactions with other groups continue to grow.
    a. Interactions between the WH-COHH and the other COHH. The Director and Deputy Director, with the help of the whole Woods Hole Center, hosted the third joint Center Directors' and Investigators meeting, in Woods Hole, in April, 2007. This meeting was highly successful and resulted in 5 manuscripts focused on progress in different areas of OHH, which have been completed and are under review for publication in the BiomedCentral journal *Environmental Health*, a journal with a very good impact factor. The meeting also resulted in a collation of outlines of several academic courses on Oceans and Human Health that have been developed over the past years by investigators in Woods Hole Center and in other OHH centers.
    b. We have continued to participate in monthly conference calls among the leadership of the four NSF-NIH Centers, and a second monthly call that includes the NOAA OHHI leaders, to discuss points in collaboration and interaction in all aspects of the Centers' activities.
    c. Interactions between the WH-COHH and other NIEHS Centers: We are continuing to explore opportunities for jointly sponsored enrichment activities, as an outgrowth of the Center Directors meeting in April. These interactions are being facilitated by the Supplemental funds obtained from NSF in 2007.

Outreach and Impact:

The Center has been involved in informing the wider scientific community. We also have continued to encourage the involvement of center investigators in community outreach and education efforts. There have been several major activities this year.

a. Drs. Stegeman and Gast made presentations to the Massachusetts Shellfish Officers Association quarterly meeting, describing COHH activities and discussing possible interactions.

b. Dr. Stegeman and Dr. Lora Fleming of the Miami COHH have worked to develop an exciting agenda for the new Gordon Research Conference supported meeting on Oceans and Human Health, scheduled for 2008. We have enlisted and confirmed participation of the full slate of speakers and discussions leaders and are seeking funding to support this exciting new conference.
c. Dr. Amaral-Zettler gave a lecture on Oceans and Human Health in a course for physicians and health care professionals, at Falmouth Hospital. The Course, entitled "Molecular Biology's role in Modern Medicine, was organized by Genomics Core Director Mitch Sogin.
d. The Administrative Core and the particularly the Director helped Center Investigators with the preparation of the SGER grant application submitted to the NSF.
e. The Center also sponsored several seminars by researchers from other institutions, and provided support fore a visit by Dr. Debashish Battacharya (University of Iowa) to the Center.

Training and Development:
The Center supported an undergraduate Summer Fellow in 2007, through the NSF-REU grant. The fellow was Lara Polansky, University of Miami, who worked with pilot project recipient Porter Hoagland on assessing the role of *K. brevis* blooms on emergency department respiratory diagnoses in Florida. Another student was supported through the MBL to work in Becky Gast's lab. Other students worked on COHH projects in the lab of Center Investigator Martin Polz at MIT. The Woods Hole Center also provided support to two students to attend the fall Harmful Algal Bloom conference in Woods Hole.

Significance:
As before, the Administrative Core has overseen growth of the Center and facilitated the activities of all of the units, through communication and support. The Center has had important successes in the research projects and the pilot project program. Inter-center communication is robust. Funds continue to be leveraged to support and expand the activities of the Center. Center activities include research that has had direct consequences for the public health. The Administrative Core is the focal point for all Center activities.

Plans:
As in the past, during the next project period we will continue to oversee the management of the grant and the Center, to foster intra- and inter-center communication and collaboration, to increase the activities in out-reach and enrichment, and to seek additional funding sources that might be leveraged to maintain a robust pilot project program. During the next year the WH-COHH also will continue recruiting undergraduate student researchers through the NSF-funded REU program. We will participate in the fourth COHH meeting in Hawaii, and work to the success of the new Gordon Conference on OHH.

Research and Education Activities, Findings, Training and Development, Outreach, and Products

Anderson – Project 1

Activities:
Work in Year 4 has focused on Specific Aims 3 and 5 of the project:
3) Characterize the relationships between toxicity, physiological variability and genotype in *Alexandrium* spp. from the Gulf of Maine; and
5) track changes in the genotypic diversity of *Alexandrium* populations through time throughout the Gulf of Maine.

Findings:
In this project year, we continued our studies of *Alexandrium* bloom populations to include a third year of data from summer of 2007. In 2005, we collected and analyzed samples from a massive toxic *Alexandrium* bloom that occurred that summer. Results of the microsatellite genotyping showed that the late-bloom and early-bloom populations were significantly different from each other (Fisher's combined test, $p<0.05$), indicating that the genetic composition of the bloom population changed on the order of about 3 weeks. The observed change could be due to natural succession of the bloom community, or by the addition of new genotypes from a separated population, or both.

These two hypotheses were addressed during fieldwork performed during 2006, when we collected samples from across the Gulf of Maine region during a second *Alexandrium* bloom (summer of 2006) as well as from a toxic bloom in an isolated embayment, Salt Pond, MA. In the wider Gulf of Maine, populations were collected from across the region including offshore on Georges Bank, and none were significantly different from one another. A comparison of the 2005 and 2006 bloom samples showed that, in general, populations from the Gulf of Maine blooms in the two different years were not genetically distinct. Populations collected from Salt Pond, however, were genetically distinct from those in the wider Gulf of Maine, and they changed over the course of the 3-week bloom. From these two years of data, it appears that overall genetic composition of *Alexandrium* blooms in the Gulf of Maine is not significantly different from year to year. Within a year, however, we did observe changes in bloom populations on the timescale of approximately one month. This could result from the natural progression or 'turnover' of genotypes during a bloom, or from the mixing of genetically distinct cells from other (unknown) sources. Results of the 2006 analysis of the Salt Pond bloom provides support for the former hypothesis, although the mixing of different source populations cannot be discounted.

We continued our population genetic analysis in 2007, sampling repeatedly in 9 locations across the Gulf of Maine. We are in the process of genotyping over 1000 isolates from 2007, including biweekly samples from the offshore Georges Bank region and the most northerly "upstream" population in the Bay of Fundy. Although *Alexandrium* cells were not as numerous in 2007 as they were in the previous years, the monitoring cruises did reveal high numbers of cells in the offshore region of Georges Bank. The discovery of a large offshore population of cells when inshore areas had much lower numbers could have significant implications for *Alexandrium*

bloom ecology and shellfish toxicity in the region. The Georges Bank region is the focus of activity for a separate project headed by PI Anderson. Supplementary funding from NSF for COHH activities will enable us to send personnel on these cruises to collect samples to continue our population genetic analysis, with the aim of understanding the connection between inshore and offshore *Alexandrium* populations.

In this project year we have also initiated studies of the link between toxicity and physiology in *Alexandrium*. This includes paired growth rate and toxicity measurements on a large number of cultured *Alexandrium* collected from different years. The first experiment used 45 cultures isolated during 2001, and determined growth rates and toxicity under constant conditions. A 1.75-fold difference in maximum growth rates was observed amongst the cultures, ranging from 0.41 to 0.72 divisions/day. A second growth rate experiment used a total of 44 clonal isolates grown at 15°C and 6°C at a light level of either ~400 and 100 $\mu E \cdot m^{-2} \cdot s^{-1}$. Growth rates were determined for all isolates, under each of the four conditions. Toxicity measurements are currently underway. This dataset of matching information on genotype, growth rate, and toxicity will be used to examine the relationships between these three characteristics of *Alexandrium* strains. Ultimately, the information will be used in a modeling effort by WH COHH PI McGillicuddy to examine the role of genetic heterogeneity in the abundance and overall toxicity of *Alexandrium* populations in the region.

During the course of our work, we have involved a number of graduate and undergraduate students, as listed above. These students have worked on several aspects of this project, including the cruise sample collection and isolation of the *Alexandrium* cultures for microsatellite analysis.

Outreach and Impact:

Presentations:

Anderson, D.M. Don't eat the clams: Managing the threat from the New England red tide. Frontiers in Environmental Sciences, Weekly Series on Emerging Issues in Environmental Health Sciences. (Invited presentation at NIEHS headquarters in the *Frontiers in Environmental Science* seminar series. The talk was also webcast, and can be viewed at: http://www.niehs.nih.gov/news/video/science/frontiers/)

Anderson, D.M., B.A. Keafer, K. Norton, D.J. McGillicuddy, R. He, C.H. Pilskaln, and d. Couture. *Alexandrium fundyense* cyst dynamics in the Gulf of Maine. Fourth Symposium on Harmful Algae in the U.S. October 2007, Woods Hole, MA.

Anderson, D.M., B.A. Keafer, D.J. McGillicuddy, and R. He. The 2005 and 2006 New England red tides: mechanisms, management challenges, and implications for future forecasting capabilities", Estuarine Research Foundation meeting, November 2007, Providence, RI.

Erdner, D.L., L.A.R. McCauley, K. Libera, and D.M. Anderson. 2007. Population genetics of toxic *Alexandrium* blooms in the Gulf of Maine. Fourth Symposium on Harmful Algae in the U.S. October 2007, Woods Hole, MA.

Erdner, D.L., L.A.R. McCauley, K. Libera, and D.M. Anderson. 2007. A real-time PCR assay for the toxic dinoflagellate *Alexandrium fundyense*: laboratory studies and field validation. American Society of Limnology and Oceanography international meeting, February 2007, Santa Fe, NM.

Posters:

Brosnahan, M.L., R.J. Olson, D.L. Erdner, and D.M. Anderson. Evidence of a self-recognition system in the sexual life-cycle of *Alexandrium tamarense*. Fourth Symposium on Harmful Algae in the U.S. October 2007, Woods Hole, MA.

Products:

Websites Created

The Northeast PSP site (http://www.whoi.edu/sbl/liteSite.do?litesiteid=3230&articleId=13371)

Research and Education Activities, Findings, Training and Development, Outreach, and Products

McGillicuddy – Project 2

Research and Education:

1. Formulate a suite of population dynamics models for the various genotypes of *A. fundyense*. Model formulation will be guided by existing observations, as well as laboratory experiments to be conducted in Project 1 (Anderson).
2. Incorporate the ensemble of population models into existing models of Gulf of Maine coastal hydrodynamics.
3. Use the coupled physical-biological models to construct hindcast simulations of *A. fundyense* survey observations to be collected jointly with Project 1 (Anderson).
4. Diagnose the simulations to determine the processes regulating the space/time expression of the different genotypes in terms of *A. fundyense* abundance.
5. Utilize toxigenicity data for the various genotypes (provided by Project 1 (Anderson)) together with the coupled physical-biological models (Aim 3) to make predictions of shellfish toxicity along the coast. Toxicity predictions will be tested with observations from ongoing shellfish monitoring programs.

Findings:

Our effort this year was focused on execution of two research cruises, voyages 435 and 437 of R/V *Endeavor*. The first survey (Figure 1, left panel) revealed surprisingly low cell concentrations in the Gulf of Maine, in contrast to the severe blooms that took place in 2005 (Anderson et al., 2005) and 2006. For example, we can compare the southwestern block of our survey plan, from Cape Cod to Cape Ann, with a nearly identical survey conducted on R/V *Tioga* during exactly the same time period in 2006 (Figure 2). Surface live counts in 2007 were all zeros except for one station in which a single cell was observed (corresponding to a concentration of 14 cells per liter). This constitutes a dramatic change from this same time period in 2006, when most of the area was covered by several hundred to several thousands of cells per liter. In interpreting this comparison, keep in mind that the 2007 observations were collected during and in the aftermath of significant wind forcing, resulting in surface mixed layers of up to 40m in some places. Therefore the possibility of subsurface populations cannot be discounted.

Another major finding of our 2007 field work was a large bloom of *A. fundyense* on Georges Bank (Figure 1, left panel). Highest abundances occurred along the Southern Flank, with peak concentrations just over 13,000 cells per liter. The population extended southwest of Georges Bank along the outer continental shelf, consistent with the southwestward exit pathway from the bank at the southern end of the Great South Channel.

Our second survey (EN437) revealed persistence of the bloom on Georges Bank (Figure 1, right panel). Surface live counts indicate that cell concentrations on Georges Bank dropped somewhat since EN435, although there are several stations with thousands of cells per liter and a peak concentration near 10,000 cells per liter. The bank-wide pattern also changed, insofar as the central crest of the bank became a local minimum in concentration. Cell concentrations on the

southern New England shelf rose from near zero during EN435 to several hundred cells per liter during EN437.

Whereas the EN435 survey revealed near absence of *A. fundyense* in the coastal Gulf of Maine, cell concentrations in that area during EN437 were fairly similar to what has been observed in the past this time of year (Figure 1, right panel). A key question is how did the system "catch up" to "normal" conditions given such a late start? This question is even more puzzling in light of the severe blooms that took place in 2005 and 2006.

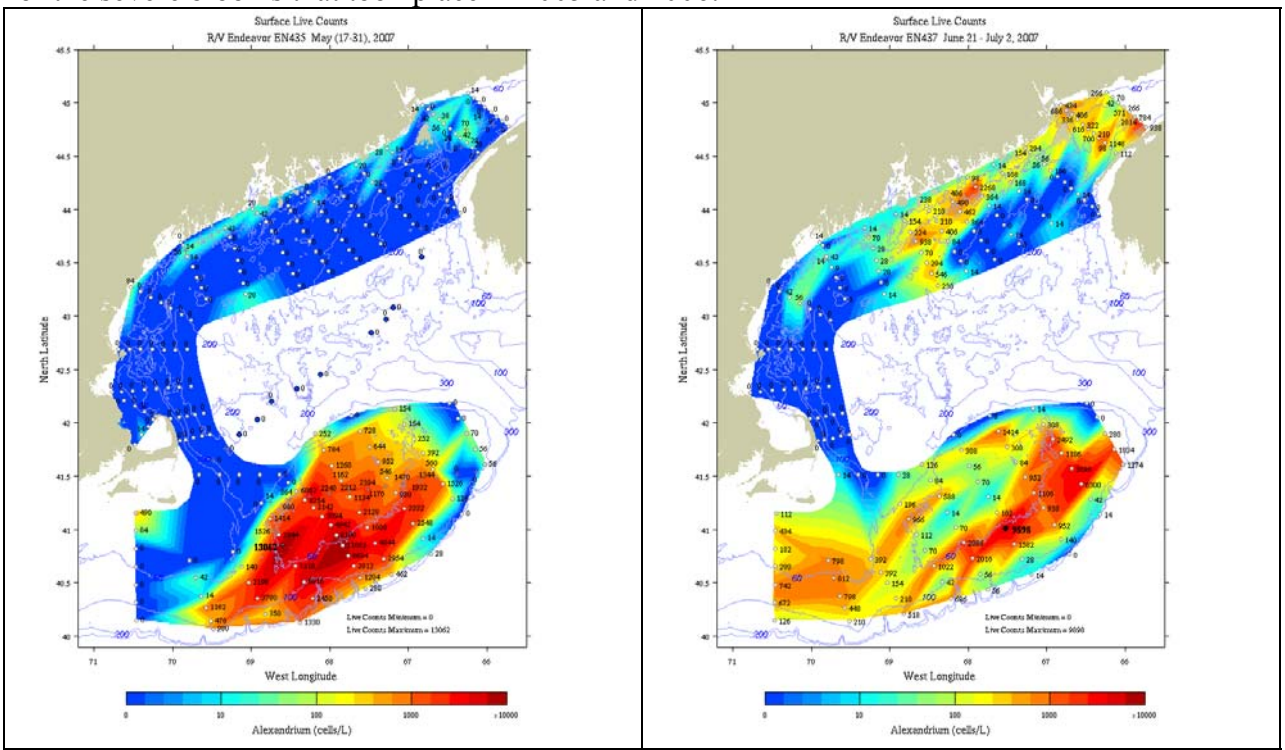


Figure 1. *Alexandrium* concentration (cells l^{-1}) from surface live counts on EN435 (left) and EN437 (right).

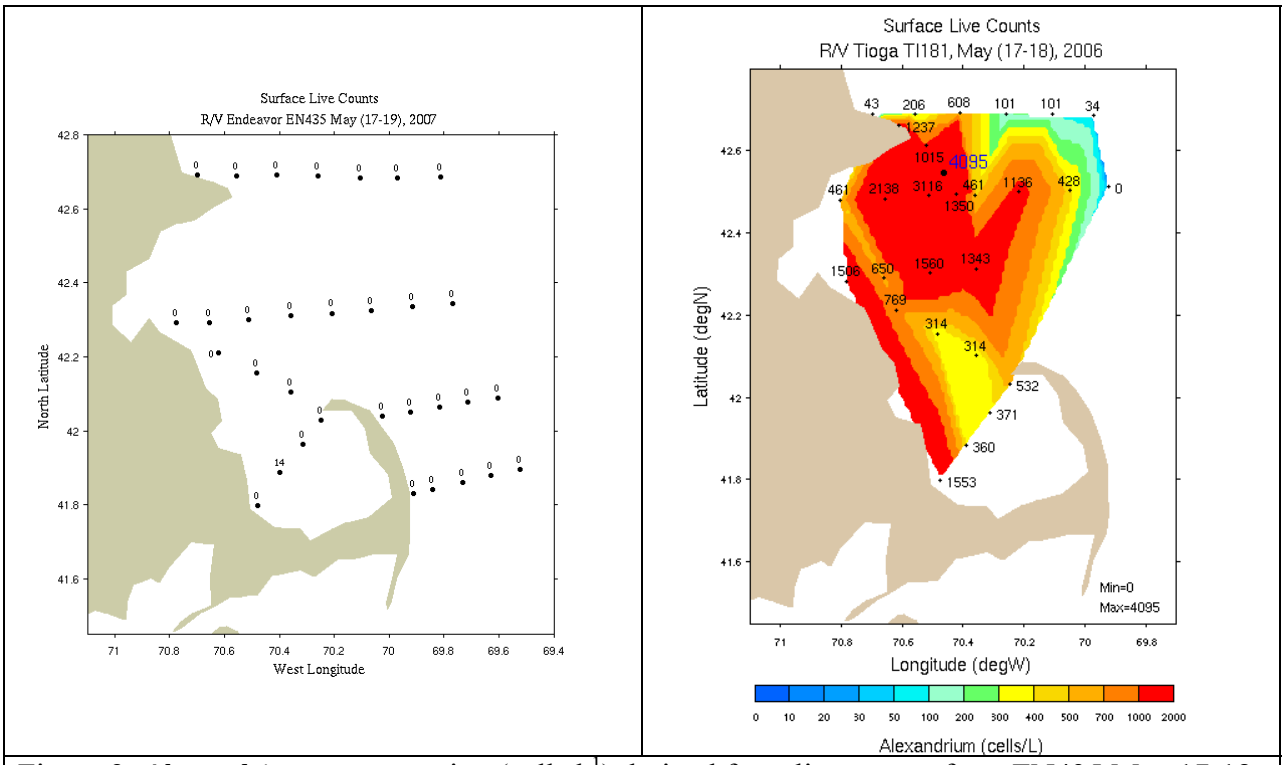


Figure 2: *Alexandrium* concentration (cells l^{-1}) derived from live counts from EN435 May 17-18 2007 (left) versus TI181 May 17-18 2006 (right).

