



Evaluation Report

Fast, Free Phylogenies: HPC for Phylogenetics Tutorial

October 13-15, 2010

Pamela Bishop
Program Evaluation Coordinator
National Institute for Mathematical and Biological Synthesis
November, 2010

Table of Contents

Executive Summary	iii
Brief Synopsis of Event	iii
Evaluation Design	iii
Highlights of Results	v
Conclusions and Recommendations	vi
Background.....	1
Introduction	1
Participant Demographics	1
Evaluation Design	3
Evaluation Questions.....	3
Evaluation Procedures	4
Data Analysis	4
Evaluation Findings.....	5
Respondent Satisfaction.....	5
Pre-tutorial Webinar	5
Overall Satisfaction with Tutorial	5
Satisfaction with Accommodations.....	6
Tutorial Content and Format.....	7
Format	7
Content	7
Participant Knowledge.....	10
NIMBioS Cluster	11
Impact on Future Research	12
Suggestions for Future Tutorials.....	12
Conclusions and Recommendations	14
Appendix A	16
Appendix B	19
Appendix C	22

Table of tables

Table 1. Participant fields of study and areas of concentration	3
---	---

Table of figures

Figure 1. Racial composition of program respondents (n =39)	2
Figure 2. Status of respondents (n= 39)	2
Figure 3. Respondent satisfaction with various aspects of tutorial.....	6
Figure 4. Respondent satisfaction with accommodations	7
Figure 5. Amount of content offered	8
Figure 6. Most useful aspects of the tutorial	8
Figure 7. Participant knowledge gains.....	10
Figure 8. Impact of NIMBioS cluster on learning experience	11
Figure 9. Suggestions for future tutorials.....	12

Executive Summary

Brief Synopsis of Event

The “Fast, Free Phylogenies: HPC for Phylogenetics Tutorial” (Phylogenetics tutorial) took place at NIMBioS on the campus of the University of Tennessee October 13-15, 2010. The event was co-sponsored by NIMBioS, iPlant, and the National Institute for Computational Sciences. The tutorial focused on how to use TeraGrid, the CIPRES Portal, the iPlant Discovery environment, university clusters, and other typically free HPC resources for phylogenetic analysis. The event was geared primarily toward biologists (including students, postdocs and faculty) who were at least moderately experienced with phylogenetic analysis and who had datasets to run but were typically running analyses on their own desktops, though other researchers, such as statisticians or mathematicians working in phylogenetics, were encouraged to apply. Participants were encouraged to apply as teams (such as a pairing of a biologist and a statistician who collaborate in their work). Participants had the option of receiving instruction in basic Unix usage via webinar and self-paced online resources before the actual tutorial, so lack of experience in this area was not a barrier to participation. Co-organizers of the tutorial were Eric Carr (NIMBioS); Jim Ferguson (NICS UTK/ORNL); Susan Holmes (Statistics, Stanford); Brian O’Meara (Ecology & Evolutionary Biology, UTK); Sheldon McKay (iPlant); Alexis Stamatakis (Computer & Communication Sciences, Tech. Univ. of Munich); Dan Stanzione (Adv. Computing Ctr/iPlant, Texas); Robert Thomson (Center for Population Biology, UC Davis); and James Wilgenbusch (Scientific Computing, Florida State Univ.).

Evaluation Design

An electronic survey aligned to the following evaluation questions was designed by the NIMBioS Evaluation Coordinator with input from the NIMBioS Director and Deputy Director:

1. Were participants satisfied with the tutorial overall?
2. Did the tutorial meet participant expectations?
3. Was the tutorial appropriate to the participants’ levels of expertise?
4. Did participants feel they learned an appropriate amount of information?
5. Were participants satisfied with the amount of content and format of the tutorial?
6. Were participants satisfied with the accommodations offered by NIMBioS?
7. What changes in accommodations, group format, and/or content would participants like to see at future similar meetings?

The final instrument was hosted online via the University of Tennessee’s online survey host mrlInterview. Links to the survey were sent to 30 tutorial participants on October 18, 2010 (tutorial organizers and participants associated with NIMBioS were excluded from the evaluation). Reminder emails were sent to non-responding participants on October 25 and 28, 2010. By November 4, 2010, 29 participants had given their feedback, for a response rate of 97%.

An electronic demographic survey aligned to the reporting requirements of the National Science Foundation was designed by the NIMBioS Evaluation Coordinator with input from the NIMBioS Director. The final instrument was hosted online via the University of Tennessee's online survey host mrlInterview. Links to the survey were sent to the 35 tutorial participants for whom NIMBioS did not have complete information on September 22, 2010. Reminder emails were sent to non-responding participants on September 28 and October 5, 2010. By October 10, 2010, 35 participants had filled out the survey for a response rate of 100%. Demographic questions regarding gender, race, and ethnicity, and disability status were optional (disability status is not reported in this evaluation report). All demographic information is confidential, and results are reported only in the aggregate. When feasible, the evaluator filled in missing demographic data from other sources (e.g. address, institution, field of study). The evaluator did not assume race, ethnicity, or disability status for any participant who did not report this information.

Highlights of Results

- 79% of respondents attended the pre-tutorial webinar, 96% of whom indicated it was useful in preparing them for the tutorial.
- 100% of participants agreed they would be more likely to include high performance computing resources in their future research as a result of attending the tutorial.
- The majority of respondents agreed the tutorial met their expectations (83%) and was appropriate to their level of expertise (90%).
- 97% of respondents either agreed or strongly agreed that they would recommend participating in NIMBioS tutorials to their colleagues.
- 100% of respondents thought the instructors were very knowledgeable about their topics, and 93% thought the presentations were useful.
- The majority of participants thought the hands-on exercises and group discussions were useful (97%).
- 90% of respondents agreed that the format of the tutorial was very effective for achieving its goals, and 100% were satisfied with the opportunities provided during the tutorial to ask questions and/or make comments.
- The majority of respondents (59%) indicated they felt the amount of content offered during the tutorial was “just right.”
- Respondents reported the greatest knowledge gains regarding where and how to submit jobs to online resources, followed by how to navigate Unix, alternatives for job management, and how to submit jobs to a cluster.
- Overall, respondents reported being satisfied with the travel, housing, and other amenities provided by NIMBioS.

Conclusions and Recommendations

Overall satisfaction with the tutorial was high among respondents, the majority of whom indicated that the tutorial met their expectations and was appropriate to their level of expertise. Respondents were also satisfied with the travel, housing, and other amenities offered by NIMBioS.

The majority of respondents thought that the tutorial format was effective for achieving its goals and were satisfied with the opportunities provided during the tutorial to ask questions and/or make comments. Many respondents indicated the exposure to new information about the resources available for research was the most useful aspect of the tutorial, while others found learning to compile jobs on the cluster useful. The majority of respondents felt the amount of content covered during the tutorial was “just right.” When asked what topics they would have liked to have covered in this tutorial if given more time, the most common responses were related to making comparisons among the different software packages available for phylogenetic analyses. Several respondents indicated they would have liked to include more hands-on sessions, and also more information about R.

Respondents reported the greatest knowledge gains in learning where and how to submit jobs to online resources, followed by how to navigate Unix, alternatives for job management, and how to submit jobs to a cluster. Fewer agreed that they had a better understanding of how to program and how to compile. Almost all respondents indicated that using the NIMBioS cluster enhanced their learning experience during the tutorial.

Several suggestions were made for improving future tutorials, including providing more handouts to participants, allowing more time for running jobs on clusters, and longer participant introductions. Other suggestions included separating participants by experience level and having a longer tutorial.

Based on analysis of respondent response data, the recommendations for future tutorials are as follows:

- The pre-tutorial webinar was very well-received as an introduction to the tutorial material. Consider offering introductory webinars for future similar events.
- Consider offering more handouts to participants containing steps for the hands-on sessions, software used, and definitions of terms that may be new to many participants.
- Consider setting aside time at the beginning of tutorials to allow for longer participant introductions where participants can indicate their background and interest in the tutorial topic.
- For future tutorials on high-performance computing, consider a presentation comparing the various software packages (what each can and cannot do, availability, cost, etc.) available for analysis on the tutorial's topic.
- Participants repeatedly suggested allowing for more hands-on time. If feasible, work more of this type of activity into future tutorials.

