DEVELOPMENT OF TUTORIAL MATERIALS TO ENHANCE TRAINING IN PHYLOGENETIC IDENTIFICATIONS OF MICROORGANISMS
Amy Apprill, University of Hawaii at Manoa

Introduction
The study of microorganisms and their processes receives considerable attention from various fields including oceanography, limnology, engineering, geochemistry and geology – among others. Students enrolled in graduate programs within these fields are frequently required to obtain expertise outside their program or advisor’s laboratory to learn the tools and techniques necessary for microbial-based research. The Agouron Institute’s ‘Microbial Oceanography’ and the Marine Biological Laboratory’s ‘Microbial Diversity’ courses are excellent examples of training courses aimed at enhancing the education of interdisciplinary microbiologists. However, not all graduate students are able to commitment to the 6-week timeframe of these intensive training courses, and these courses are only able to train a limited number of students. The C-MORE educational program can help bridge this training gap by offering web-based materials, short tutorials and workshops designed to train the next generation of microbial oceanographers in cutting-edge tools and techniques for microbial oceanography research.

An example of the success of a short workshop-based training course was the microscopy workshop for phytoplankton and bacterioplankton identification held at the Hawaii Institute of Marine Biology and led by C-MORE post-doctoral scholar Dr. Rachel Foster and Drs. Michael Rappé, John Waterbury and Edward Carpenter during 2008. The course brought together students and educators for one week of lectures and practical laboratory tutorials. The participants gained an overview of microscopy methodology, received hands-on training in traditional and modern techniques, and acquired the confidence to explore new questions using their newly developed toolbox. At the end of the course, the students attested to the importance of this unique educational opportunity in a survey.

One particular area of training now necessitated by most microbial-based research is the analysis of nucleotide sequence data, and particularly sequences from the SSU ribosomal RNA gene. SSU rRNA data are by far the most common type of sequence data generated to explore the identity, function or ecology of microorganisms, and their cost-effectiveness now enables laboratories to generate hundreds to thousands of sequences per study. Methods of analyzing sequence data are not difficult; however, the range of tools and the learning curve required to usefully employ these tools can often be overwhelming to students. The purpose of this proposal is to develop a handbook of educational materials and a two-day tutorial workshop to train students in methods to analyze nucleotide sequence data. The handbook will focus on analysis of the SSU rRNA gene and will provide background information and computer tutorials for using the ARB software package. ARB is the most commonly used software for SSU rRNA gene analysis because it allows alignment of sequences with consideration of the secondary structure, permits users to personalize and build searchable databases, offers primer and probe design capabilities, and builds phylogenetic trees of sequence data using a variety of calculations. The workshops will provide an opportunity for students to gain hands-on training in SSU rRNA gene data analysis using ARB.
Objectives of this proposal

1. Develop a handbook and a practice data set to guide the novice user through analysis of SSU ribosomal RNA gene sequences from Bacteria and Archaea using the ARB software package.

2. Provide free access to the educational materials by making the handbook and practice dataset downloadable from the C-MORE educational webpage. Advertise the available handbook and practice dataset to interdisciplinary students, post-docs and young scientists.

3. Develop a two-day training workshop to train students in the analysis of SSU rRNA gene sequence data using the ARB software package. An initial one day training session will be held for the Agouron Institute’s ‘Microbial Oceanography’ course in Honolulu, HI, and additional 2 day training workshops will be held in both Honolulu, HI and Woods Hole, MA for locally based students.

Development of handbook and tutorial materials

In addition to providing some basic background information regarding phylogenetic theory and the use of rRNA as a gene marker, the handbook materials will guide the user step-by-step through the process of sequence analysis, creating phylogenetic trees, and primer and probe design using the ARB software. The handbook will include captured computer screen images which will visually guide the user through the different menu options of the software. The handbook materials related to use of the ARB software will be based on materials previously developed by Dr. Michael Rappé for a training course held at Cornell University and also materials from a course held at the University of Queensland.

The handbook will focus on the use of a practice dataset of 50 raw chromatograms of SSU rRNA gene sequences. The handbook will consist of several modules (or tutorials) that will instruct the user in the clean up and quality control of raw sequence chromatographs, assembling partial sequences into full length sequences, importing sequences into an existing SSU rRNA gene database using the ARB software package, aligning sequences, manual quality control of automatic alignments, creating phylogenetic comparisons using the different available methods, and designing probes and primers to target certain sequences. The SSU rRNA gene database offered for practice use will be a pared down version of the latest release of the publicly available SILVA SSU database. This will provide users with the most up to date phylogenetic comparisons but by using a smaller database the calculations involved will be processed quickly on a standard computer. An appendix of additional resources including available gene SSU rRNA gene databases (e.g., Greengenes, Silva, Ribosomal Database Project II) and web-based phylogenetic resources (e.g., CIPRES Project) will be assembled and listed in the handbook with short descriptions of their offerings.