Also in the past year, we developed a simple theory for the observed bimodal vertical distribution of motile phytoplankton (Ralston et al., 2007). Some motile phytoplankton have the capability to exploit deep sources of nutrients in a vertical migration cycle: photosynthesis in the near-surface layer, transit to depth, uptake of the limiting nutrient, and transit back to the surface layer. If all four steps can be completed within 24 hours, then migrations can be synchronized to the day/night cycle to maximize photosynthetic efficiency. Alternatively, if physiological, behavioral, or environmental factors make it impossible for the cycle to be completed in 24 hours, then migration may be asynchronous. Many observations of phytoplankton reveal bimodal vertical distributions of organisms, with maxima near the surface and the nutricline. We demonstrated how bimodal vertical distributions of phytoplankton may be symptomatic of asynchronous vertical migration using a Lagrangian Ensemble numerical model. We simulated vertical migration of the dinoflagellate *A. fundyense* in conditions similar to those in the Gulf of Maine, where bimodal distributions of *A. fundyense* have been observed. Migration is regulated by internal nutritional state – organisms swim down toward the nitracline when depleted of nitrogen, and return to the surface after nutrient uptake. We tested the sensitivity of the results to growth rate, nitrogen uptake rate, and swimming speed, and found that organism distributions can be bimodal or unimodal depending on conditions. Finally, we developed an analytical estimate for population distribution based on organism characteristics and nutricline depth.

Outreach and Impact:

10/06 Toxic algae in the Gulf of Maine: observations and models. WHOI Topics in Oceanography workshop for middle-school and high-school science teachers, WHOI Exhibit Center.

4/07 Modeling Blooms of *Alexandrium fundyense* in the Gulf of Maine: From Climatology to Forecasting. Environmental fluid mechanics seminar, Department of Civil and Environmental Engineering, MIT, Cambridge, MA (Invited).

4/07 Skill Assessment for Coupled Biological/Physical Models of Marine Systems. Ocean Color Research Team Meeting, Seattle, WA.

10/07 Observations and models of *Alexandrium fundyense* blooms in the Gulf of Maine and Georges Bank: From Climatology to Forecasting. Fourth Symposium on Harmful Algae in the U.S., Woods Hole, MA.

Products:

Websites Created:

1. Data from R/V *Oceanus* voyage 412, and R/V *Endeavor* cruises EN435 and EN437:
   http://science.whoi.edu/users/mcgillic/cohh/oc412/data/
   http://science.whoi.edu/users/mcgillic/en435/
   http://science.whoi.edu/users/mcgillic/en437/

2. Near-real-time nowcasting and forecasting of the 2005, 2006, and 2007 blooms:
   http://science.whoi.edu/users/ruoying/Redtide_05/movie.html
   http://omgrhe.meas.ncsu.edu/Redtide/Redtide_06/
   http://omgrhe.meas.ncsu.edu/Redtide/Redtide_07/

Research and Education Activities, Findings, Training and Development, Outreach, and Products

Gast – Project 3

Research and Education Activities:

This proposal examines the diversity, distribution and persistence of human pathogens in the coastal marine environment of Mt. Hope Bay, Massachusetts. This estuarine environment has been heavily impacted by human activity, including significant thermal pollution, several sewage outfalls and the effects of past industrial activities. Monitoring for bacterial contamination occurs in the shellfish areas, but outside of these regions, very little is known about the occurrence of either bacterial pathogens or protistan parasites that are likely introduced into the Bay through sewage or runoff. The potential impact of the thermal output of the Brayton Point Power Plant on these microbial populations has not been examined at all. We hypothesize that both bacterial and protistan human pathogens are present in Mt. Hope Bay and nearby rivers, and that their distribution and persistence are affected by physico-chemical condition in the Bay. To examine this hypothesis we propose to accomplish the following specific aims:

Aim 1: Determine the distribution and persistence of human protistan pathogens in Mt. Hope Bay. We will conduct a molecular survey of human pathogens within the Mt. Hope Bay estuarine system, including sewage outfalls near and removed from thermal point sources, contaminated shellfish beds and shorebird nesting grounds, to better characterize the types of protistan pathogens present in this coastal environment. Our project will specifically target *Giardia*, *Cryptosporidium*, *Naegleria*, and *Acanthamoeba*, but will also perform general microbial eukaryotic and prokaryotic diversity assessments to determine whether there are novel sequence types present that are related to known pathogens (might represent unidentified human pathogens).

Aim 2: Determine whether naturally occurring marine amoebae can serve as reservoirs for pathogenic bacteria, and look for evidence of the natural occurrence of these associations in Mt. Hope Bay. We will test whether locally collected marine amoebae are capable of harboring the pathogenic bacteria *Vibrio parahaemolyticus*, *Vibrio vulnificus*, and *Legionella pneumophila*. The research will test for the induced presence and replication of bacteria in actively growing and dormant phases of the amoebae, and will utilize both PCR-based and in situ hybridization based methods to detect naturally occurring associations.

Aim 3: Establish physico-chemical parameters associated with pathogen presence in Mt. Hope Bay and test predictive capability of water quality models regarding these correlations.
We will correlate measurements of water and sediment physico-chemical parameters with microbial community structure, including the presence of human pathogens. These measurements and our biological data will be used to assist in establishing and validating a water quality model for Mount Hope Bay that may also be used to understand and predict the potential presence of pathogens.

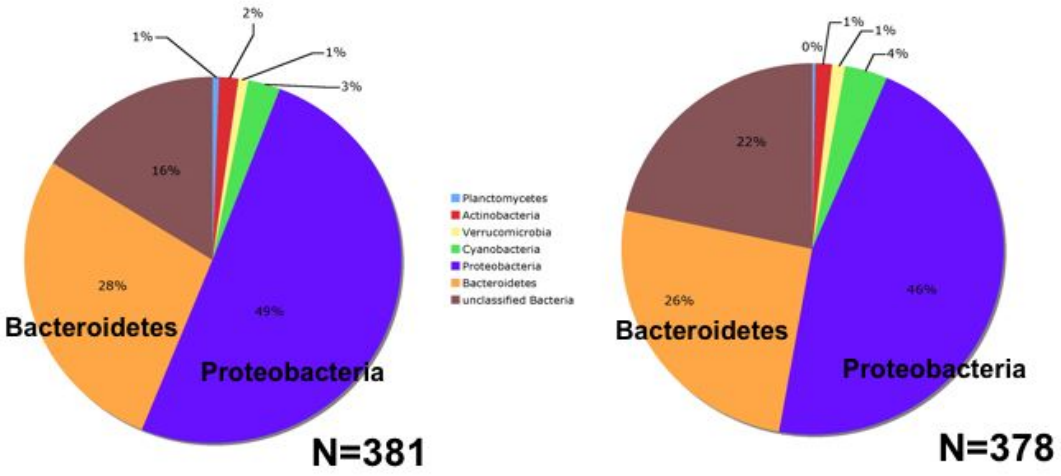
Findings:

Fieldwork. Four sites in Mt. Hope Bay were selected for our survey of human parasites and the effect of physico-chemical parameters on their distribution and persistence. One is located within the thermal plume of the Brayton Point power plant outlet. The second site (Braga Bridge) is located at an underwater sewage outfall that releases secondarily treated sewage into the Taunton River just a bit north of the Bay proper. A third site is farther up the Taunton River (Dighton Bridge) at a brackish water/marsh site, and the fourth site (Common Fence) is near the bottom of the Bay at an area not directly impacted by the thermal plume or the sewage outfall. Sediment and water samples were collected in triplicate at all four sites. Our seasonal sampling officially began in November 2004, and continued through April, June, July, August, November 2005 and February 2006. An additional sampling was carried out in June 2006 in response to large amounts of precipitation and increased runoff. In addition to our four sites we also conducted two transect sampling trips in August of 2005 and February of 2006. Each time our transect began at our Brayton Point site and included 7 stations along a direct line to our control site Common Fence. Transect sampling included water samples (collected at the surface and at depth just above the sediment) and sediment samples. Sediment and water samples were again also obtained and processed from five sites in the Great Salt Lake, Utah.

Results:

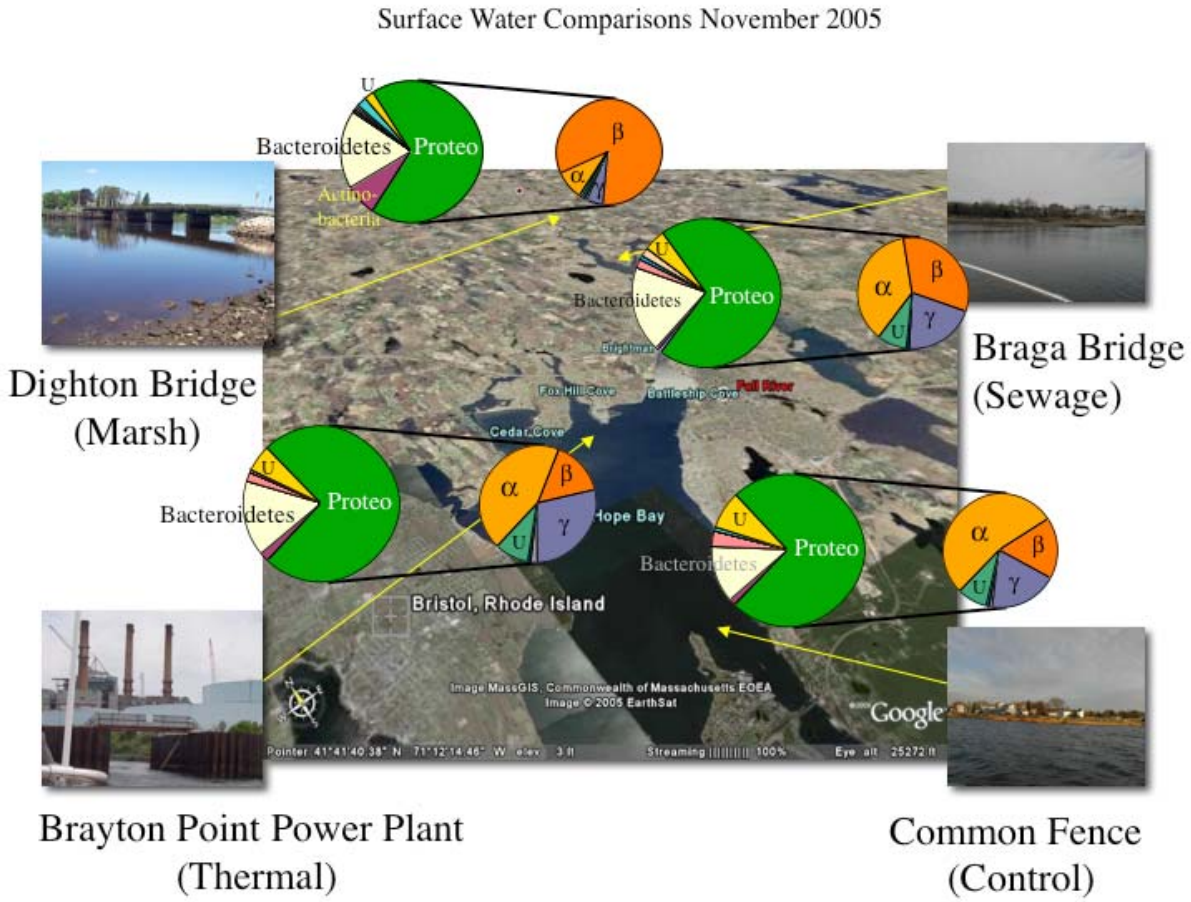
Microbial Diversity Surveys.

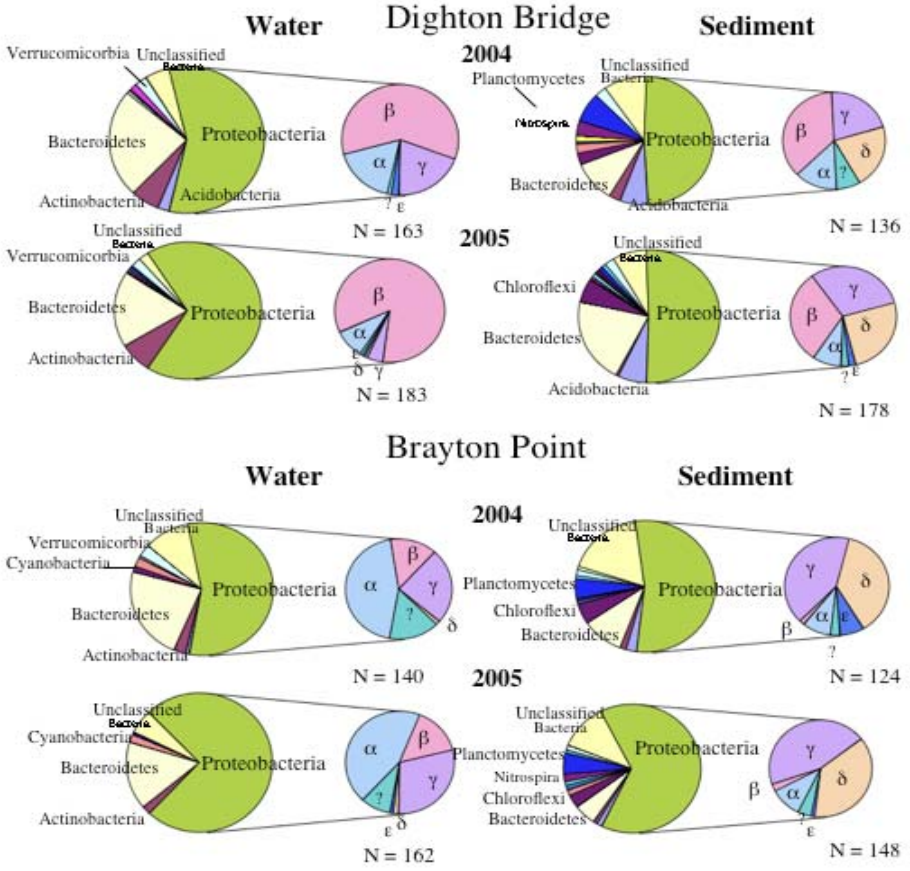
Separate Domain-specific Libraries. We have compiled the first comprehensive (eukaryal, bacterial, archaeal) data from small-subunit ribosomal RNA (SSU rRNA) gene clone libraries for water and sediment samples collected near the thermal plume and underlying sediments of the Brayton Point Power Plant (abbreviated BP in figures below) during November 2004. We have partial sequences of nearly 4,000 clones from 2 different sites and have further sequenced 1,000 unique clones from these to full-length. Not surprising, our findings reveal a highly diverse consortium of the three domains including relatives of sludge bacteria, polyaromatic hydrocarbon-degrading bacteria, and representatives related to the genera *Staphylococcus, Streptococcus, and Clostridium*. Comparisons between overall diversity in water samples for two replicate samples collected at the Brayton Point Power Plant site yielded very similar phylum-distributions. Our phylum-level comparisons between sediment samples, however, revealed significant differences in the microbial assemblages recovered. Preliminary phylogenetic placement of these environmental sequences revealed that the largest number of potential human pathogens were detected in sediment samples and not in the water-column samples. The two subsequent diagrams contrast water versus sediment replicates, asterisks indicate significant differences between groups based on the Ribosomal Database Project II Library Compare Program.



Serial Analysis of Ribosomal Sequences Tag (SARST) analyses: Our clone library constructs provide us with a reference library of full-length sequences for our Brayton Point Power Plant site. This serves as a baseline dataset with which to compare sequences obtained from higher-throughput methods that target short variable regions of small-subunit rRNA genes. Given that our initial phylogenetic analyses identified the largest number of potential human pathogens in sediment samples, we sought to further explore the microbial diversity in these samples using methods that would allow us to sample more deeply. One such method that we have begun applying to our sediment samples is called SARST (Serial Analysis of Ribosomal Sequence Tags). We currently have a total of 2,365 tags from the V6 region of SSU rRNA from one of our Brayton Point sediment samples. 513 of these possessed 100% matches to known organisms in GenBank. Of these we further identified potential pathogens belonging to the following bacterial genera that possess pathogenic representatives: *Shigella, Shewanella, Pseudomonas and Francisella*. The occurrence of *Francisella* in estuarine environments is not well-documented. As a result of this finding, we will begin to screen our samples for this pathogen in both environmental samples and amoeba cultures.

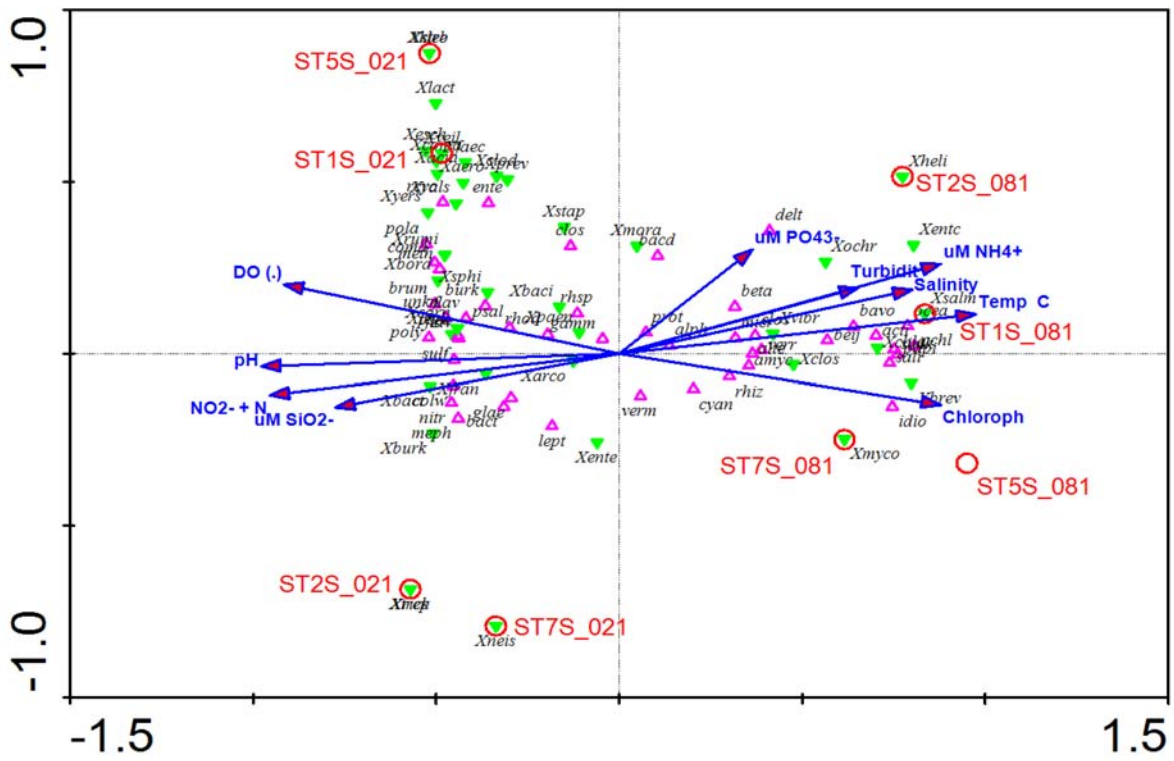
Comparison of 3-domain targeted Clone Libraries: To explore the seasonality and distribution of microbial populations and their corresponding potentially pathogenic members, we constructed clone libraries using primers that target all three domains of life simultaneously and capture the region in between the V4 through V8 hypervariable region of SSU rRNA genes. We constructed libraries from both water and sediment samples from November 2004 and November 2005 sampling seasons for the Brayton Point, Braga Bridge, Dighton Bridge and Common Fence locations. For each library we selected 192 clones for DNA sequencing in the forward and reverse orientations. Clone library composition was dominated by bacteria although some sediment samples yielded archaeal hits. The data have passed through initial quality control screening but still need to be screened for possible chimeric sequences. The pie charts below show preliminary data at the phylum level that were generated using the Ribosomal Database Project II Classifier program that assigns taxonomy to rRNA gene sequences as an example of the kind of data that are forthcoming from our analyses.