HPC for Phylogenetics Tutorial Evaluation Report

Background

Introduction

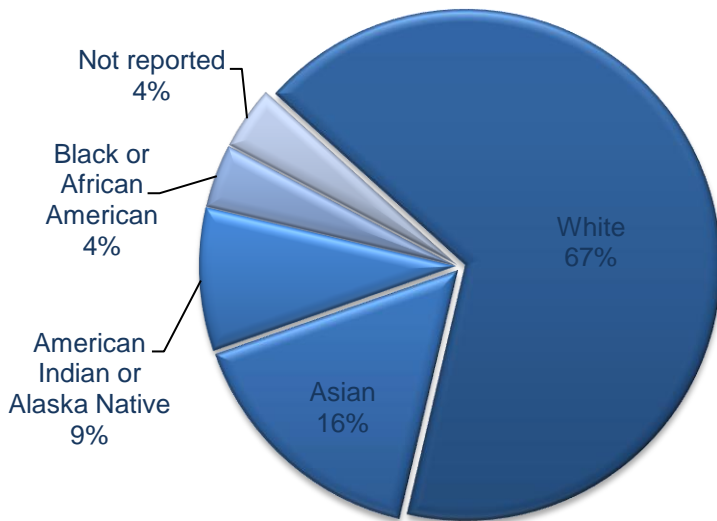
This report is an evaluation of a NIMBioS tutorial entitled “Fast, Free Phylogenies: HPC for Phylogenetics Tutorial” (Phylogenetics tutorial), which took place at NIMBioS October 13-15, 2010. NIMBioS tutorials are relatively large (30-40 participants), and serve as short introductions to particular research methods.

The event, which was attended by 39 participants, was co-sponsored by NIMBioS, iPlant, and the National Institute for Computational Sciences. The tutorial focused on how to use TeraGrid, the CIPRES Portal, the iPlant Discovery environment, university clusters, and other typically free HPC resources for phylogenetic analysis. The event was geared primarily toward biologists (including students, postdocs and faculty) who were at least moderately experienced with phylogenetic analysis and who had datasets to run but were typically running analyses on their own desktops, though other researchers, such as statisticians or mathematicians working in phylogenetics, were encouraged to apply. Participants were encouraged to apply as teams (such as a pairing of a biologist and a statistician who collaborate in their work). Participants had the option of receiving instruction in basic Unix usage via webinar and self-paced online resources before the actual workshop, so lack of experience in this area was not a barrier to participation. Co-organizers of the tutorial were Eric Carr (NIMBioS); Jim Ferguson (NICS UTK/ORNL); Susan Holmes (Statistics, Stanford); Brian O’Meara (Ecology & Evolutionary Biology, UTK); Sheldon McKay (iPlant); Alexis Stamatakis (Computer & Communication Sciences, Tech. Univ. of Munich); Dan Stanzione (Adv. Computing Ctr/iPlant, Texas); Robert Thomson (Center for Population Biology, UC Davis); and James Wilgenbusch (Scientific Computing, Florida State Univ.)

Participant Demographics

The 16 females and 23 males (four of whom self-identified as Hispanic/Latino) came from diverse racial backgrounds (Figure 1).

Figure 1. Racial composition of program respondents (n =39)



Participants in the tutorial were graduate students, college/university faculty and staff, postdoctoral researchers, and those from non-profit organizations(Figure 2).

Primary fields of study for the 39 participants included biological/biomedical sciences, computer and information sciences, engineering, mathematics, marine science, and social science (Table 1).

Figure 2. Status of respondents (n= 39)

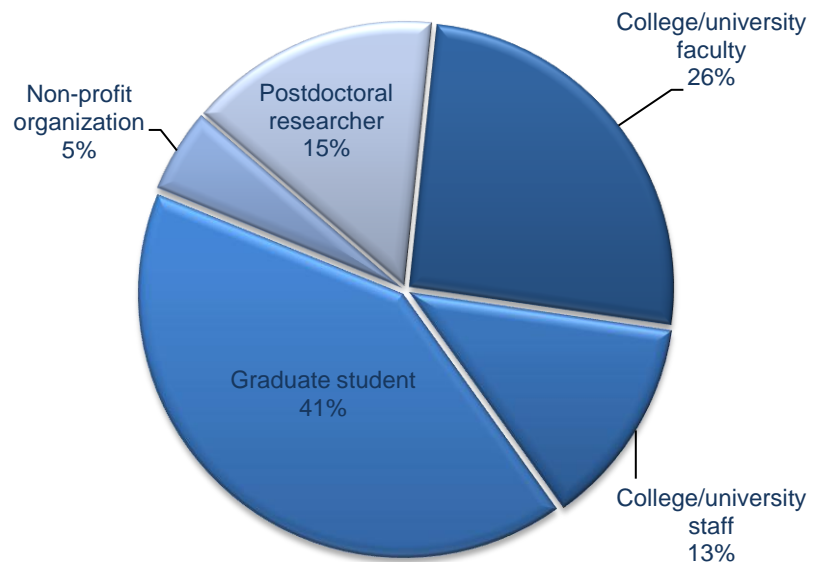


Table 1. Participant fields of study and areas of concentration

Field of Study	Concentration	# Participants
Biological/Biomedical Sciences	Biology/Biomedical Sciences, General	1
	Biomedical Sciences	1
	Botany/Plant Biology	1
	Entomology	2
	Evolutionary Biology	25
	Molecular Ecology	1
Computer & Information Sciences	Computer & Information Science, Other	1
	Computer Science	1
Engineering	Computer	1
Mathematics	Statistics	1
Ocean/Marine Sciences	Ocean/Marine, Other	1
Social Sciences	Economics	1
Not reported		1

Participants represented 34 unique institutions across five countries, including Columbia, Germany, the Netherlands, South Africa, and the United States. Within the U.S., 19 different states were represented. Included in the institutions were 32 unique colleges/universities and two non-profit organizations. Of the 32 colleges/universities, most were classified as comprehensive schools (having undergraduate and graduate programs).

Evaluation Design

Evaluation Questions

The evaluation of the tutorial was both formative and summative in nature, in that the data collected from respondents was intended to both gain feedback from respondents about the quality of the current tutorial and also to inform future similar events. The evaluation framework was guided by Kirkpatrick's Four Levels of Evaluation model for training and learning programs (Kirkpatrick, 1994¹). Several questions constituted the foundation for the evaluation:

¹ From Kirkpatrick, D.L. (1994). *Evaluating Training Programs: The Four Levels*. San Francisco, CA: Berrett-Koehler.

1. Were participants satisfied with the tutorial overall?
2. Did the tutorial meet participant expectations?
3. Was the tutorial appropriate to the participants' levels of expertise?
4. Did participants feel they learned an appropriate amount of information?
5. Were participants satisfied with the amount of content and format of the tutorial?
6. Were participants satisfied with the accommodations offered by NIMBioS?
7. What changes in accommodations, group format, and/or content would participants like to see at future similar meetings?

Evaluation Procedures

An electronic survey aligned to the evaluation questions was designed by the NIMBioS Evaluation Coordinator with input from the NIMBioS Director and Deputy Director. The final instruments were hosted online via the University of Tennessee's online survey host mrlInterview. Links to the survey were sent to 30 tutorial participants on October 18, 2010 (tutorial organizers and participants associated with NIMBioS were excluded from the evaluation). Reminder emails were sent to non-responding participants on October 25 and 28, 2010. By November 4, 2010, 29 participants had given their feedback, for a response rate of 97%.