The handbook will be used as the basis for the ARB workshops and will be overhauled based on suggestions generated from the workshops. After completion, the handbook and sample dataset will be uploaded to the C-MORE website and made freely available for downloading. The availability of the handbook will be advertised through postings to the ARB users list-serve, microbial-based electronic list-serves, and direct e-mails from A.
Apprill aimed at faculty poised in positions to broadly distribute the message about the available resources.

**ARB tutorial workshops**

Tutorial workshops will be developed and offered to students, post-docs and young investigators to provide hands-on training and discussion on the analysis of sequence data using the ARB software package. The workshops will be offered at three different times:

1) The first workshop will be offered as part of the Agouron Institute’s ‘Microbial Oceanography course on June 18, 2009 at the University of Hawaii at Manoa (UH-M). The workshop will be fully organized and taught by A. Apprill. This first workshop will draw upon the expertise of the experienced Agouron teaching staff for recommendations of tutorial and workshop modifications.

2) The second workshop will be offered to local students in the Providence/Boston/Woods Hole, MA area over 2 days at the Woods Hole Oceanographic Institution (WHOI) (tentative dates are July 30-31). C-MORE faculty Dr. Sonya Dyrhman has agreed to sponsor the workshop at WHOI. WHOI maintains 6 Macintosh teaching computers, and the course will be made available for at least 6 students and up to 4 more that have their own computer. The workshop will be fully organized and taught by A. Apprill. Additionally, WHOI/MIT senior-level graduate student and MBL ‘Microbial Diversity’ alum Erin Banning will assist with questions during the tutorials.

3) The third workshop will be offered to local students in the Oahu, HI area over 2 days at the UH-M (tentative dates are August 13-14). SOEST maintains 10 Macintosh teaching computers in the Meteorology Department which are available for use during the workshops. The workshop will be fully organized and taught by A. Apprill. Additional assistance will be provided by post-doctoral scholar Dr. Megan Huggett.

The workshops will each accommodate up to 10 students. The workshops will be based on the materials provided in the newly developed handbook, and students will have access to a computer with the latest ARB software. The workshops will provide snacks and lunch for the students during the two days of training (except for the Agouron course workshop). Student travel expenses will not be covered and therefore the workshops will be aimed primarily at students local to the workshop area. An informal polling of graduate students at the WHOI/MIT/MBL and UH-M revealed that 6-8 students and young investigators in each location are currently interested in the workshop. I anticipate that advertising the course widely among each institution via e-mail advertisements and bulletin board postings will bring in interest from additional students.

**Recruiting diverse students and underrepresented minorities**

The workshops will be broadly advertised to attract a diverse assemblage of students representing different academic departments. Within the University of Hawaii at Manoa, e-mails will be sent to students, post-docs and faculty in the Oceanography, Zoology,
Microbiology, Chemistry, Geology and Geophysics, Information and Computer Sciences, Molecular Biosciences and Biosystems Engineering, and research degree programs within the John A. Burns School of Medicine (including Cell and Molecular Biology, Epidemiology, and Tropical Medicine). The Woods Hole course will be advertised via e-mails to students, post-docs and faculty within WHOI, the Marine Biological Laboratory, select MIT programs (Biology, Civil and Environmental Engineering). In order to recruit Native Hawaiians and Pacific Islanders (NHPI) and other underrepresented minorities to participate in the workshops, I will advertise in ways to reach this audience including 1) emailing NHPI-serving organizations and members of the C-MORE diversity council, and asking them to forward to their mailing lists and/or post on their websites, 2) word of mouth advertising, through members of the NHPI community, and 3) posting an advertisement on national minority-frequented websites such as LSAMP, AGEP and IBP.

Preliminary workshop syllabus
Day 1:
9 – 9:30am: Welcome, Overview of workshop and brief introduction to the theory of phylogenetic analyses (lecture)
9:30 – 10:30am: Practical 1 – Clean-up, assemblage and quality control of sequence data (tutorial)
10:30-noon: Overview of SSU rRNA gene databases and introduction to the ARB software environment (lecture & tutorial)
Noon – 1pm: Lunch break
1 – 3:00pm: Practical 2 - Importing and aligning sequences into an ARB database (tutorial)
3 – 3:30: Afternoon break
3:30 – 4:30pm: Overview of phylogenetic comparisons and steps to creating a phylogenetic tree (lecture & tutorial)
4:30 – 6:00pm: Practical 3 – Creating a phylogenetic comparison of the newly imported and aligned sequences (tutorial)

Day 2:
9 – 10:00am: Recap of day 1, Introduction to web-based phylogenetic resources (lecture & tutorial)
10:00 –noon: Free time to complete Practical 3, question and answer session
Noon – 1pm: Lunch break
1 – 2:00pm: Primer and probe design theory (lecture)
2:00 – 4:00: Practical 4: Primer and probe design (tutorial)
4 – 4:30pm: Afternoon break
4:30 – 5:30pm: Brief student presentations about the phylogenetic trees and probes they designed, Discussion from participants
5:30 – 6pm: Overview of advanced features in ARB

At the completion of the workshop, the participants will be asked to complete a survey of the course and teaching staff. The results of the survey and the impressions of the teaching faculty will be compiled into a report to C-MORE assembled by A. Apprill. Following the workshops, the materials (syllabus, tutorial and Powerpoint presentation lectures) will be
uploaded to the C-MORE website and made freely available. The availability of these materials will be advertised similarly to the handbook advertisements, as described above.