Seasonal Extreme Transect Study applying 454 Tag Sequencing:

Our final and most powerful approach employs a method developed by Mitchell Sogin and colleagues as part of the WHCOHH Genomics facility. This approach targets the same hypervariable region as the SARST-V6 method but circumvents the cloning step by proceeding directly to tag sequencing via pyrosequencing chemistry and 454 sequencing technology. As in SARST-V6, microbial tag taxonomic assignments are based on similarity to a reference SSU rRNA V6 library. Both SARST-V6 and 454 V6 tag sequencing allow for the recovery of both kinds of microbes present in a sample and their relative abundances. Our first run of the 454 machine has resulted in the recovery of 113,893 tags and included 4 summer and 4 winter surface water samples along a transect away from the power plant. The Canonical Correspondence Analysis triplot (CANOCO 4.5) below shows the relationship between Operational Taxonomic Units (OTUs = bacterial "species"), winter (STXS_021) and summer stations (STX_081) and environmental parameters. The plot includes the overall top 50 most abundant OTUs and those of any potential pathogens (indicated by an inverted green triangle and containing an "X" prefix). This analysis is preliminary but is compelling in that it reveals an abundance of potential pathogens falling in the area near Station 1 (the same as Bp 1), winter (ST1S_021). A second 454 run has just been completed with water column samples from the surface and at depth, as well as sediment samples. This second run was conducted with redesigned primers to target a broader range of bacteria. Analyses of these data are planned for the next several months.



Detection of specific pathogens. We continue in the process of setting up amplification arrays with our extracted samples for use with organism specific primers. These include primers for *Giardia, Cryptosporidium, Acanthamoeba, Naegleria* and *Legionella*. The *Legionella* primers were used in nested amplifications of our seawater, sediment and ameba culture samples. Sediment and ameba cultures from Mt. Hope Bay showed positive amplification results, and there appears to be a seasonal pattern to the prevalence of *Legionella*-like species in amebas. All of the water and sediment samples from the Great Salt Lake were positive using the *Legionella* amplification method. Both MHB and GSL samples are not often positive for *L. pneumophila*, but it has been detected. *Giardia* is also being detected in sediment and water samples of MHB, primarily in the summer months.

Legionella and amoeba sequences. Samples that were positive for the presence of legionellae were amplified, cloned and sequenced. In this work we have now newly described both the significant diversity of legionellae in the marine environment and have documented the seasonal pattern of their abundance. During the past year we have focussed on analyzing the *Legionella* sequence types present within amoeba cultures positive for the genus. Overall, 41% of MHB amoeba cultures were positive for the *Legionella* genus, while 57% of the GSL were positive. Of the positive cultures, 10% of the MHB amoeba cultures and 20% of the GSL amoeba cultures

were positive for *L. pneumophila*. It is of note that all of the amoebae that were positive for *L. pneumophila* were growing on either marine or brackish water media. This indicates that marine amoebae are very capable of supporting the growth of *Legionella pneumophila* in the marine environment. When we examined sequence types present in the non-pneumophila positive amoebae, we found that 31% were related to legionella-like amoebal pathogens, while 16% were related to other *Legionella* species that are also human pathogens. Seventy percent of the sequence types recovered from amoebae were also recovered directly from the environment, suggesting that they comprise a significant proportion of the *Legionella* population. We also investigated what types of amoebae were able to harbor *Legionella* species. Based upon small subunit ribosomal sequences, *Acanthamoeba* and *Hartmannella* were identified, along with several other genera of amoebae that have not previously been identified harboring these bacteria. This suggests that most species of amoebae have the potential to harbor bacterial pathogens in the natural environment (including the marine environment).

Doheny Beach/Avalon Beach Epidemiology Study. We have become involved in the Southern California Coastal Water Research Project (SCCWRP) beach epidemiology study. While the primary objective of the study is to correlate new indicator studies with human disease incidence, we requested to be included in the project to examine the incidence of *L. pneumophila* with respect to respiratory symptoms. We have also acquired beach sediment samples for analysis of the presence of human pathogens, including *Brucella* and *Campylobacter*. All of the water samples have been processed for Legionella detection, and those results have been submitted to SCCWRP for inclusion in the epi analysis. The sediment work was started in the fall by Elizabeth Halliday, a graduate student supported by the grant. She has developed quantitative PCR methodology that allows the identification of percent recovery and PCR inhibition so that more accurate values can be obtained, and has begun analyzing the beach sand samples from Avalon. We will also participate in the Miami COHH group beach epi study currently underway.

Significance. We have completed over a year's worth of sampling within the Mt. Hope Bay system and have been processing these samples for pathogen presence and microbial community structure. Although the detection of legionellae in marine environments was documented, the extensive presence of Legionella-like sequences in Mt Hope Bay was unexpected. This is also true for the Great Salt Lake samples. These organisms appear to be fairly prevalent in amoebae recovered from the sediments of these environments, and the number of sites that they were recovered from increased as the year progressed. This suggests that legionellae can persist and spread within the marine environment.

Our clone library analyses are still in progress, so we cannot make any conclusions at this time, however, the baseline data recovered from our study will provide a valuable comparative dataset for other studies of thermally and sewage impacted environments. Our microbial diversity data suggest that there are a diversity of potential human pathogens that can be detected using our employed approaches. Whether or not these potential pathogens can further become associated with free-living amoebae is a question we still need to explore. Of particular interest are

sequence tags recovered from both SARST and 454 tag sequencing that have a 100% match to sequences in the genus *Francisella*. *Francisella* is an obligate endosymbiont and *F. tularensis* a known human pathogen and causative agent of Tularemia. Tularemia is a localized problem on the nearby island of Martha's Vineyard, so ascertaining whether the human disease-causing strain is harbored in association with estuarine environments will be an important goal for the project over the next year.

Plans:
For the next 12 months we will be analyzing samples from the Doheny/Avalon study and summarizing the microbial diversity data from our combined approaches. *Giardia, Cryptosporidium, Naegleria and Acanthamoeba* amplifications will be completed for water, sediment and guano samples. We will also explore the relationship between our microbial diversity data and corresponding environmental data, and will begin working with the MHB FvCOM model to examine the impact of hydrodynamics on microbial populations and distribution. Manuscripts will be prepared that summarize the results of our Mt. Hope Bay study.

Training:
Support of a graduate student (Halliday), training of a postdoc (Del Castillo), mentoring of undergraduate students through REU programs, informal epidemiology training, and participation in SCCWRP California beach study and Miami beach study.

Gast has constructed and taught a graduate level course on Oceans and Human Health. Lectures are available to other COHH researchers.

Gast was a speaker at the WHOI HARP undergraduate career symposium "Global Environmental Challenges in Oceanography: Climate Change and Oceans and Human Health" June 27, 2007

Outreach and Impact:
Gast is a new Subject Editor (zoonotic diseases, protists) for Diseases of Aquatic Organisms, and is a co-editor for a special issue of Diseases of Aquatic Organisms on Marine Vertebrate Zoonoses, and will serve as a co-editor for a special issue on Marine Vertebrate Zoonoses.

Gast participated in a Massachusetts shellfish officers meeting with Stegeman during the summer to discuss potential interactions and impacts of COHH projects.

Products:
Websites Created:
1. COHH website (www.whoi.edu/science/cohh/whcohh/)
2. Gast homepage (www.whoi.edu/people/rgast/index.html)
3. Zoonotic disease website (www.whoi.edu/sites/zoonoses)

Research and Education Activities, Findings, Training and Development, and Outreach

Polz – Project 4

Research and Education Activities:
Our research pursues two overarching goals. The first is to explore environmental forcings on multiple spatial and temporal scales leading to abundance of pathogenic vibrios (Aim 1 and 2); the second is to understand the nature of genetic differences and gene flow between co-existing pathogenic and non-pathogenic variants of vibrios (Aim 3 and 4). All our aims are pursued within the Plum Island estuary and Parker River where we determine whether changes in estuarine physics, water chemistry and biology correlate with transitions in diversity and abundance of different *Vibrio* species and to determine estuarine conditions that might trigger the emergence of pathogenic strains of *Vibrio*. Overall, we are nearing completion of the tasks within AIM 1-3, and report progress on AIM 4 since last year's report.

Findings:

AIM 1 and 2. To characterize and model dynamics and reservoirs of *V. vulnificus* and *V. parahaemolyticus* populations over seasonal cycles, and to test the link between estuarine physics, nutrient and particle abundance and growth patterns of *Vibrio* species over tidal cycles
We summarize here progress on both aims. The sampling for both aims is complete and was spelled out in the last report. The analysis of population structure by culture-independent assays is currently in progress and will be finished by early 2008.

To determine ecologically coherent groups among co-existing vibrios in the water column, we examined the temporal and spatial distribution of *Vibrionaceae* genotypes. We sampled the water column to differentiate the free-living and attached (to both particles and zooplankton) compartments of the planktonic community under different macroecological conditions (spring and fall). Particle-attached and free-living cells were separated (in 4 replicates each) into a total of four consecutive size

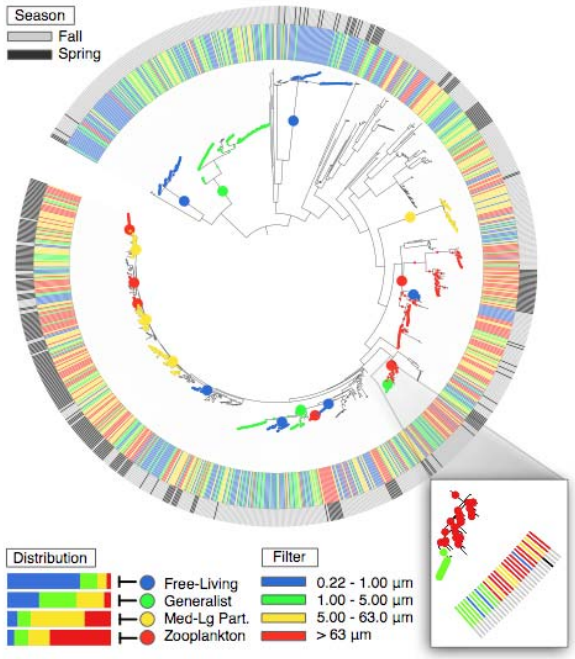


Fig. 1 Mapping genotypes across environments. Shown is a phylogeny of *Vibrio* isolates based on the *hsp60* gene. The inner ring depicts by color the size fraction from which each strain was isolated (see legend in figure), and the outer ring shows the season. The *AdaptML* software package was used to infer four commonly occurring distributions for phylogenetic groups (see legend in figure), and ecological populations demonstrating one of these characteristic distributions are indicated by colored dots (large dots indicate the ancestor of the group). The inset figure shows two closely reated groups (nested clades) that have distinct ecological distributions.

fractions, which are enriched in zooplankton (≥63 μm), large (63-5 μm) and small (5-1 μm) particles, and free-living cells (1-0.22 μm). The 5-1 μm size fraction is somewhat ambiguous, likely containing cells attached to small particles, as well as large or dividing cells; however, it provides a firm buffer between obviously particle-attached (>5 μm) and free-living (<1μm) cells.

Roughly 1,000 isolates were characterized by sequencing of a protein-coding gene (*hsp60*). To confirm relationships, between 1 and 3 additional gene fragments (mdh, adk and pgi) were sequenced for all *V. splendidus*, the dominant taxon during warm water conditions (Thompson et al., 2005b), and several other microdiverse groups. These data allow conservative estimation of ecological differentiation because inadvertent mixing of strains among microhabitats and homologous recombination among strains homogenize rather than create associations.

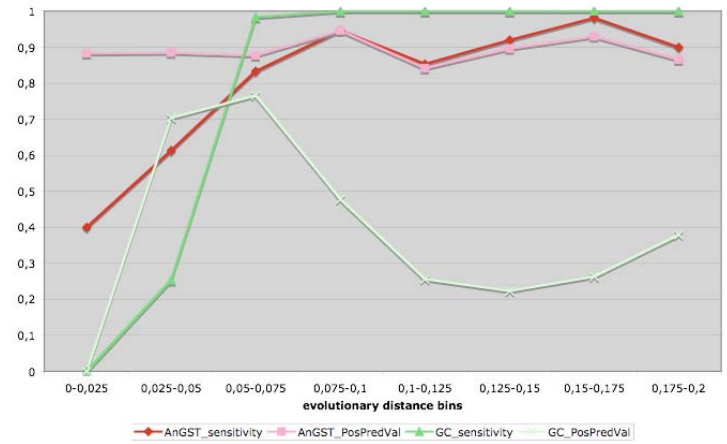


Fig 2. *AnGST-HR* performance on simulated vibrio data compared with pairwise approach. Simulated gene sequences were simulated according to an actual vibrio phylogeny inferred from 100 MLSA samples, and recombination events between leafs were randomly inserted into the tree among taxa at different evolutionary distances. *GENECONV* (GC) and *AnGST-HR* were evaluated based on their sensitivity (1- fraction of false negatives) and positive predictive value (PPV = fraction of true predictions that are correct). Both methods show a 'blind spot' for detecting recombination between very similar sequences, but *AnGST-HR* shows a high PPV across the entire range of distances, while *GENECONV* tends to make false positive predictions when highly divergent sequences are considered.

In collaboration with Eric Alm of the MIT Systems Microbiology group, we developed a software package (*AdaptML*) to identify associations between phylogenetically-defined groups and microhabitats that combines a 'model-free' empirical approach and an explicit evolutionary model of habitat association. Briefly our empirical framework establishes whether or not each clade in our strain phylogeny shows a non-random distribution across environments, while the explicit evolutionary model estimates population boundaries, rates of habitat 'switching', and the distribution of habitats across a set of measured environmental parameters.

Our data reveal important aspects of vibrio ecology in the water column. First, phylogenetic groups with distinct ecological preferences are clearly identifiable among all clades with sufficient representation in the dataset (Fig. 1). There is strong temporal structure evident where populations occurring in the spring and fall are frequently distinct at deep phylogenetic levels (Fig. 1). However, spatial partitioning is also evident for many groups, including *V. splendidus*, which contain many shallow clades with displaying distinct distributions (*V. splendidus* occupies the lower section in the tree spanning ~4-10 o'clock, Fig. 1). We were even able to identify nested clades where a single nucleotide change in several housekeeping genes is correlated to a switch in environmental distribution and thus suggests that this group is diversifying (one of these is highlighted in the inset to Fig. 1).

Second, although the currently crude spatial sampling (by size fraction) does not allow assignment of specific habitats (e.g. specific types of organic particles, zooplankton body sections), the analysis indicates differentiation into populations with broad ('generalist') and more narrow ('specialist') distributions, respectively. For example, *V. ordalii* was identified as specialized to free-living lifestyle, while the potentially pathogenic *V. parahaemolyticus* appears to be a generalist inhabiting particles, zooplankton and the free-living fraction of the water.

AIM 3 and 4. To determine the 'rules' of recombination, gene transfer and clonality in co-occurring *V. vulnificus* and *V. parahaemolyticus* and related vibrios, and to assess the diversity, mobility and evolutionary history of genes implicated in pathogenicity
Completion of this aim required the development of a new tool for recombination determination (Fig. 2). AnGST-HR is a modification of the AnGST software, which can identify gene dynamics within a comparative genomics framework. Fig. 2 shows that the method already performs better than other current methods in simulated datasets. The software is now being adapted to multi-locus sequence data sets to allow determination of recombination rates against sequence distance, which represents a critical parameter in judging the genetic isolation of different populations of microbes.

We have leveraged the work carried out under the auspices of the COHH to obtain funding from the Moore Foundation for sequencing of several of our genomes. We already have one *V. splendidus* and *V. alginolyticus* genome in hand. Technical developments now enable us to sequence 21 additional genomes by a combination of 454 and Solexa sequencing. This will provide an unprecedented dataset for both aims. Moreover, we are in the process of determining correlation of O-antigen diversity to population structure in the populations isolated from diverse habitats (water column, animal associated, etc.).

Significance:
The last few years have seen rising concern about the emergence of new variants of pathogens and spread of existing pathogens due to local or global environmental change. This has focused attention on the ecological context of pathogens in both the human body and the environment. Advances in population biology, aided by genomics, have demonstrated that many closely related (genomic) variants of microbial species exist in the environment. Furthermore, it has been shown that virulent bacteria frequently emerge from non-virulent strains via lateral gene transfer and it has been suggested that bacterial genomes are capable of extensive recombination. This raises the fundamental question to what extent observed genomic variants represent ecological and evolutionary units that can be seen as the bacterial equivalent to the eukaryotic sexual species. Do observed genomic variants occupy different environmental niches or do they represent a common gene pool capable of rapid 'assembly' of new variants in response to environmental challenges? These questions are crucial for interpretation of pathogen biology, risk assessment of emergence, and insights into how representative currently extensively studied strains (e.g., of *E. coli* or of pathogenic species) are for the 'species' they represent.

Products:
We are currently collaborating with colleagues in Austria and Germany to develop widely applicable microarray software based on our paper Marcelino et al. (2006).

Further, we have developed the AdaptML software, which can be used to detect ecologically cohesive bacterial populations among strains sampled from varied environmental samples.

Research and Education Activities, Findings, Training and Development, and Outreach

Sogin – Genomics Core

Research and Education Activities:

a. Facility operations and administration:

The genomics core provides DNA sequencing services and computational support to investigators in the Woods Hole Center for Oceans and Human Health (WH-COHH). Many of the Center's projects take advantage of our strengths in molecular microbial ecology and the bioinformatics. Over the past calendar year, the genomics core facility has provided WH-COHH projects with just over 31,000 reads corresponding to ~23 million base pairs. **Figure 1** Shows the usage pattern for the genomics core in 2007 and **Figure 2** describes the use of the Genome Core Facility by different WH-COHH projects.

Figure 1. WH-COHH sequencing usage pattern.

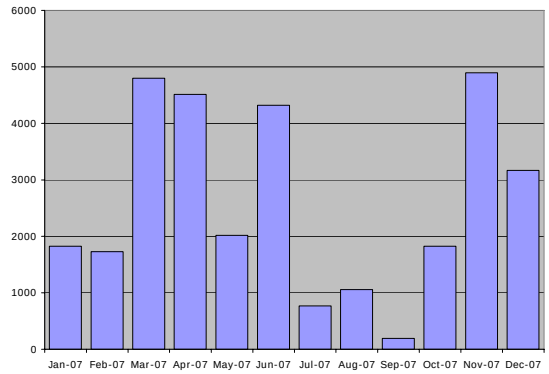
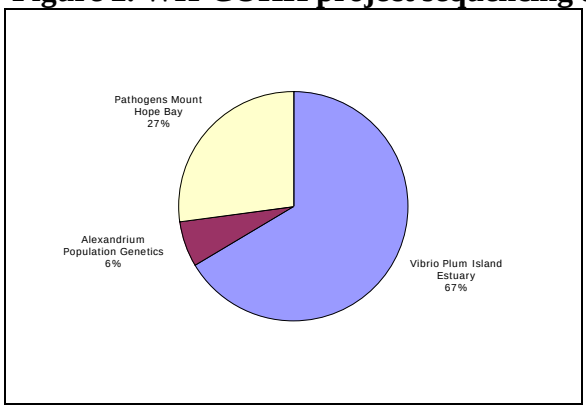


Figure 2. WH-COHH project sequencing activity.