An electronic demographic survey aligned to the reporting requirements of the National Science Foundation was designed by the NIMBioS Evaluation Coordinator with input from the NIMBioS Director. The final instrument was hosted online via the University of Tennessee's online survey host mrlInterview. Links to the survey were sent to the 35 tutorial participants for whom NIMBioS did not have complete information on September 22, 2010. Reminder emails were sent to non-responding participants on September 28 and October 5, 2010. By October 10, 2010, 35 participants had filled out the survey for a response rate of 100%. Demographic questions regarding gender, race, and ethnicity, and disability status were optional (disability status is not reported in this evaluation report). All demographic information is confidential, and results are reported only in the aggregate. When feasible, the evaluator filled in missing demographic data from other sources (e.g. address, institution, field of study). The evaluator did not assume race, ethnicity, or disability status for any participant who did not report this information.

Data Analysis

Data from the electronic surveys included both forced-response and supply-item questions. All data were downloaded from the online survey host into the statistical software package SPSS for analysis. Quantitative data were analyzed using SPSS, while qualitative data were analyzed in SPSS Text Analysis for Surveys. Qualitative responses were categorized by question and analyzed for trends.

Evaluation Findings

Respondent Satisfaction

Pre-tutorial Webinar

Seventy-nine percent of respondents attended the pre-tutorial webinar, 96% of whom indicated it was useful in preparing them for the tutorial. Some respondent comments about the webinar:

“I got a clear clue about what was going to happen during the tutorial. It's very useful.”

“I had only a very rudimentary acquaintance with UNIX language and this tutorial gave me a heads up on what kinds of research projects might require this material.”

“It was great--thanks for setting this up. This was a useful way to get lots of folks together in an easy-to-use interactive, online environment. “

“Brian is a great teacher, his patience helped me learn.”

Some respondents offered suggestions for improving future similar webinars, including having handouts available to participant before the webinar, and slowing the pace a little:

“I think written instructions should be provided before the webinar. It's too easy to fall behind when you have nothing to refer to for instructions.”

“It would have been helpful to have a pdf ahead of time of the commands we were working with.”

“Brian was going to fast during the webinar which made it difficult for the ones that had less experience writing commands.”

“Somewhat too fast to follow. I think a self-paced tutorial would work better.”

Overall Satisfaction with Tutorial

Overall satisfaction with the tutorial was high among respondents, the majority of whom indicated they either agreed or strongly agreed that the tutorial met their expectations (83%) and was appropriate to their level of expertise (90%). Some general respondent comments:

“I knew I needed HPC for my research, but before the tutorial I had no idea that there were so many ways it would be useful at many levels of my project. I'm hoping to convince other people in my field about how useful these skills and resources are.”

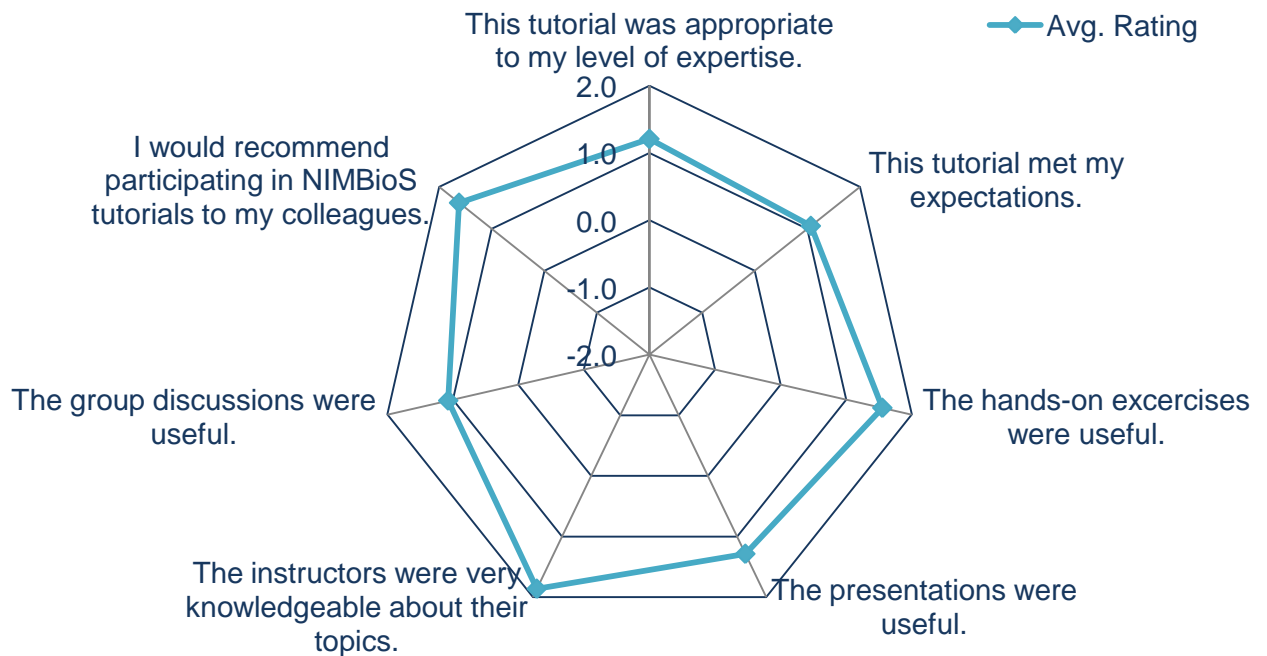
“This HPC tutorial opened up a new world for me. It will enhance my research program and that of my students and collaborators by several orders of magnitude. My lab will be learning Perl and "R". I now understand the relevance and importance of the these

tools. I am convinced that I must master the tools introduced in this tutorial if I am to remain competitive for certain kinds of funding.”

“Thank you for the opportunity to participate! I am more than ever motivated to learn more Unix and PERL and dive into larger data set analyses. I hope to get to the point where I can teach this to our undergraduate students (I know no one else of my colleagues will do). I think it would be great (possibly for a grant?) to be paired with a computer science/math faculty or computational biology faculty who could co-teach a course in bioinformatics/genomics/phylogenetics etc. Probably with rotating locations. Most small, primarily undergrad schools don't often have faculty in all areas of expertise, limiting collaborations of that kind.”

All respondents thought the instructors were very knowledgeable about their topics, and 93% thought the presentations were useful. The majority of respondents thought the hands-on exercises (97%) and group discussions (76%) were useful as well. Additionally, 97% of respondents either agreed or strongly agreed that they would recommend participating in NIMBioS tutorials to their colleagues (Figure 3, answered on a 5-point Likert scale from -2 to 2 for “strongly disagree” to “strongly agree”).

Figure 3. Respondent satisfaction with various aspects of tutorial

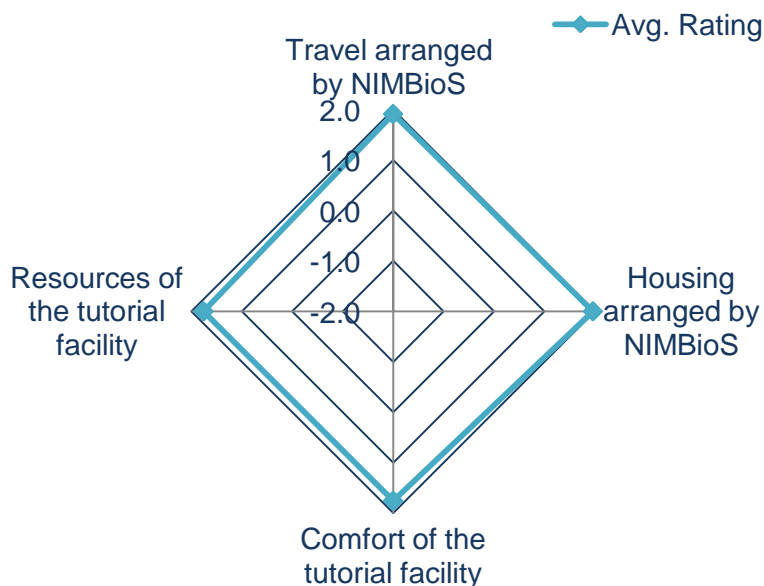


Satisfaction with Accommodations

Overall, respondents reported being satisfied with the travel, housing, and other accommodations provided by NIMBioS during the Workshop. Twenty-four respondents answered questions about their travel and housing, all of whom said they were satisfied with

their accommodations. (Figure 4, answered on a 5-point Likert scale from -2 to 2 for “strongly disagree” to “strongly agree”).