**Feasibility and support**

This project has a high potential for success due to my commitment to building and distributing the proposed educational materials and holding the workshops. I attended an ARB tutorial as part of the 2005 ‘Microbial Diversity’ course offered by the Marine Biological Laboratory, and co-taught a similar tutorial with Dr. Thomas Schmidt (Michigan State University) during the 2008 course. Additionally, I have the support of several experienced faculty and post-doctoral students interested in teaching the workshops including Dr. Megan Huggett who recently attending a three-day ARB workshop at the University of Queensland. Importantly, I have spoken with 6-8 students each from Honolulu, HI and Woods Hole, MA interested in attending the workshops.

This project requires logistical support from the C-MORE educational and administration staff. Specifically, assistance and advice will be solicited in designing advertisements and the workshop website so they are inline with the C-MORE image. Additionally, assistance from willing C-MORE faculty and post-doctoral scholars is requested to deliver word of mouth information about the available educational materials to their respective institutions.

**Summary and timeline of expected deliverables from these activities**

1. A handbook, practice dataset and tutorial for sequence data analysis using ARB will be freely available for the public (available online by 10/1/2009).
2. Three workshops will be held offering specialized training in ARB for 30-35 students (completed by 9/1/2009).
3. The workshop syllabus, Powerpoint presentation lectures and tutorials will be made freely available for future formal and informal training in ARB (available online by 10/1/2009).
4. A report will be generated from the workshops and submitted to C-MORE. The report will summarize the events, provide recommendation for future workshops, and include comments from the participants (completed by 10/1/2009).
5. The students will be surveyed 6 months and 1 year following the workshops. These surveys and a report summarizing these surveys will be submitted to C-MORE (completed by 9/1/2010).

**Minimizing the carbon footprint of the proposed activities**

The activities proposed here come with a limited carbon footprint. Specifically, the educational materials will be freely offered on the internet in an electronic format. For the workshops, airplane travel is only required by A. Apprill because additional faculty and participants will be locally recruited, and students will be locally based. The workshops will utilize 100% recycled copy paper (Staples) for the printed handbooks and biodegradable plates and utensils (Styrophobia) products for the meals.
Budget and justification

<table>
<thead>
<tr>
<th>Item</th>
<th>Cost ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salary for A. Apprill (2 months based on $55,000 annual post-doc salary)</td>
<td>9,167</td>
</tr>
<tr>
<td>Travel expenses for A. Apprill (Woods Hole, MA to Honolulu, HI)</td>
<td></td>
</tr>
<tr>
<td>- Roundtrip airfare and baggage fee ($900)</td>
<td>1,150</td>
</tr>
<tr>
<td>- Reduced per diem (5 days at $50/day)</td>
<td></td>
</tr>
<tr>
<td>Coffee break and lunch for workshops (does not include Agouron course workshop)</td>
<td>960</td>
</tr>
<tr>
<td>- 12 people @ $20/day (2, 2 day workshops)</td>
<td></td>
</tr>
<tr>
<td>Copying and binding fees for handbooks for 3 workshops (40 handbooks @ $5/handbook)</td>
<td>200</td>
</tr>
<tr>
<td>TOTAL REQUESTED</td>
<td>$11,477</td>
</tr>
</tbody>
</table>

The budget includes two months of post-doctoral salary support for A. Apprill stretched over the months of June - August, 2009. As soon as funding for this proposal is acknowledged, the workshops will be scheduled and advertised and construction of the handbook and tutorial materials will begin. Follow-up activities for the workshops will extend into August, 2010, and tutorial materials will continuously be updated by A. Apprill as needed. A. Apprill will be based in Woods Hole, MA during July and August and travel support is requested for A. Apprill from Woods Hole, MA to Honolulu, HI to teach the Honolulu course in August (a reduced per diem is necessary to cover a rental car and dorm lodging at the Hawaii Institute of Marine Biology). The C-MORE review committee recommended that A. Apprill teach the course for Honolulu students prior to moving to Woods Hole, but she would prefer to teach it after having sufficient time to revise the workshop materials and become proficient at installing the software. There are no charges for classroom, internet or computer usage at UH-M or WHOI. Lunch and snacks are not requested for the Agouron course workshop.

Broader impacts of this proposed activity and relation to C-MORE objectives

The educational materials and ARB workshops proposed here will importantly provide advanced and practical skills for the next generation of microbial oceanographers and interdisciplinary microbiologists. The workshops provide tremendous opportunities for early career scholars A. Apprill, M. Huggett and E. Banning to gain experience educating graduate level and higher students. The coordination of the ARB workshops at UH-M and WHOI promotes interactions and the building of educational programs between these geographically separated C-MORE institutions. Lastly, the proposed activity will provide a proof of concept for future C-MORE ARB workshops and may lead to the development of a separate proposal to the National Science Foundation for future workshop funding.