Sequencing costs have changed little over the past year. Each template costs 30 cents using a protocol run on our BiomekFX platform. This is still a relatively inexpensive template production cost. In July, our Roche Genome Systems20 (GS20) pyrosequencing system (funded by a award titled ***Microbial population structure of the world's oceans*** from the W.M. Keck Foundation to the Marine Biological Laboratory at Woods Hole) was upgraded to a GS-FLX, which averages ~400,000 reads of ~238 nucleotides/read in each run. We anticipate that this

system will continue to be a key element of future COHH investigations. With this system, we have generated V6-tag datasets (as described below) for several of the MHB transect samples, providing very deep sampling to complement the traditional clone libraries outlined in the original proposal. We have also sequenced the genomes of six *Vibrio* phages in collaboration with Drs. Jonathan King and Peter Weigele (WH-COHH pilot project) using the GS-FLX.

b. Scientific and technology developments in the genome core.

Using the GS20 (and later, the GSFLX) sequencer the core facility continued development of a new, massively parallel DNA sequencing approach to characterize microbial populations with unprecedented level of sensitivity for detecting low abundance organisms. The International Census of Marine Microbes (ICOMM) provided the test samples and funds for reagents and the WH-COHH provided computational support for analysis of the data. We described in our last report the use of this approach to assess the species richness and eveness in water samples from the North Atlantic Deep Water flow and from Axial seamount (Sogin et al., 2006). During 2007, we demonstrated the feasibility of using 5-nt "keys" to multiplex reactions on the same region of a plate, deconvoluting the reads into sample bins informatically. This approach has been used on hundreds of samples from seawater, freshwater, sediment, soil, sewage, and human feces. We analyzed just over 1.3 million V6 tag sequences from human stool samples (500,000 from David Relman's laboratory and 800,000 generated in our laboratory as part of a collaboration with Jeffrey Gordon). This will provide a reference data for interpreting V6 tag sequences studies from human-impacted environments that we are studying in the context of a WH-COHH pilot project. Our data processing pipeline is well established, and datasets from tag and full length studies are displayed using a new web tool, Visualization of Microbial Population Structures (VAMPS), developed with support from the Sloan Foundation (awarded to M.L. Sogin and David Mark Welch).

Findings:

The ability to accurately characterize microbial population structures through the use of tag sequencing protocols is contingent upon the sequencing accuracy of this techonology. The published accuracy of 454 sequencing technology is only 96%. In genome projects, highly-redundant consensus assemblies can compensate for sequencing errors. In contrast, studies of microbial diversity that catalogue differences between PCR amplicons of ribosomal RNA genes (rDNA) or other conserved gene families cannot take advantage of consensus assemblies to detect and minimize incorrect base calls. To explore error modalities, we generated more than 340,000 reads from a PCR amplicon library that was prepared from a collection of 43 reference templates of known sequence. Each reference template contains a distinct rDNA including the V6 hypervariable region from a collection of 43 divergent bacteria. Differences between pyrosequences and their cognate reference sequences identified signatures of low quality data. We determined that 454 technology has a 99.5% accuracy rate in unassembled sequences, and we identified several factors that can be used to remove a small percentage of low-quality reads, improving the accuracy to 99.75% or better. Table 1 lists the source of those errors.

Table I.

Data Selection	Percent of Reads	Error Rate
All reads	100.0%	0.49%
Reads with no Ns	94.4%	0.24%
Reads with more than one N	5.7%	4.7%
Reads with length >=81 and <=108	98.8%	0.33%
Reads with length <81 or >108	1.2%	18.9%
Reads with no Ns and length >=81and <=108	93.3%	0.20%
Reads with no proximal errors	97.0%	0.45%
Reads with fewer than 3 proximal errors	>99.99%	0.48%
Reads with more than 3 proximal errors	<0.01%	12.2%
Reads with no Ns and length >=81and <=108 and no proximal errors	90.6%	0.16%

The simple removal of sequences that contain any unidentified ambiguities within the sequence of the target amplicon and elimination of sequences shorter than 50 bp or 20% longer than the target amplicon reduces the error rate to less than 0.25%. This is roughly ½ the error rate associated with capillary methods. This work appeared in Genome Biology (Huse, S.M., J.A. Huber , H.G. Morrison, M.L. Sogin and D. Mark Welch. Accuracy and quality of massively parallel DNA pyrosequencing. Genome Biology, 8: R143).

We have also demonstrated that the small amplicon size (~100 bp) more accurately captures the microbial diversity in a sample, most likely because small amplicons contain fewer regions of secondary structure that cause the polymerase to dissociate from the template. Results are shown in Table II. A manuscript describing this result is in progress (Huber, J.A., H.G. Morrison, S.M. Huse, P.R. Neal, M.L. Sogin, and D.B. Mark Welch. Effect of PCR amplicon size on assessments of microbial diversity and community structure, unpublished).

Table II.

Library	No. Clones	Forward Primer	Reverse Primer	% Exact Matches[a]	% Unique Sequences[b]
FS312_100bp	761	967F	1046R	44%	51%
FS312_400bp	860	967F	1391R	74%	22%
FS312_1000bp	381	337F	1391R	72%	12%
FS396_100bp	685	967F	1046R	11%	40%
FS396_400bp	866	967F	1391R	48%	22%
FS396_1000bp	677	337F	1391R	38%	19%

[a] Percent of V6 sequences within library that are exact matches to an existing sequence in the reference database

[b] Perecent of V6 sequences within library not detected by the other two libraries from the same sample

Training:

Through the genome core activities we continue to train a graduate student, Yuko Hasegawa in the use of advanced genomic techniques in the analysis of microbial communities associated with anthropogenic activities.

Outreach and Impact:

Drs. Sogin and Morrison have given lectures as part of a continuing education seminar series for Falmouth physicians and science teachers titled “Medical Genetics”. Analysis of microbial communities using the novel tag approach was included in the material presented.

Proposals Funded:
VAMPS – Visual analysis of Microbial Population Structures
Alfred P. Sloan Foundation. M.L. Sogin PI $300,000 July 1 2007- Jun 30, 2009.

Research and Education Activities, Findings, Training and Development, Outreach, and Products

Hahn – Pilot Project Program

Activities:

The specific objectives of the Pilot Project Program are:

A. To assess the feasibility of new areas of study, especially those that are not currently represented in our Research Project base but would contribute to the overall goals of the center.

B. To support the collection of preliminary data that can be used to generate full proposals to NSF, NIH, or other agencies or organizations.

C. To recruit scientists not currently involved in research on oceans and human health, but who may have expertise in one of these areas, to become participating members of the center and interact with other center investigators.

D. To stimulate collaborative and interdisciplinary research within the center.

E. To foster the application of new technologies and experimental approaches to questions concerning the impact of oceanic processes on public health.

F. To ensure the ability of the center to respond rapidly to new scientific information and emerging challenges in this field.

G. To contribute to the training of future researchers in the field of Oceans and Human Health through enhancement of graduate and postdoctoral training opportunities.

Findings:

Completed pilot projects (2004-2007) (list of publications and grants at end)

- *Characterization of a cyanobacterial anti-algal compound* (Eric Webb and Chris Reddy, WHOI).
- *Cnidarian toxins against voltage-gated Ca2+ channels* (Robert Greenberg, MBL).
- *Marine phage as vectors of gene transfer between marine bacteria and bacterial pathogens* (Peter Weigele and Jonathan King, MIT).
- *Transcriptome profiling in the harmful alga Aureococcus anophagefferens.* (Sonya Dyhrman, WHOI).
- *Beach Pathogens* (Steve Elgar, Britt Raubenheimer, & Rebecca Gast, WHOI)
- *Names-based cyberinformatics tools for rapid response communications and outreach during event management – a pilot based on harmful algal blooms in NE US coastal waters* (David J. Patterson, MBL, and Don Anderson, WHOI).

<u>Ongoing pilot projects (2006-2008)</u>

• *Anthropogenic impacts and profiling fecal microbial populations at a salt marsh* (Mitch Sogin, MBL)

Fecal coliforms are indicator organisms that warn of possible fecal contamination and its potential impact on human health. Coliform surveillance activities generally rely upon cultivation assays or molecular Microbial Source Tracking (MST) technology to assay for the presence of a particular indicator organism in environmental samples. This project will employ a

massively parallel DNA tag sequencing strategy to profile entire microbial communities in the Little Sippewissett Salt Marsh and the barrier Woodneck Beach. Possible coliform sources include human waste from failed septic systems, bird populations or other animals. The objective is to use the tag sequence data to locate specific sources of fecal contamination and identify suites of genes that could serve as multi-species indicators of human pollution. The experimental strategy takes advantage of rapidly evolving hypervariable regions in ribosomal RNAs and our ability to generate many thousands of short DNA tag sequences using 454 Life Science's pyrosequencing sequencing technology.

• *Transcriptional Markers of Life Cycle Transitions in Harmful Algal Blooms* (Don Anderson, WHOI)
Bloom dynamics of the red tide dinoflagellate, *Alexandrium fundyense* are driven in large part by transitions in its life cycle. While these stages are well documented, the biological and oceanographic forces that trigger transitions between the stages are not. A major obstacle to determining the conditions that trigger these transitions is our inability to rapidly identify sexual stage cells. Here, a novel transcriptome experiment will be used to discover molecules that are uniquely expressed by conjugating gamete cells and by germinating cysts. This experiment will utilize sequencing-by-synthesis technology that is newly available through the Genomics Core.

• *The Economic Effects of Harmful Algal Blooms: A pilot project to estimate the costs of human respiratory ailments associated with aerosolized brevetoxins* (Porter Hoagland and Di Jin, WHOI; Lora Fleming, Miami)
This study will estimate the costs-of-illness associated with human respiratory ailments that arise as the consequence of the aerosolization and coastal to inland transport of brevetoxins from blooms of the marine dinoflagellate, *Karenia brevis*, in the Gulf of Mexico. The research will develop models to link the occurrence of HAB events in the coastal-ocean with exposures to aerosolized brevetoxins. The researchers will compile datasets and develop models of illness rates that would permit historical estimates of these kinds of impacts and the simulation of future potential impacts.

• *The Economics of Human Health Risks from Pathogens and Toxins in the Marine Environment* (Hauke Kite-Powell and Porter Hoagland, WHOI)
These researchers will produce an "order of magnitude" estimate of the annual human health cost imposed on residents of the United States by exposure to pathogens and toxins from the marine environment. The estimate will be derived from a review and synthesis of information in the existing literature on (1) the spatial and temporal prevalence of marine pathogens, (2) the pathways by which they affect humans and the potentially exposed populations, (3) the human health effects of exposure, and (4) the economic cost of resulting medical conditions. The estimate will inform future research on pathogens in the marine environment, allowing scholars and public officials to target pathogens and settings where improved scientific understanding is most likely to produce significant economic benefits, and setting the stage for focused economic analyses. (co-funded with the WHOI Marine Policy Center)

Fourth call for pilot project proposals.
The fourth call was issued in September of 2007 by email to all faculty and research staff at WHOI, MBL, and MIT. On November 1, 2007, nine proposals were received and subsequently

were reviewed and scored by members of the Internal Advisory Committee and external reviewers, including several members of the other COHH Centers. The proposals requested approximately $308,000 in total costs. All applicants were provided with written reviews. Three projects were selected for funding:

Hydrodynamics and Transport Pathways for Fecal Microbial Populations in a Salt Marsh and Barrier Beach System (David Ralston, WHOI, Dept of Applied Ocean Physics and Engineering)

We propose to develop and implement a numerical model of flow and transport in a salt marsh and barrier beach system on Cape Cod, MA. The project will focus on Little Sippewissett Marsh and Wood Neck Beach, a system that during summer months frequently experiences high fecal coliform concentrations that indicate potential for significant impacts on human health and lead to restrictions on recreational use. Potential sources of fecal contamination include failed septic systems, birds, or other animals in the marsh. Recent investigations have found that fecal coliform counts are highly variable temporally (tidally, with precipitation events, and seasonally) and spatially across the marsh and beach. A COHH pilot project currently underway has begun to quantify the spatially heterogeneity of microbial communities in the system, intending to distinguish among different sources of fecal contamination and to develop a broader suite of indicator organisms (Sogin, 2006). The project proposed here will work in conjunction with that microbial community mapping project to quantify transport pathways, residence times, and exchange rates in the marsh. Combining transport mechanisms with the spatially heterogeneous source terms will permit calculation of potential exposure and risk to human health associated with the disparate sources of contamination. The model results will also provide guidance for public health officials to redesign of monitoring efforts to sample at times and locations of maximum potential exposure to elevated coliform concentrations, thereby minimizing human health risks. The tasks in the project include acquiring bathymetric data from existing sources and from new field surveys, constructing a numerical grid from the bathymetry, establishing and collecting data for model boundary conditions, and initial calibration and testing of the model based on available observations in the marsh. The model is intended to serve as a basis for future interdisciplinary studies at the study site – model results will aid in design of field observations, and more extensive field studies will aid in refinement of the model. The salt marsh and barrier beach of the study site represent a common coastal environment, and the impacts on human health of fecal contamination in coastal settings are of concern regionally and nationally. The pilot project will provide a step toward building an integrated program in Little Sippewissett to study linkages among physical, chemical, and biological processes in salt marshes and very shallow estuaries, coastal environments where frequent human interactions make potential health impacts particularly significant.

Using signature tagged mutagenesis (STM) to investigate how pandemic Vibrio parahaemolyticus persists in the bacterioplankton and associates with epithelia in the marine environment (Janelle Thompson, MIT, Dept of Civil and Environmental Engineering)
The overarching goal of this proposal is to investigate the evolution and emergence of pathogens from the marine environment by studying the interactions of the model pathogen pandemic *Vibrio parahaemolyticus* with the marine protist *Cafeteria roenbergensis* as a model for marine grazing pressure, and the starlet sea anemone *Nematostella vectensis* as an evolutionarily basal

host epithelium.

The specific objectives of this proposal will be to:
1. Develop a signature-tagged mutagenesis (STM) system for *Vibrio parahaemolyticus* using mariner-transposition and a novel PCR-tRFLP-based strategy for signature tracking.
2. Generate a library of tagged transposon mutants of *Vibrio parahaemolyticus.*
3. Characterize *Nematostella vectensis* as a model for microbe-epithelial interactions.
4. Screen a subset of the tagged *Vibrio parahaemolyticus* mutants for reduced fitness under grazing by *Cafeteria roenbergensis* and during association with *Nematostella vectensis* to demonstrate proof-of concept for use of STM to investigate the genetic mechanisms of environmental persistence and marine epithelial interactions and to generate preliminary results for future proposals.

BMAA, a cyanobacterial neurotoxin, in marine food webs: a pilot project
(Carl Lamborg, Mak Saito, Paul Drevnick; WHOI, Dept. of Marine Chemistry and Geochemistry)
ß-methylamino-L-alanine (BMAA) is a neurotoxic amino acid produced by cyanobacteria. High concentrations of BMAA in human brain tissue have been linked to neurodegenerative diseases (ALS, Alzheimer's, Parkinson's) in Guam and Canada. The source of BMAA in Guam is cyanobacteria in the roots of cycad plants and biomagnification through a unique food web. The source of BMAA in Canada is unknown. A recent study, however, reported that many marine cyanobacteria also produce BMAA. Cyanobacteria are ubiquitous in the ocean and especially abundant in coastal areas that have experienced harmful algal blooms, representing a potentially significant source of BMAA to marine food webs. Fish or shellfish that eat cyanobacteria or otherwise accumulate BMAA may thus pose a health risk to human consumers of seafood. We propose to address the most fundamental question concerning the distribution of BMAA in the temperate coastal ocean: Are BMAA concentrations in seafood high enough to be of concern for human health? We will examine fish and shellfish of commercial, recreational, and subsistence importance for BMAA concentrations. If we find BMAA concentrations that pose a human health risk, (i) this could form the basis of a human health risk assessment for BMAA and (ii) we will have preliminary data to generate a full proposal for further study.

c. Significance

The funded pilot projects address important questions regarding the interactions between the oceans and human health. Many of the projects represent new directions for the Center, expanding the scope of our research efforts into, for example, marine natural products (e.g. BMAA) and employing new technologies (e.g. signature tagged mutagenesis). Many of the funded or proposed pilot projects involve postdoctoral researchers or graduate students, and will thus contribute to the training of the next generation of researchers in this area.

d. Plans

During the next project period, we will continue the Pilot Project Program, issuing a new call for proposals early in the year. We also will follow-up the previously funded projects, to gauge their progress and success in achieving research goals and obtaining externally funded grants using preliminary data generated in the pilot projects.

e. See Publications List

f. Project-generated resources:

One objective of the pilot project program is to support the collection of preliminary data that can be used to generate full proposals to NSF, NIH, or other agencies. Proposals (funded and pending) arising from our pilot projects are listed below:

Approved for Funding:
NSF: S. Dyhrman, M. Saito - EN-GEN: Transcriptional and Proteomic Analyses of Multiple Environmental Stressors in Marine Diatoms (TP-AMES) $999,500

Florida Fish & Wildlife Conservation Commission. P. Hoagland and others. Proposal to extend the WH-COHH pilot project to the entire west coast of Florida. $84,975.

NSF: H. Kite-Powell: Supplemental COHH funding to study the economic losses associated with beach closures in the Myrtle Beach region due to contamination of marine waters ($20,000).

Pending:
EPA: S. Dyhrman - Linking Biogeochemistry to Harmful Algal Bloom Nutritional Physiology with Gene Expression Analysis: A Case Study with *Aureococcus anophagefferens* $497,821

NIH: PI Sandra McLellan and Co/I Mitchell Sogin. Microbial community profiling of sewage contamination in the Great Lakes: R21 grant submission.

Research and Education Activities, Findings, Training and Development, Outreach, and Products

Stegeman – Administrative Core

Research and Education Activities:

1. Foster the communication, planning, integration and interaction among Center members, thereby assuring that the goals developed in the Center vision are achieved.
2. Provide and maintain an administrative structure to oversee and monitor the financial aspects of the Center, including grants management.
3. Establish a structure that will support effective communication and planning with contribution from all Center Investigators.
4. Establish an External Advisory committee, which will review the Center programs and advise the Director and Deputy Director.
5. Oversee the operation of and work toward a successful Pilot Project small grant program.
6. Establish and maintain public and private web pages for the dissemination of information and for data transfer and communication.
7. Pursue interactions between the COHH and other Centers within the three Institutions, as well as with other COHH and EHS Centers, in order to expand the reach and impact of the center, and to expand the research base on issues relevant to the Center mission and themes.
8. Encourage the involvement of center investigators in community outreach and education efforts.

Findings:

1. As before, the Director and Deputy Director have continued to oversee the Center Office and the activities of each of the component projects and cores in the Center. We continue to build visibility in and impact through all three of the component institutions, the Woods Hole Oceanographic Institution (WHOI), the Marine Biological Laboratory (MBL) and the Massachusetts Institute of Technology (MIT).
2. The Center Director and staff (administrative professionals), and the respective Grants Management offices at the Woods Hole Oceanographic Institution and other institutions continue to monitor the accounts for each of the Research Projects, the Pilot Projects, and the Advanced Genomics Core.
3. Center Investigator meetings have been held monthly, on the 1st Friday of each month. In addition, the Director of the Administrator informs the Center members of activities and opportunities pertinent to Oceans and Human Health via frequent email. The meetings and notes address all aspects of the Center interactions and communications, internally as well as externally. The Director and Deputy worked to increase visibility of the Center.
4. The Administrative Core organized a day and a half retreat for the investigators in the WH-COHH in January 2008, at the nearby National Academy Study Center. The retreat focused overall progress in the research and possible new directions. There will be a similar meeting of all the Pilot Project awardees, to be held in February.
5. The Director met during the summer with the Chair of the External Advisory Committee (Dr. Michael Gallo, University of Medicine and Dentistry of New Jersey), to discuss the center's activities.