Figure 4. Respondent satisfaction with accommodations



Tutorial Content and Format

Format

The majority of respondents (90%) thought the tutorial format was effective for achieving its goals, and all were satisfied with the opportunities provided during the tutorial to ask questions and/or make comments. Common suggestions for improving the format were having handouts available during the tutorial, and having more detailed how-to's:

"[The tutorial format would have been more effective if] A [handout] was given to us. This would have helped me to follow the instructors in their demonstrations, and more importantly, I would be able to do it myself when get back to my university."

"For the most part, the format of the tutorial was fine. I would point out that many of the "hands on" sessions should have been - "watch this" as the abilities of the class differed so markedly and there was not the time for all to follow. What are needed are a better introduction and overview of the resources available and the software. One example would be 1) here is Condor. 2) This is how one utilizes Condor - i.e. you need a pool of computers 3) This is how you find resources for setting up Condor 4) these are the applications you can run in Condor."

Content

Fifty-nine percent of respondents felt the amount of content covered during the tutorial was “just right,” while the remainder felt there was either “too much” or “too little” content covered (Figure 5).

Qualitative analysis of responses revealed that most respondents felt the exposure to new information about the resources available for research was the most useful aspect of the tutorial (Figure 6):

[The most useful aspect was]

“The breadth of resources and tools that were covered. It whetted my appetite for learning Perl and R, for adding 'shared computer resources' to my current collaborations and proposals for funding. The instructors were top notch and just meeting them and picking their brains about computing topics was a wonderful resource.”

“It gave me a lot of information about phylogenetic tools I can use now to develop my research project.”

“It made me aware of my options.”

Figure 5. Amount of content offered

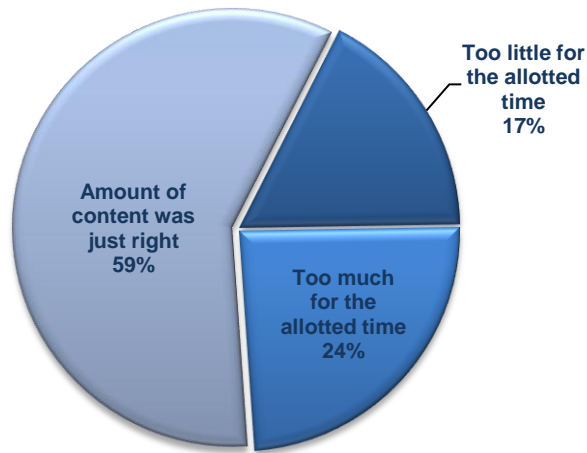
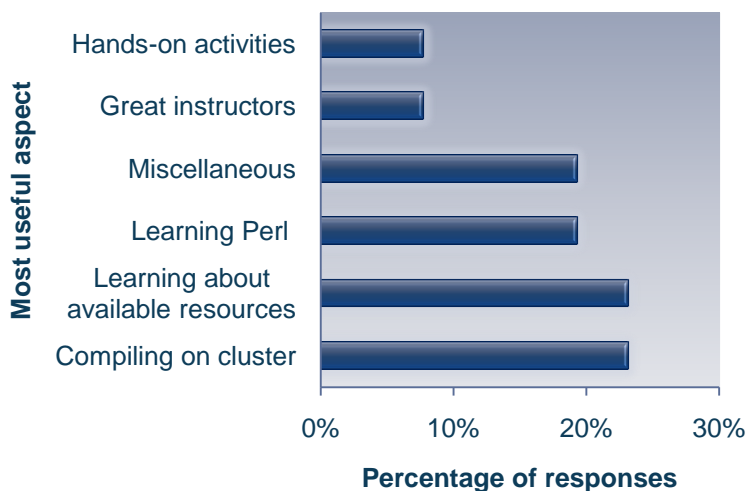


Figure 6. Most useful aspects of the tutorial



Many respondents indicated that learning to compile jobs on the cluster was the most useful aspect of the tutorial:

[The most useful aspect was]

“How to compile programs on a cluster and how to get a job started on a cluster.”

“How to deal with a cluster from compiling to submitting a job.”

The third most common response was learning Perl scripting:

“The Perl tutorial was particularly helpful, as was the introduction to the Condor, TeraGrid and the CIPRES portal.”

“The UNIX/PERL demonstrations and conversing with others about common problems.”

Other respondents felt the great instructors and hands-on activities were the most useful aspects:

“The fantastic instructors (both in the content of their lectures and in their willingness to wander around and help us individually), the diverse nature of the participants, and hands on opportunities.”

“The most useful aspect was to interact with the professionals who develop tools.”

“The most useful aspect was hands-on activities, e.g. Unix commands, submitting jobs to the cluster, etc. I appreciated that Brian slowed down and waited for everyone to get on the same page.”

When asked what topics they would have liked to have covered in this tutorial if given more time, the most common responses were suggestions related to making comparisons among the different software packages available for phylogenetic analyses:

“1) a more thorough discussion of the programs applicable to HPC resources 2) continuing with that line of thought - a better understanding of what resources are needed by these programs so that I may have an informed discussion with a cluster systems administrator.”

“Current state of various commonly used phylogenetics programs - how to configure each to run more parallelized.”

“It was a little too focused on introductory stuff like UNIX. It would have been better to have one model dataset set up, representing a typical phylogenetics problem which you might want HPC for, and then run that dataset through all of the various options we discussed. This would much more directly inspire future use...”

Other participants would have liked to have had more hands-on activities:

“...I wish we had a bit more hands on experiences and some more handouts with basic Unix, Perl (incl. reg. expressions), and python commands..”

“I would like to have had more details on how to run, get data and visualize trees in different hands-on session.”

“More practice about submitting projects on a cluster and a hands-on experience using the cloud in Amazon, for example..”

“The topics were perfect. However, it would have been great if there were a couple more days to do more hands-on exercises.”

Still others would have liked more information on R:

“A bit more information on R would have been useful.”

“...An analysis using R and ape/picante using a cluster...”

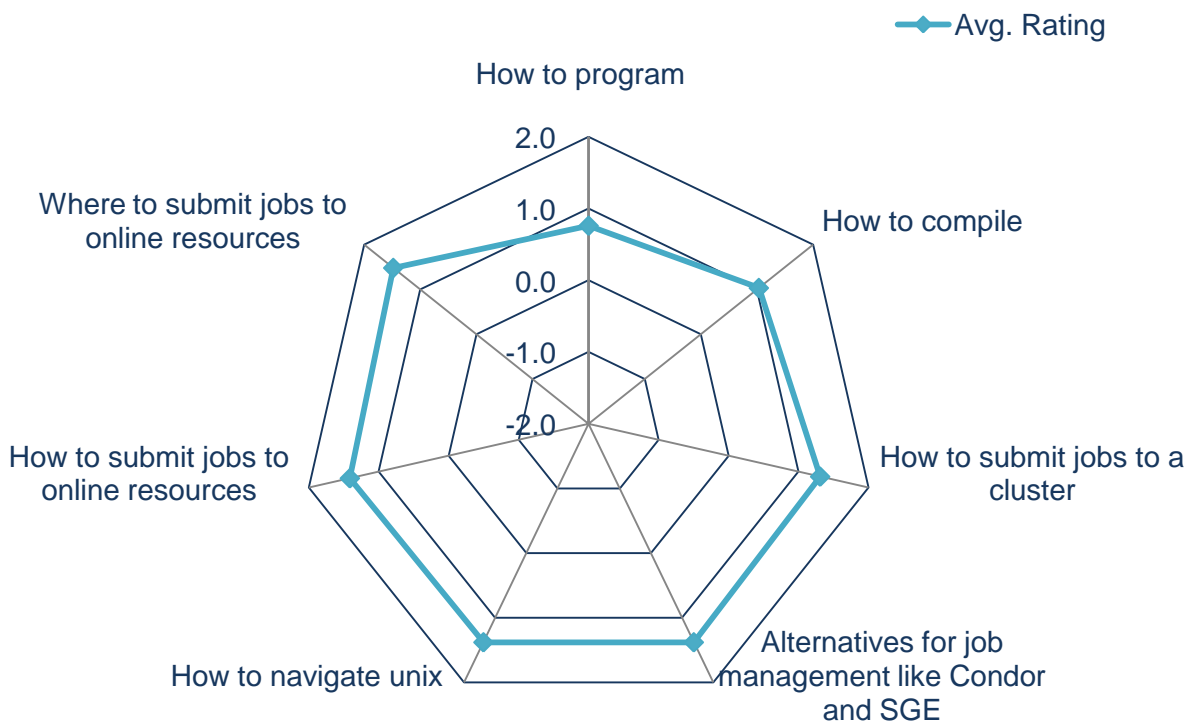
“I would like to see more of R and Perl scripting.”