6. The Director and Deputy continued to work together with the director of the Pilot Project program (Dr. Mark Hahn), to coordinate the operation of the successful Pilot Project program. Four pilot projects were funded with 2006 funds, including one involving a joint project with the Miami Center for Oceans and Human Health. In excess of $40,000 was leveraged to augment these projects. The fourth call was issued in September of 2007 by email to all faculty and research staff at WHOI, MBL, and MIT. On November 1, 2007, nine proposals were received and subsequently were reviewed and scored by members of the Internal Advisory Committee and external reviewers, including several members of the other COHH Centers and other Ad Hoc reviewers. The proposals requested approximately $308,000 in total costs. All applicants were provided with written reviews. Three projects were selected for funding: Three were recommended for funding, two from WHOI and one from MIT. Additional funds in the program may be used to fund partially two other highly meritorious proposals, one from MIT and one from MBL. The distribution of awards to all three of the constituent institutions is a strength of the program.
7. The COHH website developed at WHOI has continued to serve the needs of the four primary OHH Centers, Woods Hole, Hawaii, Miami and Washington (http://www.whoi.edu/science/cohh/). COHH links continue to be added and an ftp site (internal) is in use.
8. Interactions with other groups continue to grow.
    a. Interactions between the WH-COHH and the other COHH. The Director and Deputy Director, with the help of the whole Woods Hole Center, hosted the third joint Center Directors' and Investigators meeting, in Woods Hole, in April, 2007. This meeting was highly successful and resulted in 5 manuscripts focused on progress in different areas of OHH, which have been completed and are under review for publication in the BiomedCentral journal *Environmental Health*, a journal with a very good impact factor. The meeting also resulted in a collation of outlines of several academic courses on Oceans and Human Health that have been developed over the past years by investigators in Woods Hole Center and in other OHH centers.
    b. We have continued to participate in monthly conference calls among the leadership of the four NSF-NIH Centers, and a second monthly call that includes the NOAA OHHI leaders, to discuss points in collaboration and interaction in all aspects of the Centers' activities.
    c. Interactions between the WH-COHH and other NIEHS Centers: We are continuing to explore opportunities for jointly sponsored enrichment activities, as an outgrowth of the Center Directors meeting in April. These interactions are being facilitated by the Supplemental funds obtained from NSF in 2007.

Outreach and Impact:

The Center has been involved in informing the wider scientific community. We also have continued to encourage the involvement of center investigators in community outreach and education efforts. There have been several major activities this year.

a. Drs. Stegeman and Gast made presentations to the Massachusetts Shellfish Officers Association quarterly meeting, describing COHH activities and discussing possible interactions.

b. Dr. Stegeman and Dr. Lora Fleming of the Miami COHH have worked to develop an exciting agenda for the new Gordon Research Conference supported meeting on Oceans and Human Health, scheduled for 2008. We have enlisted and confirmed participation of the full slate of speakers and discussions leaders and are seeking funding to support this exciting new conference.
c. Dr. Amaral-Zettler gave a lecture on Oceans and Human Health in a course for physicians and health care professionals, at Falmouth Hospital. The Course, entitled "Molecular Biology's role in Modern Medicine, was organized by Genomics Core Director Mitch Sogin.
d. The Administrative Core and the particularly the Director helped Center Investigators with the preparation of the SGER grant application submitted to the NSF.
e. The Center also sponsored several seminars by researchers from other institutions, and provided support fore a visit by Dr. Debashish Battacharya (University of Iowa) to the Center.

Training and Development:
The Center supported an undergraduate Summer Fellow in 2007, through the NSF-REU grant. The fellow was Lara Polansky, University of Miami, who worked with pilot project recipient Porter Hoagland on assessing the role of *K. brevis* blooms on emergency department respiratory diagnoses in Florida. Another student was supported through the MBL to work in Becky Gast's lab. Other students worked on COHH projects in the lab of Center Investigator Martin Polz at MIT. The Woods Hole Center also provided support to two students to attend the fall Harmful Algal Bloom conference in Woods Hole.

Significance:
As before, the Administrative Core has overseen growth of the Center and facilitated the activities of all of the units, through communication and support. The Center has had important successes in the research projects and the pilot project program. Inter-center communication is robust. Funds continue to be leveraged to support and expand the activities of the Center. Center activities include research that has had direct consequences for the public health. The Administrative Core is the focal point for all Center activities.

Plans:
As in the past, during the next project period we will continue to oversee the management of the grant and the Center, to foster intra- and inter-center communication and collaboration, to increase the activities in out-reach and enrichment, and to seek additional funding sources that might be leveraged to maintain a robust pilot project program. During the next year the WH-COHH also will continue recruiting undergraduate student researchers through the NSF-funded REU program. We will participate in the fourth COHH meeting in Hawaii, and work to the success of the new Gordon Conference on OHH.

Research and Education Activities, Findings, Training and Development, Outreach, and Products

Anderson – Project 1

Activities:
Work in Year 4 has focused on Specific Aims 3 and 5 of the project:
3) Characterize the relationships between toxicity, physiological variability and genotype in *Alexandrium* spp. from the Gulf of Maine; and
5) track changes in the genotypic diversity of *Alexandrium* populations through time throughout the Gulf of Maine.

Findings:
In this project year, we continued our studies of *Alexandrium* bloom populations to include a third year of data from summer of 2007. In 2005, we collected and analyzed samples from a massive toxic *Alexandrium* bloom that occurred that summer. Results of the microsatellite genotyping showed that the late-bloom and early-bloom populations were significantly different from each other (Fisher's combined test, $p<0.05$), indicating that the genetic composition of the bloom population changed on the order of about 3 weeks. The observed change could be due to natural succession of the bloom community, or by the addition of new genotypes from a separated population, or both.

These two hypotheses were addressed during fieldwork performed during 2006, when we collected samples from across the Gulf of Maine region during a second *Alexandrium* bloom (summer of 2006) as well as from a toxic bloom in an isolated embayment, Salt Pond, MA. In the wider Gulf of Maine, populations were collected from across the region including offshore on Georges Bank, and none were significantly different from one another. A comparison of the 2005 and 2006 bloom samples showed that, in general, populations from the Gulf of Maine blooms in the two different years were not genetically distinct. Populations collected from Salt Pond, however, were genetically distinct from those in the wider Gulf of Maine, and they changed over the course of the 3-week bloom. From these two years of data, it appears that overall genetic composition of *Alexandrium* blooms in the Gulf of Maine is not significantly different from year to year. Within a year, however, we did observe changes in bloom populations on the timescale of approximately one month. This could result from the natural progression or 'turnover' of genotypes during a bloom, or from the mixing of genetically distinct cells from other (unknown) sources. Results of the 2006 analysis of the Salt Pond bloom provides support for the former hypothesis, although the mixing of different source populations cannot be discounted.

We continued our population genetic analysis in 2007, sampling repeatedly in 9 locations across the Gulf of Maine. We are in the process of genotyping over 1000 isolates from 2007, including biweekly samples from the offshore Georges Bank region and the most northerly "upstream" population in the Bay of Fundy. Although *Alexandrium* cells were not as numerous in 2007 as they were in the previous years, the monitoring cruises did reveal high numbers of cells in the offshore region of Georges Bank. The discovery of a large offshore population of cells when inshore areas had much lower numbers could have significant implications for *Alexandrium*

bloom ecology and shellfish toxicity in the region. The Georges Bank region is the focus of activity for a separate project headed by PI Anderson. Supplementary funding from NSF for COHH activities will enable us to send personnel on these cruises to collect samples to continue our population genetic analysis, with the aim of understanding the connection between inshore and offshore *Alexandrium* populations.

In this project year we have also initiated studies of the link between toxicity and physiology in *Alexandrium*. This includes paired growth rate and toxicity measurements on a large number of cultured *Alexandrium* collected from different years. The first experiment used 45 cultures isolated during 2001, and determined growth rates and toxicity under constant conditions. A 1.75-fold difference in maximum growth rates was observed amongst the cultures, ranging from 0.41 to 0.72 divisions/day. A second growth rate experiment used a total of 44 clonal isolates grown at 15°C and 6°C at a light level of either ~400 and 100 $\mu E \cdot m^{-2} \cdot s^{-1}$. Growth rates were determined for all isolates, under each of the four conditions. Toxicity measurements are currently underway. This dataset of matching information on genotype, growth rate, and toxicity will be used to examine the relationships between these three characteristics of *Alexandrium* strains. Ultimately, the information will be used in a modeling effort by WH COHH PI McGillicuddy to examine the role of genetic heterogeneity in the abundance and overall toxicity of *Alexandrium* populations in the region.

During the course of our work, we have involved a number of graduate and undergraduate students, as listed above. These students have worked on several aspects of this project, including the cruise sample collection and isolation of the *Alexandrium* cultures for microsatellite analysis.

Outreach and Impact:

Presentations:

Anderson, D.M. Don't eat the clams: Managing the threat from the New England red tide. Frontiers in Environmental Sciences, Weekly Series on Emerging Issues in Environmental Health Sciences. (Invited presentation at NIEHS headquarters in the *Frontiers in Environmental Science* seminar series. The talk was also webcast, and can be viewed at: http://www.niehs.nih.gov/news/video/science/frontiers/)

Anderson, D.M., B.A. Keafer, K. Norton, D.J. McGillicuddy, R. He, C.H. Pilskaln, and d. Couture. *Alexandrium fundyense* cyst dynamics in the Gulf of Maine. Fourth Symposium on Harmful Algae in the U.S. October 2007, Woods Hole, MA.

Anderson, D.M., B.A. Keafer, D.J. McGillicuddy, and R. He. The 2005 and 2006 New England red tides: mechanisms, management challenges, and implications for future forecasting capabilities", Estuarine Research Foundation meeting, November 2007, Providence, RI.

Erdner, D.L., L.A.R. McCauley, K. Libera, and D.M. Anderson. 2007. Population genetics of toxic *Alexandrium* blooms in the Gulf of Maine. Fourth Symposium on Harmful Algae in the U.S. October 2007, Woods Hole, MA.

Erdner, D.L., L.A.R. McCauley, K. Libera, and D.M. Anderson. 2007. A real-time PCR assay for the toxic dinoflagellate *Alexandrium fundyense*: laboratory studies and field validation. American Society of Limnology and Oceanography international meeting, February 2007, Santa Fe, NM.

Posters:

Brosnahan, M.L., R.J. Olson, D.L. Erdner, and D.M. Anderson. Evidence of a self-recognition system in the sexual life-cycle of *Alexandrium tamarense*. Fourth Symposium on Harmful Algae in the U.S. October 2007, Woods Hole, MA.

Products:

Websites Created

The Northeast PSP site (http://www.whoi.edu/sbl/liteSite.do?litesiteid=3230&articleId=13371)

Research and Education Activities, Findings, Training and Development, Outreach, and Products

McGillicuddy – Project 2

Research and Education:

1. Formulate a suite of population dynamics models for the various genotypes of *A. fundyense*. Model formulation will be guided by existing observations, as well as laboratory experiments to be conducted in Project 1 (Anderson).
2. Incorporate the ensemble of population models into existing models of Gulf of Maine coastal hydrodynamics.
3. Use the coupled physical-biological models to construct hindcast simulations of *A. fundyense* survey observations to be collected jointly with Project 1 (Anderson).
4. Diagnose the simulations to determine the processes regulating the space/time expression of the different genotypes in terms of *A. fundyense* abundance.
5. Utilize toxigenicity data for the various genotypes (provided by Project 1 (Anderson)) together with the coupled physical-biological models (Aim 3) to make predictions of shellfish toxicity along the coast. Toxicity predictions will be tested with observations from ongoing shellfish monitoring programs.

Findings:

Our effort this year was focused on execution of two research cruises, voyages 435 and 437 of R/V *Endeavor*. The first survey (Figure 1, left panel) revealed surprisingly low cell concentrations in the Gulf of Maine, in contrast to the severe blooms that took place in 2005 (Anderson et al., 2005) and 2006. For example, we can compare the southwestern block of our survey plan, from Cape Cod to Cape Ann, with a nearly identical survey conducted on R/V *Tioga* during exactly the same time period in 2006 (Figure 2). Surface live counts in 2007 were all zeros except for one station in which a single cell was observed (corresponding to a concentration of 14 cells per liter). This constitutes a dramatic change from this same time period in 2006, when most of the area was covered by several hundred to several thousands of cells per liter. In interpreting this comparison, keep in mind that the 2007 observations were collected during and in the aftermath of significant wind forcing, resulting in surface mixed layers of up to 40m in some places. Therefore the possibility of subsurface populations cannot be discounted.

Another major finding of our 2007 field work was a large bloom of *A. fundyense* on Georges Bank (Figure 1, left panel). Highest abundances occurred along the Southern Flank, with peak concentrations just over 13,000 cells per liter. The population extended southwest of Georges Bank along the outer continental shelf, consistent with the southwestward exit pathway from the bank at the southern end of the Great South Channel.

Our second survey (EN437) revealed persistence of the bloom on Georges Bank (Figure 1, right panel). Surface live counts indicate that cell concentrations on Georges Bank dropped somewhat since EN435, although there are several stations with thousands of cells per liter and a peak concentration near 10,000 cells per liter. The bank-wide pattern also changed, insofar as the central crest of the bank became a local minimum in concentration. Cell concentrations on the

southern New England shelf rose from near zero during EN435 to several hundred cells per liter during EN437.

Whereas the EN435 survey revealed near absence of *A. fundyense* in the coastal Gulf of Maine, cell concentrations in that area during EN437 were fairly similar to what has been observed in the past this time of year (Figure 1, right panel). A key question is how did the system "catch up" to "normal" conditions given such a late start? This question is even more puzzling in light of the severe blooms that took place in 2005 and 2006.

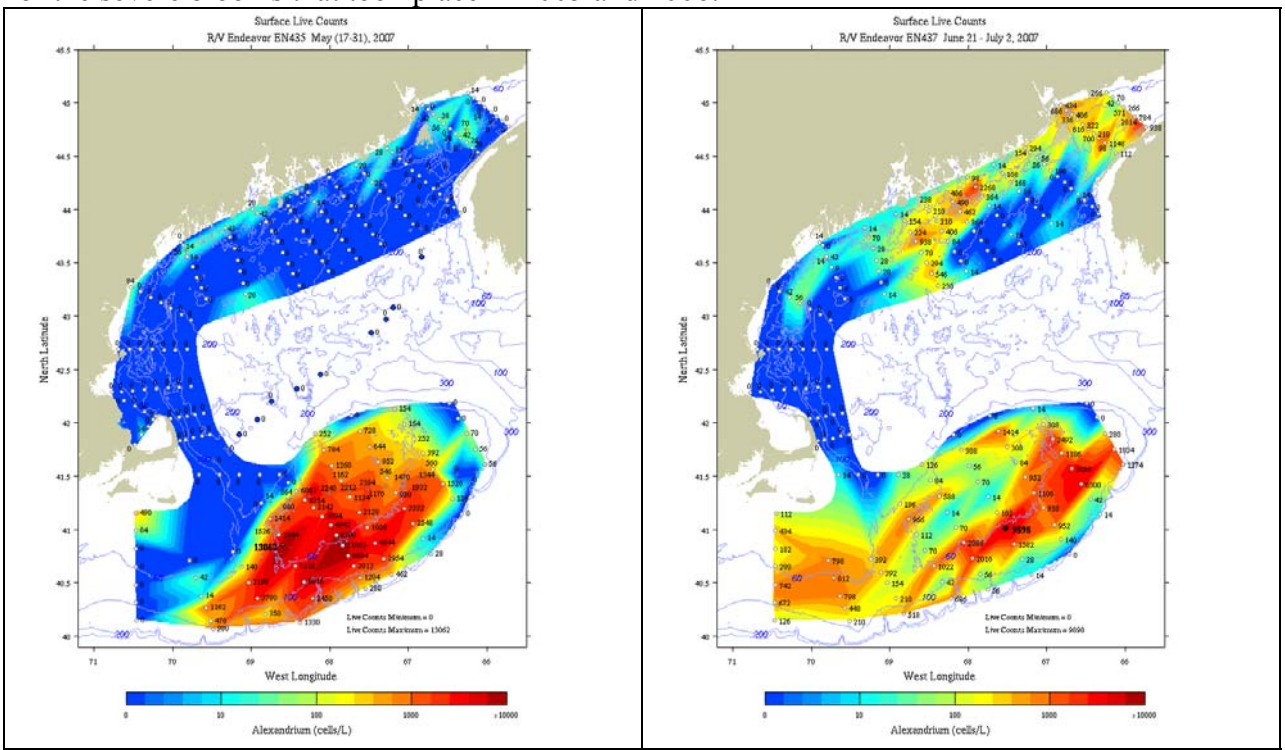


Figure 1. *Alexandrium* concentration (cells l^{-1}) from surface live counts on EN435 (left) and EN437 (right).

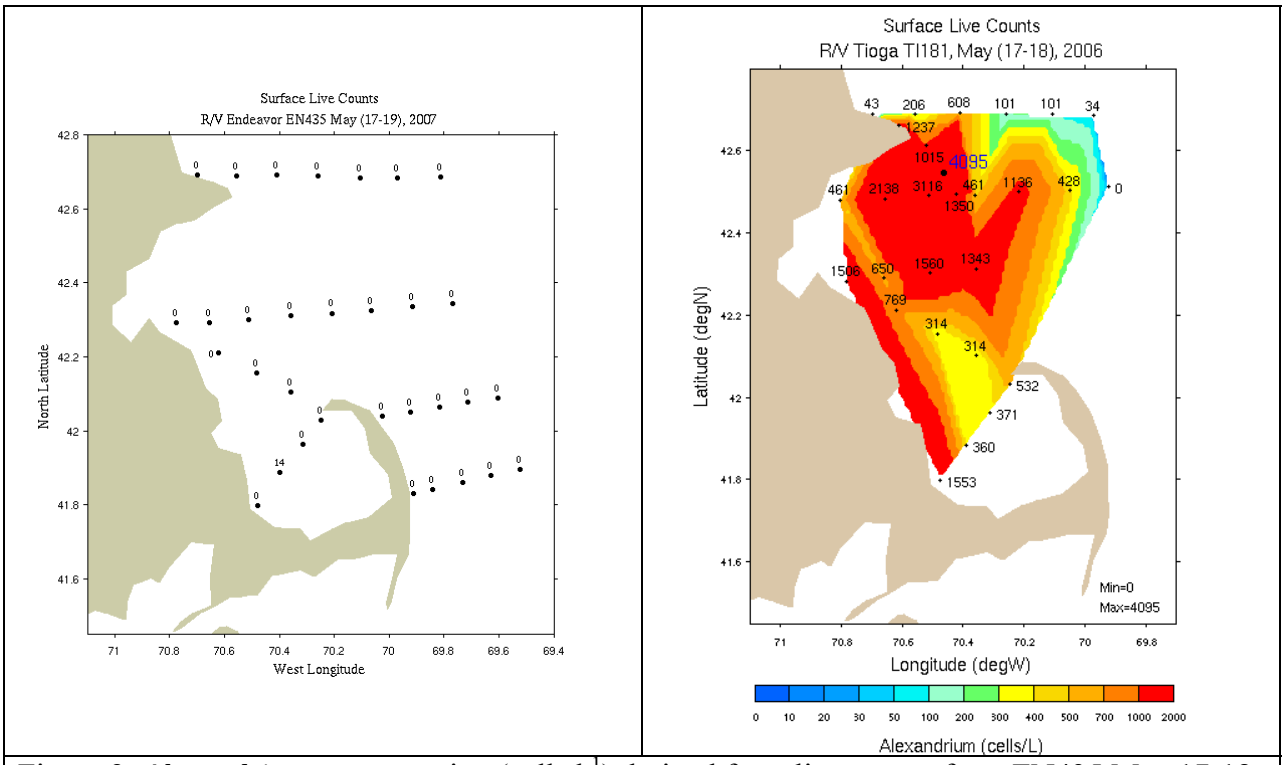


Figure 2: *Alexandrium* concentration (cells l^{-1}) derived from live counts from EN435 May 17-18 2007 (left) versus TI181 May 17-18 2006 (right).