Participant Knowledge

Respondents reported the greatest knowledge gains in learning where and how to submit jobs to online resources, followed by how to navigate Unix, alternatives for job management, and how to submit jobs to a cluster. Fewer agreed that they had a better understanding of how to program and how to compile (Figure 7, answered on a 5-point Likert scale from -2 to 2 for “strongly disagree” to “strongly agree”).

Figure 7. Participant knowledge gains

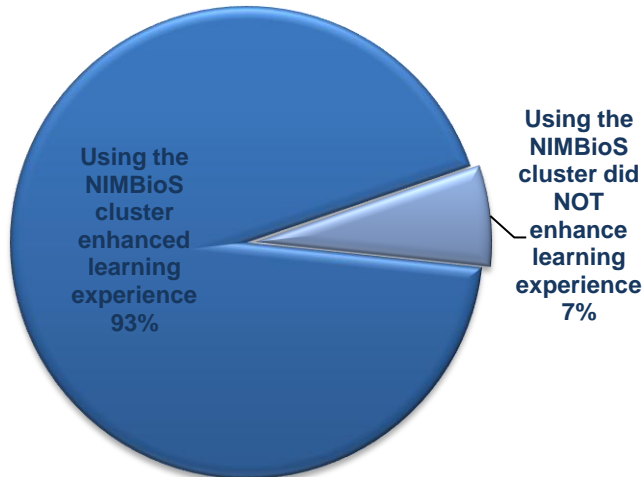
As a result of attending the tutorial, I have a better understanding of:



NIMBioS Cluster

Almost all respondents indicated that using the NIMBioS cluster enhanced their learning experience during the tutorial (Figure 8). Most respondents indicated that the hands-on

Figure 8. Impact of NIMBioS cluster on learning experience



experience of using the cluster was good practice:

“Having group access to the NIMBioS cluster was quite helpful. I have a fair amount of experience accessing other clusters so much of that section was review for me, but the practice of compiling a program in the cluster was useful,

and I picked up a few other tricks throughout the group exercises”.

“Not only did my ability to use the cluster 'enhance' my learning experience, I feel that that and the other hands-on experiences were absolutely critical to learning these concepts.”

“The hands-on exercises on the NIMBioS cluster were very useful.”

“There is nothing like hands on experience...”

“... It opened for new ways of running phylogenetic analyses.”

Some respondents felt that, while accessing the NIMBioS cluster was useful, it would have been helpful to allow participants to be able to sign up to use a cluster once they were back at their home institutions:

“Yes, but more would have been better! And really, the most useful thing that was missing from the seminar was basically "Here is an account on (some big cluster somewhere) that you are free to use for e.g. 1 year or whatever". It wouldn't necessarily have to allow HUGE jobs, but running e.g. moderate phylogenetic jobs.”

“It would have been great if we had the option to be signed up for a cluster that we could use directly during the meeting and also after the meeting. Such as an

automatic Teragrid or TACC account. We would have everything set up for phylogenetic runs as soon as we came home.”

Impact on Future Research

All respondents indicated they would be more likely to include high performance computing resources in their future research as a result of attending the tutorial. Some participant comments about the impact they felt the tutorial would have on their work:

“This HPC tutorial opened up a new world for me. It will enhance my research program, and that of my students and collaborators by several orders of magnitude. My lab will be learning Perl and "R". I now understand the relevance and importance of these tools. I am convinced that I must master the tools introduced in this tutorial if I am to remain competitive for certain kinds of funding.”

“I am already using HPC resources in my research, and will certainly continue to do so, but I am probably more likely to include the resources that were discussed, like TeraGrid and the CIPRES portal. I was not previously aware of Condor, and I am quite interested in setting up a Condor cluster in my department.”

“I will be using CIPRES for sure. I will also be using RAXML.”

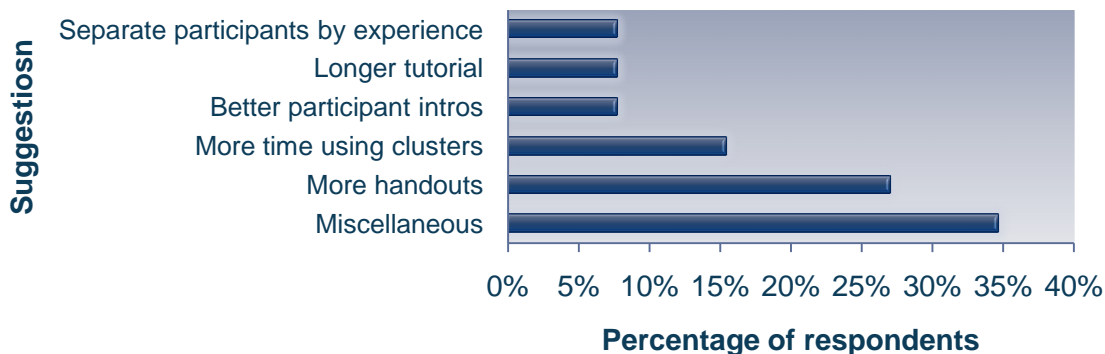
“I will definitely include it. I generated a matrix of 450 DNA sequences. The only way to make analyses quick and save time, is using HPC.”

“Plan to install CONDOR to create a local network; have also enrolled in iPlant.”

Suggestions for Future Tutorials

Respondents were asked for suggestions for improving future tutorials (Figure 9).

Figure 9. Suggestions for future tutorials



The most common theme that emerged from analysis of responses was to provide more handouts to participants (See Appendix C for complete list of suggestions):

"...Perhaps a printed tutorial given before the course began..with a specific example/data set...would help..."

"If the list of commands is available before the tutorial that would be very very useful. During the hands on, I could not decide whether I would rather copy all the commands (history) (which is what some people are doing) rather than follow the commands on my laptop. When I followed the commands on my laptop, I wasn't able to note down all the history of commands (especially if there errors and somebody will come to you for help, but the instructor will be continuing on the lecture so you would think, you've already miss something) so when the night comes when we can practice or repeat the examples on our own, I am at lost for commands. This could have been avoided if I have the list of commands before the tutorial, comment on the list during the day when they run through it (rather than typing or copying it for the first time), then I can follow up during the hands-on at the same time and also have list to follow during the night, without worrying that I missed something."

"These types of tutorials are very difficult to pull off well. It is very easy for a person to fall behind and never get caught up, despite the helpers that are there. I suggest written instructions that a person can use to catch up when they fall behind..."

"I would like talks given by Instructors to be compiled and given to us. More importantly I would like to have a note (paper or e-versions) on processes of running analyses on cluster."