Also in the past year, we developed a simple theory for the observed bimodal vertical distribution of motile phytoplankton (Ralston et al., 2007). Some motile phytoplankton have the capability to exploit deep sources of nutrients in a vertical migration cycle: photosynthesis in the near-surface layer, transit to depth, uptake of the limiting nutrient, and transit back to the surface layer. If all four steps can be completed within 24 hours, then migrations can be synchronized to the day/night cycle to maximize photosynthetic efficiency. Alternatively, if physiological, behavioral, or environmental factors make it impossible for the cycle to be completed in 24 hours, then migration may be asynchronous. Many observations of phytoplankton reveal bimodal vertical distributions of organisms, with maxima near the surface and the nutricline. We demonstrated how bimodal vertical distributions of phytoplankton may be symptomatic of asynchronous vertical migration using a Lagrangian Ensemble numerical model. We simulated vertical migration of the dinoflagellate *A. fundyense* in conditions similar to those in the Gulf of Maine, where bimodal distributions of *A. fundyense* have been observed. Migration is regulated by internal nutritional state – organisms swim down toward the nitracline when depleted of nitrogen, and return to the surface after nutrient uptake. We tested the sensitivity of the results to growth rate, nitrogen uptake rate, and swimming speed, and found that organism distributions can be bimodal or unimodal depending on conditions. Finally, we developed an analytical estimate for population distribution based on organism characteristics and nutricline depth.

Outreach and Impact:

10/06 Toxic algae in the Gulf of Maine: observations and models. WHOI Topics in Oceanography workshop for middle-school and high-school science teachers, WHOI Exhibit Center.

4/07 Modeling Blooms of *Alexandrium fundyense* in the Gulf of Maine: From Climatology to Forecasting. Environmental fluid mechanics seminar, Department of Civil and Environmental Engineering, MIT, Cambridge, MA (Invited).

4/07 Skill Assessment for Coupled Biological/Physical Models of Marine Systems. Ocean Color Research Team Meeting, Seattle, WA.

10/07 Observations and models of *Alexandrium fundyense* blooms in the Gulf of Maine and Georges Bank: From Climatology to Forecasting. Fourth Symposium on Harmful Algae in the U.S., Woods Hole, MA.

Products:

Websites Created:

1. Data from R/V *Oceanus* voyage 412, and R/V *Endeavor* cruises EN435 and EN437:
   http://science.whoi.edu/users/mcgillic/cohh/oc412/data/
   http://science.whoi.edu/users/mcgillic/en435/
   http://science.whoi.edu/users/mcgillic/en437/

2. Near-real-time nowcasting and forecasting of the 2005, 2006, and 2007 blooms:
   http://science.whoi.edu/users/ruoying/Redtide_05/movie.html
   http://omgrhe.meas.ncsu.edu/Redtide/Redtide_06/
   http://omgrhe.meas.ncsu.edu/Redtide/Redtide_07/

Research and Education Activities, Findings, Training and Development, Outreach, and Products

Gast – Project 3

Research and Education Activities:

This proposal examines the diversity, distribution and persistence of human pathogens in the coastal marine environment of Mt. Hope Bay, Massachusetts. This estuarine environment has been heavily impacted by human activity, including significant thermal pollution, several sewage outfalls and the effects of past industrial activities. Monitoring for bacterial contamination occurs in the shellfish areas, but outside of these regions, very little is known about the occurrence of either bacterial pathogens or protistan parasites that are likely introduced into the Bay through sewage or runoff. The potential impact of the thermal output of the Brayton Point Power Plant on these microbial populations has not been examined at all. We hypothesize that both bacterial and protistan human pathogens are present in Mt. Hope Bay and nearby rivers, and that their distribution and persistence are affected by physico-chemical condition in the Bay. To examine this hypothesis we propose to accomplish the following specific aims:

Aim 1: Determine the distribution and persistence of human protistan pathogens in Mt. Hope Bay. We will conduct a molecular survey of human pathogens within the Mt. Hope Bay estuarine system, including sewage outfalls near and removed from thermal point sources, contaminated shellfish beds and shorebird nesting grounds, to better characterize the types of protistan pathogens present in this coastal environment. Our project will specifically target *Giardia*, *Cryptosporidium*, *Naegleria*, and *Acanthamoeba*, but will also perform general microbial eukaryotic and prokaryotic diversity assessments to determine whether there are novel sequence types present that are related to known pathogens (might represent unidentified human pathogens).

Aim 2: Determine whether naturally occurring marine amoebae can serve as reservoirs for pathogenic bacteria, and look for evidence of the natural occurrence of these associations in Mt. Hope Bay. We will test whether locally collected marine amoebae are capable of harboring the pathogenic bacteria *Vibrio parahaemolyticus*, *Vibrio vulnificus*, and *Legionella pneumophila*. The research will test for the induced presence and replication of bacteria in actively growing and dormant phases of the amoebae, and will utilize both PCR-based and in situ hybridization based methods to detect naturally occurring associations.

Aim 3: Establish physico-chemical parameters associated with pathogen presence in Mt. Hope Bay and test predictive capability of water quality models regarding these correlations.
We will correlate measurements of water and sediment physico-chemical parameters with microbial community structure, including the presence of human pathogens. These measurements and our biological data will be used to assist in establishing and validating a water quality model for Mount Hope Bay that may also be used to understand and predict the potential presence of pathogens.

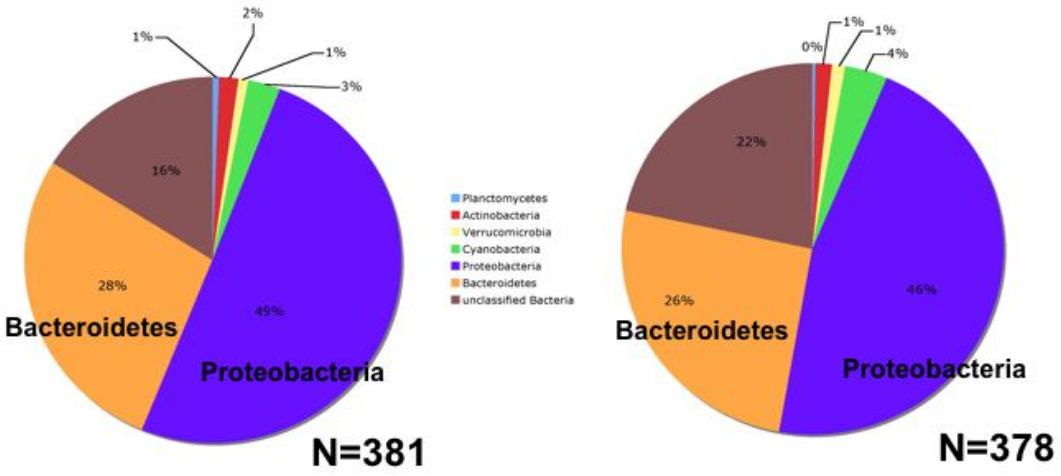
Findings:

Fieldwork. Four sites in Mt. Hope Bay were selected for our survey of human parasites and the effect of physico-chemical parameters on their distribution and persistence. One is located within the thermal plume of the Brayton Point power plant outlet. The second site (Braga Bridge) is located at an underwater sewage outfall that releases secondarily treated sewage into the Taunton River just a bit north of the Bay proper. A third site is farther up the Taunton River (Dighton Bridge) at a brackish water/marsh site, and the fourth site (Common Fence) is near the bottom of the Bay at an area not directly impacted by the thermal plume or the sewage outfall. Sediment and water samples were collected in triplicate at all four sites. Our seasonal sampling officially began in November 2004, and continued through April, June, July, August, November 2005 and February 2006. An additional sampling was carried out in June 2006 in response to large amounts of precipitation and increased runoff. In addition to our four sites we also conducted two transect sampling trips in August of 2005 and February of 2006. Each time our transect began at our Brayton Point site and included 7 stations along a direct line to our control site Common Fence. Transect sampling included water samples (collected at the surface and at depth just above the sediment) and sediment samples. Sediment and water samples were again also obtained and processed from five sites in the Great Salt Lake, Utah.

Results:

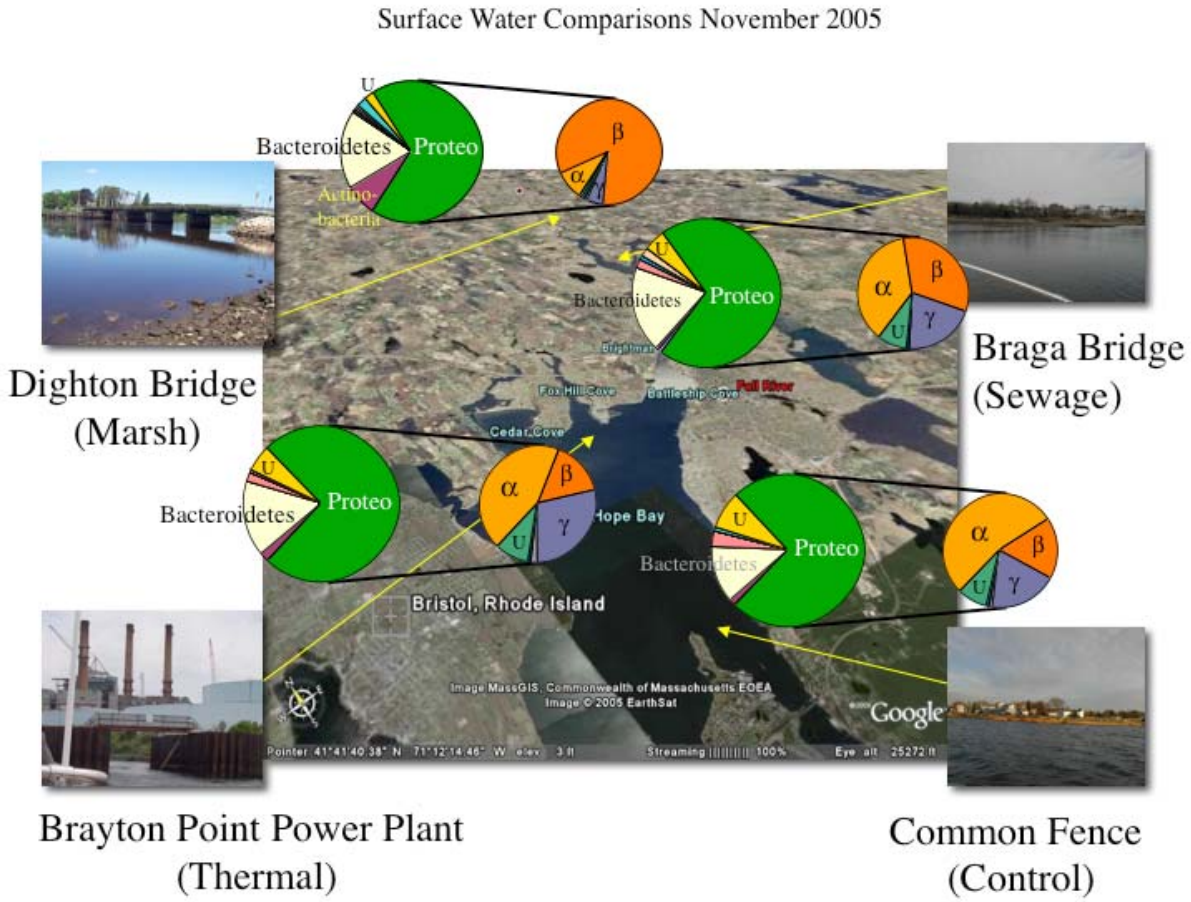
Microbial Diversity Surveys.

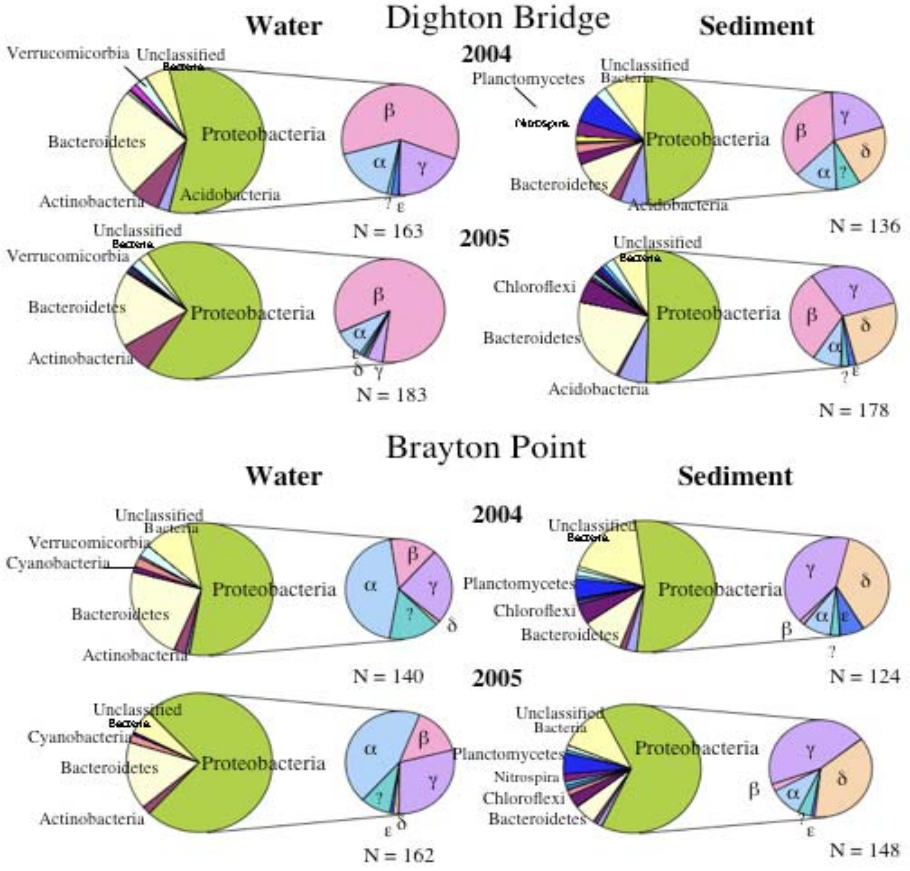
Separate Domain-specific Libraries. We have compiled the first comprehensive (eukaryal, bacterial, archaeal) data from small-subunit ribosomal RNA (SSU rRNA) gene clone libraries for water and sediment samples collected near the thermal plume and underlying sediments of the Brayton Point Power Plant (abbreviated BP in figures below) during November 2004. We have partial sequences of nearly 4,000 clones from 2 different sites and have further sequenced 1,000 unique clones from these to full-length. Not surprising, our findings reveal a highly diverse consortium of the three domains including relatives of sludge bacteria, polyaromatic hydrocarbon-degrading bacteria, and representatives related to the genera *Staphylococcus, Streptococcus, and Clostridium*. Comparisons between overall diversity in water samples for two replicate samples collected at the Brayton Point Power Plant site yielded very similar phylum-distributions. Our phylum-level comparisons between sediment samples, however, revealed significant differences in the microbial assemblages recovered. Preliminary phylogenetic placement of these environmental sequences revealed that the largest number of potential human pathogens were detected in sediment samples and not in the water-column samples. The two subsequent diagrams contrast water versus sediment replicates, asterisks indicate significant differences between groups based on the Ribosomal Database Project II Library Compare Program.



Serial Analysis of Ribosomal Sequences Tag (SARST) analyses: Our clone library constructs provide us with a reference library of full-length sequences for our Brayton Point Power Plant site. This serves as a baseline dataset with which to compare sequences obtained from higher-throughput methods that target short variable regions of small-subunit rRNA genes. Given that our initial phylogenetic analyses identified the largest number of potential human pathogens in sediment samples, we sought to further explore the microbial diversity in these samples using methods that would allow us to sample more deeply. One such method that we have begun applying to our sediment samples is called SARST (Serial Analysis of Ribosomal Sequence Tags). We currently have a total of 2,365 tags from the V6 region of SSU rRNA from one of our Brayton Point sediment samples. 513 of these possessed 100% matches to known organisms in GenBank. Of these we further identified potential pathogens belonging to the following bacterial genera that possess pathogenic representatives: *Shigella, Shewanella, Pseudomonas and Francisella*. The occurrence of *Francisella* in estuarine environments is not well-documented. As a result of this finding, we will begin to screen our samples for this pathogen in both environmental samples and amoeba cultures.

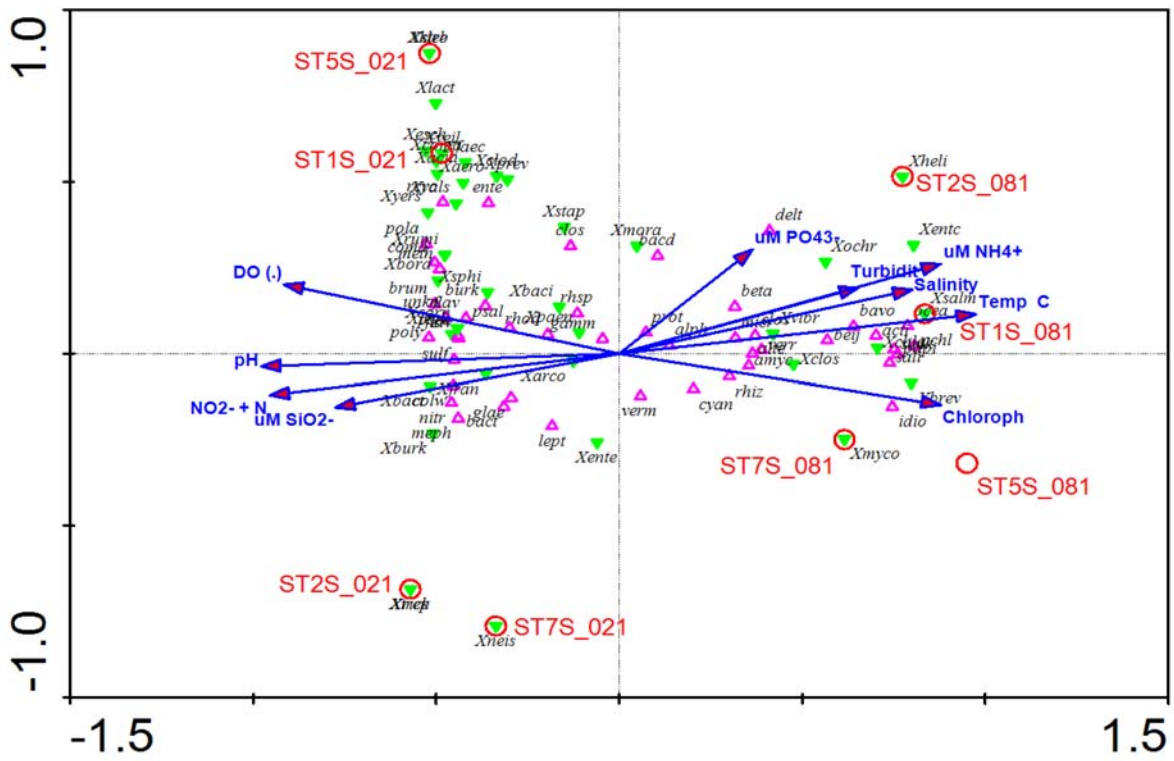
Comparison of 3-domain targeted Clone Libraries: To explore the seasonality and distribution of microbial populations and their corresponding potentially pathogenic members, we constructed clone libraries using primers that target all three domains of life simultaneously and capture the region in between the V4 through V8 hypervariable region of SSU rRNA genes. We constructed libraries from both water and sediment samples from November 2004 and November 2005 sampling seasons for the Brayton Point, Braga Bridge, Dighton Bridge and Common Fence locations. For each library we selected 192 clones for DNA sequencing in the forward and reverse orientations. Clone library composition was dominated by bacteria although some sediment samples yielded archaeal hits. The data have passed through initial quality control screening but still need to be screened for possible chimeric sequences. The pie charts below show preliminary data at the phylum level that were generated using the Ribosomal Database Project II Classifier program that assigns taxonomy to rRNA gene sequences as an example of the kind of data that are forthcoming from our analyses.





Seasonal Extreme Transect Study applying 454 Tag Sequencing:

Our final and most powerful approach employs a method developed by Mitchell Sogin and colleagues as part of the WHCOHH Genomics facility. This approach targets the same hypervariable region as the SARST-V6 method but circumvents the cloning step by proceeding directly to tag sequencing via pyrosequencing chemistry and 454 sequencing technology. As in SARST-V6, microbial tag taxonomic assignments are based on similarity to a reference SSU rRNA V6 library. Both SARST-V6 and 454 V6 tag sequencing allow for the recovery of both kinds of microbes present in a sample and their relative abundances. Our first run of the 454 machine has resulted in the recovery of 113,893 tags and included 4 summer and 4 winter surface water samples along a transect away from the power plant. The Canonical Correspondence Analysis triplot (CANOCO 4.5) below shows the relationship between Operational Taxonomic Units (OTUs = bacterial "species"), winter (STXS_021) and summer stations (STX_081) and environmental parameters. The plot includes the overall top 50 most abundant OTUs and those of any potential pathogens (indicated by an inverted green triangle and containing an "X" prefix). This analysis is preliminary but is compelling in that it reveals an abundance of potential pathogens falling in the area near Station 1 (the same as Bp 1), winter (ST1S_021). A second 454 run has just been completed with water column samples from the surface and at depth, as well as sediment samples. This second run was conducted with redesigned primers to target a broader range of bacteria. Analyses of these data are planned for the next several months.



Detection of specific pathogens. We continue in the process of setting up amplification arrays with our extracted samples for use with organism specific primers. These include primers for *Giardia, Cryptosporidium, Acanthamoeba, Naegleria* and *Legionella*. The *Legionella* primers were used in nested amplifications of our seawater, sediment and ameba culture samples. Sediment and ameba cultures from Mt. Hope Bay showed positive amplification results, and there appears to be a seasonal pattern to the prevalence of *Legionella*-like species in amebas. All of the water and sediment samples from the Great Salt Lake were positive using the *Legionella* amplification method. Both MHB and GSL samples are not often positive for *L. pneumophila*, but it has been detected. *Giardia* is also being detected in sediment and water samples of MHB, primarily in the summer months.

Legionella and amoeba sequences. Samples that were positive for the presence of legionellae were amplified, cloned and sequenced. In this work we have now newly described both the significant diversity of legionellae in the marine environment and have documented the seasonal pattern of their abundance. During the past year we have focussed on analyzing the *Legionella* sequence types present within amoeba cultures positive for the genus. Overall, 41% of MHB amoeba cultures were positive for the *Legionella* genus, while 57% of the GSL were positive. Of the positive cultures, 10% of the MHB amoeba cultures and 20% of the GSL amoeba cultures

were positive for *L. pneumophila*. It is of note that all of the amoebae that were positive for *L. pneumophila* were growing on either marine or brackish water media. This indicates that marine amoebae are very capable of supporting the growth of *Legionella pneumophila* in the marine environment. When we examined sequence types present in the non-pneumophila positive amoebae, we found that 31% were related to legionella-like amoebal pathogens, while 16% were related to other *Legionella* species that are also human pathogens. Seventy percent of the sequence types recovered from amoebae were also recovered directly from the environment, suggesting that they comprise a significant proportion of the *Legionella* population. We also investigated what types of amoebae were able to harbor *Legionella* species. Based upon small subunit ribosomal sequences, *Acanthamoeba* and *Hartmannella* were identified, along with several other genera of amoebae that have not previously been identified harboring these bacteria. This suggests that most species of amoebae have the potential to harbor bacterial pathogens in the natural environment (including the marine environment).

Doheny Beach/Avalon Beach Epidemiology Study. We have become involved in the Southern California Coastal Water Research Project (SCCWRP) beach epidemiology study. While the primary objective of the study is to correlate new indicator studies with human disease incidence, we requested to be included in the project to examine the incidence of *L. pneumophila* with respect to respiratory symptoms. We have also acquired beach sediment samples for analysis of the presence of human pathogens, including *Brucella* and *Campylobacter*. All of the water samples have been processed for Legionella detection, and those results have been submitted to SCCWRP for inclusion in the epi analysis. The sediment work was started in the fall by Elizabeth Halliday, a graduate student supported by the grant. She has developed quantitative PCR methodology that allows the identification of percent recovery and PCR inhibition so that more accurate values can be obtained, and has begun analyzing the beach sand samples from Avalon. We will also participate in the Miami COHH group beach epi study currently underway.

Significance. We have completed over a year's worth of sampling within the Mt. Hope Bay system and have been processing these samples for pathogen presence and microbial community structure. Although the detection of legionellae in marine environments was documented, the extensive presence of Legionella-like sequences in Mt Hope Bay was unexpected. This is also true for the Great Salt Lake samples. These organisms appear to be fairly prevalent in amoebae recovered from the sediments of these environments, and the number of sites that they were recovered from increased as the year progressed. This suggests that legionellae can persist and spread within the marine environment.

Our clone library analyses are still in progress, so we cannot make any conclusions at this time, however, the baseline data recovered from our study will provide a valuable comparative dataset for other studies of thermally and sewage impacted environments. Our microbial diversity data suggest that there are a diversity of potential human pathogens that can be detected using our employed approaches. Whether or not these potential pathogens can further become associated with free-living amoebae is a question we still need to explore. Of particular interest are

sequence tags recovered from both SARST and 454 tag sequencing that have a 100% match to sequences in the genus *Francisella*. *Francisella* is an obligate endosymbiont and *F. tularensis* a known human pathogen and causative agent of Tularemia. Tularemia is a localized problem on the nearby island of Martha's Vineyard, so ascertaining whether the human disease-causing strain is harbored in association with estuarine environments will be an important goal for the project over the next year.

Plans:
For the next 12 months we will be analyzing samples from the Doheny/Avalon study and summarizing the microbial diversity data from our combined approaches. *Giardia, Cryptosporidium, Naegleria and Acanthamoeba* amplifications will be completed for water, sediment and guano samples. We will also explore the relationship between our microbial diversity data and corresponding environmental data, and will begin working with the MHB FvCOM model to examine the impact of hydrodynamics on microbial populations and distribution. Manuscripts will be prepared that summarize the results of our Mt. Hope Bay study.

Training:
Support of a graduate student (Halliday), training of a postdoc (Del Castillo), mentoring of undergraduate students through REU programs, informal epidemiology training, and participation in SCCWRP California beach study and Miami beach study.

Gast has constructed and taught a graduate level course on Oceans and Human Health. Lectures are available to other COHH researchers.

Gast was a speaker at the WHOI HARP undergraduate career symposium "Global Environmental Challenges in Oceanography: Climate Change and Oceans and Human Health" June 27, 2007

Outreach and Impact:
Gast is a new Subject Editor (zoonotic diseases, protists) for Diseases of Aquatic Organisms, and is a co-editor for a special issue of Diseases of Aquatic Organisms on Marine Vertebrate Zoonoses, and will serve as a co-editor for a special issue on Marine Vertebrate Zoonoses.

Gast participated in a Massachusetts shellfish officers meeting with Stegeman during the summer to discuss potential interactions and impacts of COHH projects.

Products:
Websites Created:
1. COHH website (www.whoi.edu/science/cohh/whcohh/)
2. Gast homepage (www.whoi.edu/people/rgast/index.html)
3. Zoonotic disease website (www.whoi.edu/sites/zoonoses)

Research and Education Activities, Findings, Training and Development, and Outreach

Polz – Project 4

Research and Education Activities:
Our research pursues two overarching goals. The first is to explore environmental forcings on multiple spatial and temporal scales leading to abundance of pathogenic vibrios (Aim 1 and 2); the second is to understand the nature of genetic differences and gene flow between co-existing pathogenic and non-pathogenic variants of vibrios (Aim 3 and 4). All our aims are pursued within the Plum Island estuary and Parker River where we determine whether changes in estuarine physics, water chemistry and biology correlate with transitions in diversity and abundance of different *Vibrio* species and to determine estuarine conditions that might trigger the emergence of pathogenic strains of *Vibrio*. Overall, we are nearing completion of the tasks within AIM 1-3, and report progress on AIM 4 since last year's report.

Findings:

AIM 1 and 2. To characterize and model dynamics and reservoirs of *V. vulnificus* and *V. parahaemolyticus* populations over seasonal cycles, and to test the link between estuarine physics, nutrient and particle abundance and growth patterns of *Vibrio* species over tidal cycles
We summarize here progress on both aims. The sampling for both aims is complete and was spelled out in the last report. The analysis of population structure by culture-independent assays is currently in progress and will be finished by early 2008.

To determine ecologically coherent groups among co-existing vibrios in the water column, we examined the temporal and spatial distribution of *Vibrionaceae* genotypes. We sampled the water column to differentiate the free-living and attached (to both particles and zooplankton) compartments of the planktonic community under different macroecological conditions (spring and fall). Particle-attached and free-living cells were separated (in 4 replicates each) into a total of four consecutive size

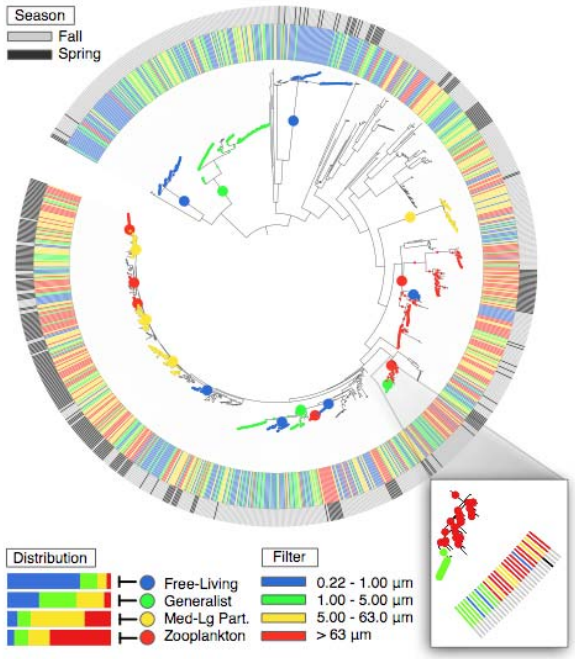


Fig. 1 Mapping genotypes across environments. Shown is a phylogeny of *Vibrio* isolates based on the *hsp60* gene. The inner ring depicts by color the size fraction from which each strain was isolated (see legend in figure), and the outer ring shows the season. The *AdaptML* software package was used to infer four commonly occurring distributions for phylogenetic groups (see legend in figure), and ecological populations demonstrating one of these characteristic distributions are indicated by colored dots (large dots indicate the ancestor of the group). The inset figure shows two closely reated groups (nested clades) that have distinct ecological distributions.

fractions, which are enriched in zooplankton (≥63 μm), large (63-5 μm) and small (5-1 μm) particles, and free-living cells (1-0.22 μm). The 5-1 μm size fraction is somewhat ambiguous, likely containing cells attached to small particles, as well as large or dividing cells; however, it provides a firm buffer between obviously particle-attached (>5 μm) and free-living (<1μm) cells.

Roughly 1,000 isolates were characterized by sequencing of a protein-coding gene (*hsp60*). To confirm relationships, between 1 and 3 additional gene fragments (mdh, adk and pgi) were sequenced for all *V. splendidus*, the dominant taxon during warm water conditions (Thompson et al., 2005b), and several other microdiverse groups. These data allow conservative estimation of ecological differentiation because inadvertent mixing of strains among microhabitats and homologous recombination among strains homogenize rather than create associations.

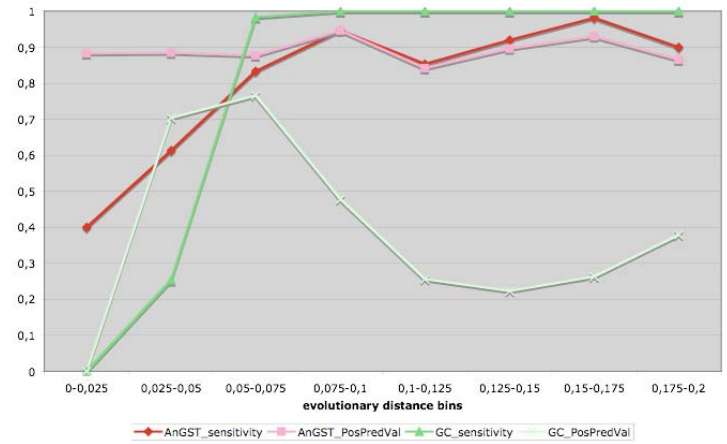


Fig 2. *AnGST-HR* performance on simulated vibrio data compared with pairwise approach. Simulated gene sequences were simulated according to an actual vibrio phylogeny inferred from 100 MLSA samples, and recombination events between leafs were randomly inserted into the tree among taxa at different evolutionary distances. *GENECONV* (GC) and *AnGST-HR* were evaluated based on their sensitivity (1- fraction of false negatives) and positive predictive value (PPV = fraction of true predictions that are correct). Both methods show a 'blind spot' for detecting recombination between very similar sequences, but *AnGST-HR* shows a high PPV across the entire range of distances, while *GENECONV* tends to make false positive predictions when highly divergent sequences are considered.

In collaboration with Eric Alm of the MIT Systems Microbiology group, we developed a software package (*AdaptML*) to identify associations between phylogenetically-defined groups and microhabitats that combines a 'model-free' empirical approach and an explicit evolutionary model of habitat association. Briefly our empirical framework establishes whether or not each clade in our strain phylogeny shows a non-random distribution across environments, while the explicit evolutionary model estimates population boundaries, rates of habitat 'switching', and the distribution of habitats across a set of measured environmental parameters.

Our data reveal important aspects of vibrio ecology in the water column. First, phylogenetic groups with distinct ecological preferences are clearly identifiable among all clades with sufficient representation in the dataset (Fig. 1). There is strong temporal structure evident where populations occurring in the spring and fall are frequently distinct at deep phylogenetic levels (Fig. 1). However, spatial partitioning is also evident for many groups, including *V. splendidus*, which contain many shallow clades with displaying distinct distributions (*V. splendidus* occupies the lower section in the tree spanning ~4-10 o'clock, Fig. 1). We were even able to identify nested clades where a single nucleotide change in several housekeeping genes is correlated to a switch in environmental distribution and thus suggests that this group is diversifying (one of these is highlighted in the inset to Fig. 1).

Second, although the currently crude spatial sampling (by size fraction) does not allow assignment of specific habitats (e.g. specific types of organic particles, zooplankton body sections), the analysis indicates differentiation into populations with broad ('generalist') and more narrow ('specialist') distributions, respectively. For example, *V. ordalii* was identified as specialized to free-living lifestyle, while the potentially pathogenic *V. parahaemolyticus* appears to be a generalist inhabiting particles, zooplankton and the free-living fraction of the water.

AIM 3 and 4. To determine the 'rules' of recombination, gene transfer and clonality in co-occurring *V. vulnificus* and *V. parahaemolyticus* and related vibrios, and to assess the diversity, mobility and evolutionary history of genes implicated in pathogenicity
Completion of this aim required the development of a new tool for recombination determination (Fig. 2). AnGST-HR is a modification of the AnGST software, which can identify gene dynamics within a comparative genomics framework. Fig. 2 shows that the method already performs better than other current methods in simulated datasets. The software is now being adapted to multi-locus sequence data sets to allow determination of recombination rates against sequence distance, which represents a critical parameter in judging the genetic isolation of different populations of microbes.

We have leveraged the work carried out under the auspices of the COHH to obtain funding from the Moore Foundation for sequencing of several of our genomes. We already have one *V. splendidus* and *V. alginolyticus* genome in hand. Technical developments now enable us to sequence 21 additional genomes by a combination of 454 and Solexa sequencing. This will provide an unprecedented dataset for both aims. Moreover, we are in the process of determining correlation of O-antigen diversity to population structure in the populations isolated from diverse habitats (water column, animal associated, etc.).

Significance:
The last few years have seen rising concern about the emergence of new variants of pathogens and spread of existing pathogens due to local or global environmental change. This has focused attention on the ecological context of pathogens in both the human body and the environment. Advances in population biology, aided by genomics, have demonstrated that many closely related (genomic) variants of microbial species exist in the environment. Furthermore, it has been shown that virulent bacteria frequently emerge from non-virulent strains via lateral gene transfer and it has been suggested that bacterial genomes are capable of extensive recombination. This raises the fundamental question to what extent observed genomic variants represent ecological and evolutionary units that can be seen as the bacterial equivalent to the eukaryotic sexual species. Do observed genomic variants occupy different environmental niches or do they represent a common gene pool capable of rapid 'assembly' of new variants in response to environmental challenges? These questions are crucial for interpretation of pathogen biology, risk assessment of emergence, and insights into how representative currently extensively studied strains (e.g., of *E. coli* or of pathogenic species) are for the 'species' they represent.

Products:
We are currently collaborating with colleagues in Austria and Germany to develop widely applicable microarray software based on our paper Marcelino et al. (2006).

Further, we have developed the AdaptML software, which can be used to detect ecologically cohesive bacterial populations among strains sampled from varied environmental samples.

Research and Education Activities, Findings, Training and Development, and Outreach

Sogin – Genomics Core

Research and Education Activities:

a. Facility operations and administration:

The genomics core provides DNA sequencing services and computational support to investigators in the Woods Hole Center for Oceans and Human Health (WH-COHH). Many of the Center's projects take advantage of our strengths in molecular microbial ecology and the bioinformatics. Over the past calendar year, the genomics core facility has provided WH-COHH projects with just over 31,000 reads corresponding to ~23 million base pairs. **Figure 1** Shows the usage pattern for the genomics core in 2007 and **Figure 2** describes the use of the Genome Core Facility by different WH-COHH projects.