Other participants suggested allowing more time for running jobs on clusters:

"...I would have preferred to have spent considerably more time working online in a cluster environment and dealing with queuing, bash scripts etc"

"Less about clusters in general, and more about their use for phylogenetics."

"I think I would take out the talk about how to put together a cluster and will expand more on running analyzes on the clusters and the cloud."

Others felt that longer participant introductions would have been helpful:

"It would have been very useful to have some means of knowing the research interests of the other participants beforehand or during the tutorial. There were a number of individuals who were sitting at the other end of the room who I never interacted with, and I have no idea whether or not they may have been a great contact for me. If we had been given a bit more of an opportunity to introduce ourselves and our research to the group, that would have been helpful, or if some form of organized working groups had been implemented, that would have been a better use of some of the "free interaction time".

“When introducing participants, have them give a 1-2 explanation of what they do and where they are - just having a name is next to useless”

“When we did introductions of the participants, it would have been useful to have had a little bit of introduction to what each of the participants was working on. My interactions with other participants would have been more productive if I had known from the beginning who might be conducting similar research to me...”

Other suggestions included separating participants by experience level and having a longer tutorial.

Conclusions and Recommendations

Overall satisfaction with the tutorial was high among respondents, the majority of whom indicated that the tutorial met their expectations and was appropriate to their level of expertise. Respondents were also satisfied with the travel, housing, and other amenities offered by NIMBioS.

The majority of respondents thought that the tutorial format was effective for achieving its goals and were satisfied with the opportunities provided during the tutorial to ask questions and/or make comments. Many respondents indicated the exposure to new information about the resources available for research was the most useful aspect of the tutorial, while felt learning to compile jobs on the cluster was useful. The majority of respondents felt the amount of content covered during the tutorial was “just right.” When asked what topics they would have liked to have covered in this tutorial if given more time, the most common responses were related to making comparisons among the different software packages available for phylogenetic analyses. Several respondents indicated they would have liked to include more hands-on sessions, and also more information about R.

Respondents reported the greatest knowledge gains in learning where and how to submit jobs to online resources, followed by how to navigate Unix, alternatives for job management, and how to submit jobs to a cluster. Fewer agreed that they had a better understanding of how to program and how to compile. Almost all respondents indicated that using the NIMBioS cluster enhanced their learning experience during the tutorial.

Several suggestions were made for improving future tutorials, including providing more handouts to participants, allowing more time for running jobs on clusters, and longer participant introductions. Other suggestions included separating participants by experience level and having a longer tutorial.

Based on analysis of respondent response data, the recommendations for future tutorials are as follows:

- The pre-tutorial webinar was very well-received as an introduction to the tutorial material. Consider offering introductory webinars for future similar events.

- Consider offering more handouts to participants containing steps for the hands-on sessions, software used, and definitions of terms that may be new to many participants.
- Consider setting aside time at the beginning of tutorials to allow for longer participant introductions where participants can indicate their background and interest in the tutorial topic.
- For future tutorials on high-performance computing, consider a presentation comparing the various software packages (what each can and cannot do, availability, cost, etc.) available for analysis on the tutorial's topic.
- Participants repeatedly suggested allowing for more hands-on time. If feasible, work more of this type of activity into future tutorials.

Appendix A

List of Participants

Participants

Last name	First name	Institution
Bond	Jason	East Carolina University
Bravo	Gustavo	Louisiana State University Baton Rouge
Cannon	Johanna	Auburn University
*Carr	Eric	University of Tennessee Knoxville
Chung	Kyong-Sook	Morton Arboretum
Cranston	Karen	Duke University
*Ferguson	Jim	National Institute for Computational Sciences
Harris	Phillip	University of Alabama Tuscaloosa
Heinz	Heather	Villanova University
Hodkinson	Brendan	Duke University
Jhwueng	Dwueng-Chwuan (Tony)	Indiana University Bloomington
Long	Elizabeth	University of California Davis
Matheny	Patrick	University of Tennessee Knoxville
Matzke	Nicholas	University of California Berkeley
Mazzoni	Camila	Leibniz Institute for Zoo and Wildlife Research
*McKay	Sheldon	University of Arizona
Miranda-Esquivel	Daniel	Universidad Industrial de Santander
Moore	Michael	Oberlin College
Munro	James	University of Maryland Baltimore
Nielsen	Stuart	University of Mississippi
*O'Meara	Brian	University of Tennessee Knoxville
O'Neill	Eric	University of Kentucky
Renner	Susanne	University of Munich
Rodriguez	Josephine	University of California Santa Barbara
Salgado	Catalina	University of Maryland College Park

Sanciango	Millicent	Old Dominion University
Schwarzbach	Andrea	University of Texas Brownsville
Shen	Chengli	University of Pittsburgh Bradford
Smith	Michelle	University of Tennessee Knoxville
*Stamatakis	Alexandros	Heidelberg Institute for Theoretical Studies
*Stanzione	Dan	University of Texas Austin
*Thomson	Robert	University of California Davis
van Velzen	Robin	Wageningen University
Velez Zuazo	Ximena	University of Puerto Rico Río Piedras
*Wilgenbusch	James	Florida State University
Wyngaard	Grace	James Madison University
Yang	Ya	University of Michigan Ann Arbor
Yang	Xi	Cornell University
Yessoufou	Kowiyou	University of Johannesburg South Africa

* **Organizer** (Organizer Susan Holmes did not attend)

Appendix B

HPC for Phylogenetics Tutorial Evaluation Survey

HPC for Phylogenetics Tutorial Evaluation Survey

Thank you for taking a moment to complete this survey. Your responses will be used to improve the tutorials hosted by the National Institute for Mathematical and Biological Synthesis. Information supplied on the survey will be confidential, and results will be reported only in the aggregate.

Pre-tutorial Webinar

Did you participate in the pre-tutorial webinar?

Yes

No->*skip to tutorial evaluation*

How useful do you feel the webinar was in preparing you for the tutorial?

Comments about the webinar:

Tutorial

Please check the appropriate box to indicate your level of agreement with the following statements about this tutorial: (Very satisfied, Satisfied, Neutral, Dissatisfied, Very dissatisfied)

The tutorial was appropriate to my level of expertise.

The tutorial met my expectations.

The hands-on exercises were useful.

The presentations were useful.

The instructors were very knowledgeable about their topics.

I would recommend participating in NIMBioS tutorials to my colleagues.

How do you feel about the amount of content offered during the tutorial?

Too little for the allotted time

Too much for the allotted time

Amount of content was just right

Please check the appropriate box to indicate your level of agreement with the following statements. As a result of participating in this tutorial, I have a better understanding of: (Strongly agree, Agree, Neutral, Disagree, Strongly disagree)

where to submit jobs to online resources

how to submit jobs to online resources

how to submit jobs to a cluster

alternatives for job management like Condor and SGE

how to compile

how to program

how to navigate Unix

Did the ability to utilize the NIMBioS cluster enhance your learning experience during the tutorial?

Yes

No

Comments:

Will you be more likely to include High Performance Computing resources in future research as a result of attending the tutorial?

Yes

No

Comments:

What topics would you have liked to have covered in this tutorial if given more time?

What do you feel was the most useful aspect of the tutorial?

What would you change about the tutorial?

How do you feel about the format of the tutorial?

This was a very effective format

This was not a very effective format

The tutorial format would have been more effective if:

Please indicate your level of satisfaction with the tutorial accommodations:
(Very satisfied, Satisfied, Neutral, Dissatisfied, Very dissatisfied, Not applicable)

Travel arranged by NIMBioS

Housing arranged by NIMBioS

Comfort of the facility in which the tutorial took place

Resources of the facility in which the tutorial took place

How satisfied were you with the opportunities provided during tutorial presentations and discussions to ask questions and/or make comments?

Very satisfied

Satisfied

Neutral

Dissatisfied

Very Dissatisfied

Please provide any additional comments about your overall experience with the tutorial:

Appendix C

Open-ended Survey Responses

What do you feel was the most useful aspect of the tutorial?

Access clusters

How to compile programs on a cluster and how to get a job started on a cluster.

How to deal with a cluster from compiling to submitting a job.

running analyses in cluster and submitting jobs to portal sites

That they taught us how to access, use and submit jobs to a cluster.

The use of cluster

The fantastic instructors (both in the content of their lectures and in their willingness to wander around and help us individually), the diverse nature of the participants, and hands on opportunities.

The most useful aspect was to interact with the professionals who develop tools.

The hands-on practice, and evening discussions

The most useful aspect was hands-on activities, e.g. Unix commands, submitting jobs to the cluster, etc. I appreciated that Brian slowed down and waited for everyone to get on the same page.

It gave me a lot of information about phylogenetic tools I can use now to develop my research project.

It made me aware of my options.

Networking and getting to know all the available resources.

The breadth of resources and tools that were covered. It whetted my appetite for learning Perl and R, for adding 'shared computer resources' to my current collaborations and proposals for funding. The instructors were top notch and just meeting them and picking their brains about computing topics was a wonderful resource.

The hands on job submission tutorial on the NIMBioS cluster, and the lectures by Alexandros Stamatakis.

UNIX commands, Perl scripting, RAxML, CIPRES, CONDOR and the availability of the instructors to discuss issues.

Alexis Stamatakis's lecture; the tutorial in Perl-scripting; the example of compiling MrBayes on a cluster, and running an analysis.

Bob's Perl programming sessions, information on how to compile, Alexis' session

I am familiar with Perl and R now.

The Perl tutorial was particularly helpful, as was the introduction to the Condor, TeraGrid and the CIPRES portal.

The UNIX/PERL demonstrations and conversing with others about common problems.

How to install CONDOR.

HPC in R, general background lectures on HPC

Linux language, group discussion

The opening day morning session was exceedingly useful. Thomson especially did a great job. O'Meara's frustration with those who fell behind could have been better hidden. (Although to a certain extent - it was justified). Talks and demonstrations by Wilgenbusch and Stamatakis - both during the day sessions and especially in the evening open sessions were exceedingly useful. Stanzione's talks were also of interest.

O'Meara's sessions were the most useful.

What would you change about the tutorial?

When introducing participants, have them give a 1-2 explanation of what they do and where they are - just having a name is next to useless.

When we did introductions of the participants, it would have been useful to have had a little bit of introduction to what each of the participants were working on. My interactions with other participants would have been more productive if I had known from the beginning who might be conducting similar research to me. Some sort of structured working groups would have been helpful - phylogenomics, population genetics, phylogenies w/ large numbers of species, etc. Also, much of the discussion of computing architecture was a little too dense for me. I felt that in those presentations, too much was covered too quickly, and that for many of the participants (myself included), the level of detail may have been a bit too much, and the presentations became difficult to follow.

Due to the amount of information we received I would suggest to make the tutorial last longer, 3 or 4 full days....

I would extend the time. Instead of a 2.5 day tutorial, I would have a little longer to have more hands-on exercises.

Where the tutorial fell down a little was on organization. It seemed to me like each presenter was simply given a topic to cover, and then was allowed free rein to cover it however he chose. Thus we had some presentations that did not seem terribly appropriate to the level of student attending, e.g. Alexis Stamatakis' presentation, and Dan Stanzione's presentation on supercomputer infrastructure (although don't get me wrong--I was glad they were present). I felt the workshop would have been much stronger had it focused much more on hands-on activities, supplemented by short background lectures. Bob's programming lectures went too fast for me. I thought he covered the appropriate amount of material, but I would have preferred it to have been stretched out some more. I also would have rearranged the order of presentations a little, so that the more whiz-bang stuff (supercomputer infrastructure, e.g.) was saved for the end.

Absolutely no screen projections with orange writing on black background (why was this used in three lectures on both days, even though participants indicated they could barely read the dark

screen already the first morning)?

I wish I was told to have ModelTest and PAUP on my computer before I arrived. I have those programs on my desktop but not laptop.

I would make a previous tutorial just for Linux and Bioinformatics. Maybe I would use Linux or Mac OS X compulsively, I think Windows is not well intended for the command line use, and sooner or later you will use Linux-MacOS X. A final technical session to discuss pro/cons of some parallel programs. Is it better RAxML or GARLI? Or PHYML?

I would not have multiple lectures on the design, configuration, and energy requirements of supercomputers. From the perspective of a practicing systematist, it would have been more useful to have a lecture on when the use of a supercomputer would be appropriate, getting accounts set up, job submission, etc. I know we did that for CIPRES, but walking us through the procedure for Longhorn would have been helpful.

1) I am still unclear as to what resources there are out there that I can use, other than CIPRES, requesting a TeraGrid account, and perhaps an iPlant account. There was no summary and while learning what Kraken and the other clusters are made of is of some interest, it is of little use to me if I cannot make use of the resources. 2). The tutorial ended without a summation! Some sort of summary as to what we covered and what resources were available to us would have been appropriate and useful. 3). Condor was discussed at great length and yet it's utility to HPC phylogenetics is still unclear - i.e. what programs can I run using Condor?

I was lost in the Perl part. Would be better to study the very basics of Perl ahead of time.

I would make it a lot more practical and a lot more current. There was a lot of information about what will be available sometime soon, but I feel that other people, like me, wanted to learn how to use what is available now, with hands-on.

More hands-on activities; more discussion about when a particular resource is appropriate (i.e. choosing between options)

See previous comment, too introductory, not enough practical running of jobs on different systems.

Unfortunately, I thought most of the tutorial was entirely too general to be of any real use. I would have preferred to have spent considerably more time working online in a cluster environment and dealing with queuing, bash scripts etc.

Less about clusters in general, and more about their use for phylogenetics.

I think I would take out the talk about how to put together a cluster and will expand more on running analyzes on the clusters and the cloud.

for the hands-on tutorials, handouts available ahead of time with commands

Go more slowly when demonstrating use of programs and navigating computer systems via UNIX. Perhaps a printed tutorial given before the course began..with a specific example/data set...would help. The online tutorial recommended several days before the course was not particularly useful.

If the list of commands are available before the tutorial, that would be very very useful. During the hands on, I could not decide whether I would rather copy all the commands (history) (which is what some people are doing) rather than follow the commands on my laptop. When I followed the commands on my laptop, I wasn't able to note down all the history of commands (especially if there errors and somebody will come to you for help, but the instructor will be continuing on the lecture so you would think, you've already miss something) so when the night comes when we can practice or repeat the examples on our own, I am at lost for commands. This could have been avoided if I have the list of commands before the tutorial, comment on the list during the day when they run through it (rather than typing or copying it for the first time), then I can follow up during the hands-on at the same time and also have list to follow during the night, without worrying that I missed something.

More detailed handouts

These types of tutorials are very difficult to pull off well. It is very easy for a person to fall behind and never get caught up, despite the helpers that are there. I suggest written instructions that a person can use to catch up when they fall behind. Also using a computer lab, in which everyone is working on the same type of computer, would help eliminate the types of problems I saw.

Add more fantastic instructors, more hands on, more suggested reading and handouts, and 1-2 more days to spread out the subjects, better AV since we couldn't see the bottom 1/3 of the screen and text was sometimes very small. I wish there had been a bit more outreach so more people could understand how this stuff is useful to them.

I would like talks given by Instructors to be compiled and given to us. More importantly I would like to have a note (paper or e-versions) on processes of running analyses on cluster

Have options for various classes of varied levels/interests held simultaneously.

Trying to bring people together with more similar experience levels. e.g. real beginners vs. more experienced. It seemed like a wild mix ranging from people who did not even know Unix commands to people who have been programming PERL for years and advanced cluster users. I would have preferred less lecture on the hardware but more on the software, especially more applied aspects. More on simple PERL programming or modifying existing scripts. What was offered was okay but too advanced for the real beginner. It would have helped to give out exercises for the evenings, e.g. examples of compiling and running programs that would cover the typical problems one would encounter (such as cryptic compiling instructions), recognizing software bugs, simple programming etc. Probably different levels of exercises so something for everyone.