Figure 1. WH-COHH sequencing usage pattern.

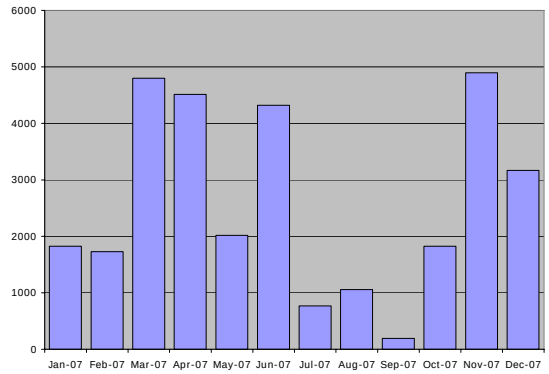
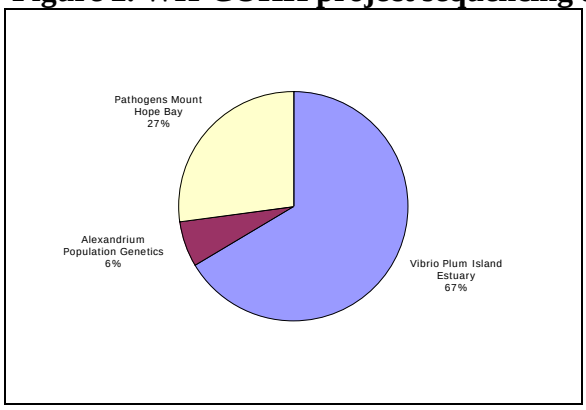


Figure 2. WH-COHH project sequencing activity.



Sequencing costs have changed little over the past year. Each template costs 30 cents using a protocol run on our BiomekFX platform. This is still a relatively inexpensive template production cost. In July, our Roche Genome Systems20 (GS20) pyrosequencing system (funded by a award titled ***Microbial population structure of the world's oceans*** from the W.M. Keck Foundation to the Marine Biological Laboratory at Woods Hole) was upgraded to a GS-FLX, which averages ~400,000 reads of ~238 nucleotides/read in each run. We anticipate that this

system will continue to be a key element of future COHH investigations. With this system, we have generated V6-tag datasets (as described below) for several of the MHB transect samples, providing very deep sampling to complement the traditional clone libraries outlined in the original proposal. We have also sequenced the genomes of six *Vibrio* phages in collaboration with Drs. Jonathan King and Peter Weigele (WH-COHH pilot project) using the GS-FLX.

b. Scientific and technology developments in the genome core.

Using the GS20 (and later, the GSFLX) sequencer the core facility continued development of a new, massively parallel DNA sequencing approach to characterize microbial populations with unprecedented level of sensitivity for detecting low abundance organisms. The International Census of Marine Microbes (ICOMM) provided the test samples and funds for reagents and the WH-COHH provided computational support for analysis of the data. We described in our last report the use of this approach to assess the species richness and eveness in water samples from the North Atlantic Deep Water flow and from Axial seamount (Sogin et al., 2006). During 2007, we demonstrated the feasibility of using 5-nt "keys" to multiplex reactions on the same region of a plate, deconvoluting the reads into sample bins informatically. This approach has been used on hundreds of samples from seawater, freshwater, sediment, soil, sewage, and human feces. We analyzed just over 1.3 million V6 tag sequences from human stool samples (500,000 from David Relman's laboratory and 800,000 generated in our laboratory as part of a collaboration with Jeffrey Gordon). This will provide a reference data for interpreting V6 tag sequences studies from human-impacted environments that we are studying in the context of a WH-COHH pilot project. Our data processing pipeline is well established, and datasets from tag and full length studies are displayed using a new web tool, Visualization of Microbial Population Structures (VAMPS), developed with support from the Sloan Foundation (awarded to M.L. Sogin and David Mark Welch).

Findings:

The ability to accurately characterize microbial population structures through the use of tag sequencing protocols is contingent upon the sequencing accuracy of this techonology. The published accuracy of 454 sequencing technology is only 96%. In genome projects, highly-redundant consensus assemblies can compensate for sequencing errors. In contrast, studies of microbial diversity that catalogue differences between PCR amplicons of ribosomal RNA genes (rDNA) or other conserved gene families cannot take advantage of consensus assemblies to detect and minimize incorrect base calls. To explore error modalities, we generated more than 340,000 reads from a PCR amplicon library that was prepared from a collection of 43 reference templates of known sequence. Each reference template contains a distinct rDNA including the V6 hypervariable region from a collection of 43 divergent bacteria. Differences between pyrosequences and their cognate reference sequences identified signatures of low quality data. We determined that 454 technology has a 99.5% accuracy rate in unassembled sequences, and we identified several factors that can be used to remove a small percentage of low-quality reads, improving the accuracy to 99.75% or better. Table 1 lists the source of those errors.

Table I.

Data Selection	Percent of Reads	Error Rate
All reads	100.0%	0.49%
Reads with no Ns	94.4%	0.24%
Reads with more than one N	5.7%	4.7%
Reads with length >=81 and <=108	98.8%	0.33%
Reads with length <81 or >108	1.2%	18.9%
Reads with no Ns and length >=81and <=108	93.3%	0.20%
Reads with no proximal errors	97.0%	0.45%
Reads with fewer than 3 proximal errors	>99.99%	0.48%
Reads with more than 3 proximal errors	<0.01%	12.2%
Reads with no Ns and length >=81and <=108 and no proximal errors	90.6%	0.16%

The simple removal of sequences that contain any unidentified ambiguities within the sequence of the target amplicon and elimination of sequences shorter than 50 bp or 20% longer than the target amplicon reduces the error rate to less than 0.25%. This is roughly ½ the error rate associated with capillary methods. This work appeared in Genome Biology (Huse, S.M., J.A. Huber , H.G. Morrison, M.L. Sogin and D. Mark Welch. Accuracy and quality of massively parallel DNA pyrosequencing. Genome Biology, 8: R143).

We have also demonstrated that the small amplicon size (~100 bp) more accurately captures the microbial diversity in a sample, most likely because small amplicons contain fewer regions of secondary structure that cause the polymerase to dissociate from the template. Results are shown in Table II. A manuscript describing this result is in progress (Huber, J.A., H.G. Morrison, S.M. Huse, P.R. Neal, M.L. Sogin, and D.B. Mark Welch. Effect of PCR amplicon size on assessments of microbial diversity and community structure, unpublished).

Table II.

Library	No. Clones	Forward Primer	Reverse Primer	% Exact Matches[a]	% Unique Sequences[b]
FS312_100bp	761	967F	1046R	44%	51%
FS312_400bp	860	967F	1391R	74%	22%
FS312_1000bp	381	337F	1391R	72%	12%
FS396_100bp	685	967F	1046R	11%	40%
FS396_400bp	866	967F	1391R	48%	22%
FS396_1000bp	677	337F	1391R	38%	19%

[a] Percent of V6 sequences within library that are exact matches to an existing sequence in the reference database

[b] Perecent of V6 sequences within library not detected by the other two libraries from the same sample

Training:

Through the genome core activities we continue to train a graduate student, Yuko Hasegawa in the use of advanced genomic techniques in the analysis of microbial communities associated with anthropogenic activities.

Outreach and Impact:

Drs. Sogin and Morrison have given lectures as part of a continuing education seminar series for Falmouth physicians and science teachers titled “Medical Genetics”. Analysis of microbial communities using the novel tag approach was included in the material presented.

Proposals Funded:
VAMPS – Visual analysis of Microbial Population Structures
Alfred P. Sloan Foundation. M.L. Sogin PI $300,000 July 1 2007- Jun 30, 2009.

Research and Education Activities, Findings, Training and Development, Outreach, and Products

Hahn – Pilot Project Program

Activities:

The specific objectives of the Pilot Project Program are:

A. To assess the feasibility of new areas of study, especially those that are not currently represented in our Research Project base but would contribute to the overall goals of the center.

B. To support the collection of preliminary data that can be used to generate full proposals to NSF, NIH, or other agencies or organizations.

C. To recruit scientists not currently involved in research on oceans and human health, but who may have expertise in one of these areas, to become participating members of the center and interact with other center investigators.

D. To stimulate collaborative and interdisciplinary research within the center.

E. To foster the application of new technologies and experimental approaches to questions concerning the impact of oceanic processes on public health.

F. To ensure the ability of the center to respond rapidly to new scientific information and emerging challenges in this field.

G. To contribute to the training of future researchers in the field of Oceans and Human Health through enhancement of graduate and postdoctoral training opportunities.

Findings:

Completed pilot projects (2004-2007) (list of publications and grants at end)

- *Characterization of a cyanobacterial anti-algal compound* (Eric Webb and Chris Reddy, WHOI).
- *Cnidarian toxins against voltage-gated Ca2+ channels* (Robert Greenberg, MBL).
- *Marine phage as vectors of gene transfer between marine bacteria and bacterial pathogens* (Peter Weigele and Jonathan King, MIT).
- *Transcriptome profiling in the harmful alga Aureococcus anophagefferens.* (Sonya Dyhrman, WHOI).
- *Beach Pathogens* (Steve Elgar, Britt Raubenheimer, & Rebecca Gast, WHOI)
- *Names-based cyberinformatics tools for rapid response communications and outreach during event management – a pilot based on harmful algal blooms in NE US coastal waters* (David J. Patterson, MBL, and Don Anderson, WHOI).

<u>Ongoing pilot projects (2006-2008)</u>

• *Anthropogenic impacts and profiling fecal microbial populations at a salt marsh* (Mitch Sogin, MBL)

Fecal coliforms are indicator organisms that warn of possible fecal contamination and its potential impact on human health. Coliform surveillance activities generally rely upon cultivation assays or molecular Microbial Source Tracking (MST) technology to assay for the presence of a particular indicator organism in environmental samples. This project will employ a

massively parallel DNA tag sequencing strategy to profile entire microbial communities in the Little Sippewissett Salt Marsh and the barrier Woodneck Beach. Possible coliform sources include human waste from failed septic systems, bird populations or other animals. The objective is to use the tag sequence data to locate specific sources of fecal contamination and identify suites of genes that could serve as multi-species indicators of human pollution. The experimental strategy takes advantage of rapidly evolving hypervariable regions in ribosomal RNAs and our ability to generate many thousands of short DNA tag sequences using 454 Life Science's pyrosequencing sequencing technology.

• *Transcriptional Markers of Life Cycle Transitions in Harmful Algal Blooms* (Don Anderson, WHOI)
Bloom dynamics of the red tide dinoflagellate, *Alexandrium fundyense* are driven in large part by transitions in its life cycle. While these stages are well documented, the biological and oceanographic forces that trigger transitions between the stages are not. A major obstacle to determining the conditions that trigger these transitions is our inability to rapidly identify sexual stage cells. Here, a novel transcriptome experiment will be used to discover molecules that are uniquely expressed by conjugating gamete cells and by germinating cysts. This experiment will utilize sequencing-by-synthesis technology that is newly available through the Genomics Core.

• *The Economic Effects of Harmful Algal Blooms: A pilot project to estimate the costs of human respiratory ailments associated with aerosolized brevetoxins* (Porter Hoagland and Di Jin, WHOI; Lora Fleming, Miami)
This study will estimate the costs-of-illness associated with human respiratory ailments that arise as the consequence of the aerosolization and coastal to inland transport of brevetoxins from blooms of the marine dinoflagellate, *Karenia brevis*, in the Gulf of Mexico. The research will develop models to link the occurrence of HAB events in the coastal-ocean with exposures to aerosolized brevetoxins. The researchers will compile datasets and develop models of illness rates that would permit historical estimates of these kinds of impacts and the simulation of future potential impacts.

• *The Economics of Human Health Risks from Pathogens and Toxins in the Marine Environment* (Hauke Kite-Powell and Porter Hoagland, WHOI)
These researchers will produce an "order of magnitude" estimate of the annual human health cost imposed on residents of the United States by exposure to pathogens and toxins from the marine environment. The estimate will be derived from a review and synthesis of information in the existing literature on (1) the spatial and temporal prevalence of marine pathogens, (2) the pathways by which they affect humans and the potentially exposed populations, (3) the human health effects of exposure, and (4) the economic cost of resulting medical conditions. The estimate will inform future research on pathogens in the marine environment, allowing scholars and public officials to target pathogens and settings where improved scientific understanding is most likely to produce significant economic benefits, and setting the stage for focused economic analyses. (co-funded with the WHOI Marine Policy Center)

Fourth call for pilot project proposals.
The fourth call was issued in September of 2007 by email to all faculty and research staff at WHOI, MBL, and MIT. On November 1, 2007, nine proposals were received and subsequently

were reviewed and scored by members of the Internal Advisory Committee and external reviewers, including several members of the other COHH Centers. The proposals requested approximately $308,000 in total costs. All applicants were provided with written reviews. Three projects were selected for funding:

Hydrodynamics and Transport Pathways for Fecal Microbial Populations in a Salt Marsh and Barrier Beach System (David Ralston, WHOI, Dept of Applied Ocean Physics and Engineering)

We propose to develop and implement a numerical model of flow and transport in a salt marsh and barrier beach system on Cape Cod, MA. The project will focus on Little Sippewissett Marsh and Wood Neck Beach, a system that during summer months frequently experiences high fecal coliform concentrations that indicate potential for significant impacts on human health and lead to restrictions on recreational use. Potential sources of fecal contamination include failed septic systems, birds, or other animals in the marsh. Recent investigations have found that fecal coliform counts are highly variable temporally (tidally, with precipitation events, and seasonally) and spatially across the marsh and beach. A COHH pilot project currently underway has begun to quantify the spatially heterogeneity of microbial communities in the system, intending to distinguish among different sources of fecal contamination and to develop a broader suite of indicator organisms (Sogin, 2006). The project proposed here will work in conjunction with that microbial community mapping project to quantify transport pathways, residence times, and exchange rates in the marsh. Combining transport mechanisms with the spatially heterogeneous source terms will permit calculation of potential exposure and risk to human health associated with the disparate sources of contamination. The model results will also provide guidance for public health officials to redesign of monitoring efforts to sample at times and locations of maximum potential exposure to elevated coliform concentrations, thereby minimizing human health risks. The tasks in the project include acquiring bathymetric data from existing sources and from new field surveys, constructing a numerical grid from the bathymetry, establishing and collecting data for model boundary conditions, and initial calibration and testing of the model based on available observations in the marsh. The model is intended to serve as a basis for future interdisciplinary studies at the study site – model results will aid in design of field observations, and more extensive field studies will aid in refinement of the model. The salt marsh and barrier beach of the study site represent a common coastal environment, and the impacts on human health of fecal contamination in coastal settings are of concern regionally and nationally. The pilot project will provide a step toward building an integrated program in Little Sippewissett to study linkages among physical, chemical, and biological processes in salt marshes and very shallow estuaries, coastal environments where frequent human interactions make potential health impacts particularly significant.

Using signature tagged mutagenesis (STM) to investigate how pandemic Vibrio parahaemolyticus persists in the bacterioplankton and associates with epithelia in the marine environment (Janelle Thompson, MIT, Dept of Civil and Environmental Engineering)
The overarching goal of this proposal is to investigate the evolution and emergence of pathogens from the marine environment by studying the interactions of the model pathogen pandemic *Vibrio parahaemolyticus* with the marine protist *Cafeteria roenbergensis* as a model for marine grazing pressure, and the starlet sea anemone *Nematostella vectensis* as an evolutionarily basal

host epithelium.

The specific objectives of this proposal will be to:
1. Develop a signature-tagged mutagenesis (STM) system for *Vibrio parahaemolyticus* using mariner-transposition and a novel PCR-tRFLP-based strategy for signature tracking.
2. Generate a library of tagged transposon mutants of *Vibrio parahaemolyticus.*
3. Characterize *Nematostella vectensis* as a model for microbe-epithelial interactions.
4. Screen a subset of the tagged *Vibrio parahaemolyticus* mutants for reduced fitness under grazing by *Cafeteria roenbergensis* and during association with *Nematostella vectensis* to demonstrate proof-of concept for use of STM to investigate the genetic mechanisms of environmental persistence and marine epithelial interactions and to generate preliminary results for future proposals.

BMAA, a cyanobacterial neurotoxin, in marine food webs: a pilot project
(Carl Lamborg, Mak Saito, Paul Drevnick; WHOI, Dept. of Marine Chemistry and Geochemistry)
ß-methylamino-L-alanine (BMAA) is a neurotoxic amino acid produced by cyanobacteria. High concentrations of BMAA in human brain tissue have been linked to neurodegenerative diseases (ALS, Alzheimer's, Parkinson's) in Guam and Canada. The source of BMAA in Guam is cyanobacteria in the roots of cycad plants and biomagnification through a unique food web. The source of BMAA in Canada is unknown. A recent study, however, reported that many marine cyanobacteria also produce BMAA. Cyanobacteria are ubiquitous in the ocean and especially abundant in coastal areas that have experienced harmful algal blooms, representing a potentially significant source of BMAA to marine food webs. Fish or shellfish that eat cyanobacteria or otherwise accumulate BMAA may thus pose a health risk to human consumers of seafood. We propose to address the most fundamental question concerning the distribution of BMAA in the temperate coastal ocean: Are BMAA concentrations in seafood high enough to be of concern for human health? We will examine fish and shellfish of commercial, recreational, and subsistence importance for BMAA concentrations. If we find BMAA concentrations that pose a human health risk, (i) this could form the basis of a human health risk assessment for BMAA and (ii) we will have preliminary data to generate a full proposal for further study.

c. Significance

The funded pilot projects address important questions regarding the interactions between the oceans and human health. Many of the projects represent new directions for the Center, expanding the scope of our research efforts into, for example, marine natural products (e.g. BMAA) and employing new technologies (e.g. signature tagged mutagenesis). Many of the funded or proposed pilot projects involve postdoctoral researchers or graduate students, and will thus contribute to the training of the next generation of researchers in this area.

d. Plans

During the next project period, we will continue the Pilot Project Program, issuing a new call for proposals early in the year. We also will follow-up the previously funded projects, to gauge their progress and success in achieving research goals and obtaining externally funded grants using preliminary data generated in the pilot projects.

e. See Publications List

f. Project-generated resources:

One objective of the pilot project program is to support the collection of preliminary data that can be used to generate full proposals to NSF, NIH, or other agencies. Proposals (funded and pending) arising from our pilot projects are listed below:

Approved for Funding:
NSF: S. Dyhrman, M. Saito - EN-GEN: Transcriptional and Proteomic Analyses of Multiple Environmental Stressors in Marine Diatoms (TP-AMES) $999,500

Florida Fish & Wildlife Conservation Commission. P. Hoagland and others. Proposal to extend the WH-COHH pilot project to the entire west coast of Florida. $84,975.

NSF: H. Kite-Powell: Supplemental COHH funding to study the economic losses associated with beach closures in the Myrtle Beach region due to contamination of marine waters ($20,000).

Pending:
EPA: S. Dyhrman - Linking Biogeochemistry to Harmful Algal Bloom Nutritional Physiology with Gene Expression Analysis: A Case Study with *Aureococcus anophagefferens* $497,821

NIH: PI Sandra McLellan and Co/I Mitchell Sogin. Microbial community profiling of sewage contamination in the Great Lakes: R21 grant submission.