The tutorial format would have been more effective if:

A note was given to us. This would have helped me to follow the instructors in their demonstrations, and more importantly, I would be able to do it myself when get back to my university.

For the most part, the format of the tutorial was fine. I would point out that many of the "hands on" sessions should have been - "watch this" as the abilities of the class differed so markedly and there was not the time for all to follow. What is needed is a better introduction and overview of the resources available and the software. One example would be 1) here is Condor. 2) This is how one utilizes Condor - i.e. you need a pool of computers 3) This is how you find resources for setting up Condor 4) these are the applications you can run in Condor.

It was more detailed and the general HPC aptitude of the class was equivalent.

What topics would you have liked to have covered in this tutorial if given more time?

I would not have expanded the content...rather I would have liked to proceed more slowly through each exercise that required we type in commands on our own computers. There were so many different platforms; I just needed more time to remember where on my computer I was and where I was trying to go.

Gene tree vs. species tree analysis methods

Speciation detection, recombination analysis...

Tree dating

Would be better if we were told to self-study the basics of Perl before the trip. Also, self-study materials would be better to be distributed at least a week ahead of time, instead of several days.

I would have liked to have had more time on programming. This segment was rushed. I understand that there's no way to teach programming in a 3-day workshop, but some more hands-on sessions dealing with programming would have been useful. It would have also been useful to discuss which programming languages are most appropriate to phylogenetics, and the pluses and minuses of each. We touched on this a little, but I would have liked more.

Additional hands-on exercises on up-loading and running data sets. I think one of the long lectures by Wilgenbusch on clusters (which mainly stated that one should not try to build them) or Dan Stanzione (which included way too many disparaging remarks about NSF and self-congratulatory attempts at humor) could have been cut to 30 minutes each.

I wish the webinar was a scheduled part of the conference so that I could have attended (It was announced too late for me to change my schedule). I wish we had a bit more hands on experiences and some more handouts with basic Unix, Perl (incl. reg. expressions), and python commands. I wish there had been more explicit suggestions for further reading material and prudent technology investments for laptops or desktops. I wish the material was spread out a bit more. I was glad we had the option to come back in the evening but I found it too exhausting. I think perhaps one more day with open ended discussion and technical help would have been very useful. I'm not sure how to accomplish this, but I would have liked to see a bit more about how to foster collaborations between computer science and biology, especially for folks whose home institutions don't have the infrastructure of things like NIMBioS. Maybe we should push for college courses like 'First-time programming for biologists'.

I would like to have had more details on how to run, get data and visualize trees in different

hands-on session.

More practice about submitting projects on a cluster and a hands-on experience using the cloud in Amazon, for example. Also, since a crash course on perl was offered it would have been good to expend more time with that.

The topics were perfect. However, it would have been great if there were a couple more days to do more hands-on exercises,

Additional time on Perl scripting would have been extremely useful. Also, for folks that have multi-processor laptops, setting up Condor on your laptop and showing us how to run a PAUP analysis would have been great.

A bit more information on R would have been useful.

An alternative to Perl as Python. A worked example of a full analysis submitted to the cluster as a single bash file, from alignment to ML or Bayesian analysis, including ModelTest. An analysis using R and ape/picante using a cluster.

I would like to see more of R and Perl scripting.

1) A more thorough discussion of the programs applicable to HPC resources 2) continuing with that line of thought - a better understanding of what resources are needed by these programs so that I may have an informed discussion with a cluster systems administrator.

Comparison of different software for phylogenetic analysis

Current state of various commonly used phylogenetics programs - how to configure each to run more parallelized.

I would have liked to have learned more about the software that can be used for parallel computation for phylogenetics: RAxML, MrBayes (others?). How they implement parallelization and what the pitfalls are

More about the analysis by using the softwares

It was a little too focused on introductory stuff like UNIX. It would have been better to have one model dataset set up, representing a typical phylogenetics problem which you might want HPC for, and then run that dataset through all of the various options we discussed. This would much more directly inspire future use...

Comments about the webinar:

Brian is a great teacher, his patience helped me learn.

I got a clear clue about what going to happen during the tutorial. It's very useful

I had only a very rudimentary acquaintance with UNIX language and this tutorial gave me a heads up on what kinds of research projects might require this material.

If you know about the commands, then it is okay if you do not attend the webinar. For me it is somewhat useful because I am familiar about what was discussed during the webinar. For others who are not familiar, the webinar should be a really good intro just to get a feel of what's ahead during the tutorial.

It was great--thanks for setting this up. This was a useful way to get lots of folks together in an easy-to-use interactive, online environment.

A 1-2 page list/description of "UNIX commands you actually need" would have been just as good, probably.

I think written instructions should be provided before the webinar. It's too easy to fall behind when you have nothing to refer to for instructions.

It would have been helpful to have a pdf ahead of time of the commands we were working with

Brian was going to fast during the webinar which made it difficult for the ones that had less experience writing commands.

Somewhat too fast to follow. I think a self-paced tutorial would work better.

I think it could be longer and include some other Linux skills needed as regex and maybe Perl

NIMBioS Cluster Comments:

For practice, the cluster was quite useful.

Having group access to the NIMBioS cluster was quite helpful. I have a fair amount of experience accessing other clusters so much of that section was review for me, but the practice of compiling a program in the cluster was useful, and I picked up a few other tricks throughout the group exercises.

I learnt many things I was not well aware of, previous to the tutorial

It helped being able to log on to the cluster. My wireless connection was somewhat unreliable - so I would lose my connection and then would get a little bit behind.

Not only did my ability to use the cluster 'enhance' my learning experience, I feel that that and the other hands-on experiences were absolutely critical to learning these concepts.

The hands-on exercises on the NIMBioS cluster were very useful.

There is nothing like hands on experience. Still, the instructors moved through the material too quickly. It would have been helpful to have a cheat sheet of UNIX commands...so that if one fell behind, one could quickly find a command...and catch up.

Yes indeed. It opened for new ways of running phylogenetic analyses

I found out that this cluster is quite different to my University's cluster. I also realized I need to get more experience with UNIX.

Yes, but more would have been better! And really, the most useful thing that was missing from the seminar was basically "Here is an account on (some big cluster somewhere) that you are free to use for e.g. 1 year or whatever". It wouldn't necessarily have to allow HUGE jobs, but running e.g. moderate phylogenetic jobs.

It would have been great if we had the option to be signed up for a cluster that we could use directly during the meeting and also after the meeting. Such as an automatic teragrid or TACC account. We would have everything set up for phylogenetic runs as soon as we came home.

I only learned how to log in to the cluster. I never really used it for anything.

More likely to use HPC resources in the future comments:

But will all depend on if I get an account on a cluster.

Although my understanding of Linux is fine, I thought that HPC was more confusing, but right now I see HPC as "the" normal way to conduct some phylogenetic analyses as bootstrapping

Absolutely!

Absolutely.

I am already using HPC resources in my research, and will certainly continue to do so, but I am probably more likely to include the resources that were discussed, like TeraGrid and the CIPRES portal. I was not previously aware of Condor, and I am quite interested in setting up a Condor cluster in my department.

I knew I needed HPC for my research, but before the tutorial I had no idea that there were so many ways it would be useful at many levels of my project. I'm hoping to convince other people in my field about how useful these skills and resources are.

I will be using CIPRES for sure. I will also be using RAxML.

I will definitely include it. I generated a matrix of 450 DNA sequences. The only way to make analyses quick and save time is using HPC.

Plan to install CONDOR to create a local network; have also enrolled in iPlant.

This HPC tutorial opened up a new world for me. It will enhance my research program and that of my students and collaborators by several orders of magnitude. My lab will be learning Perl and "R". I now understand the relevance and importance of these tools. I am convinced that I must master the tools introduced in this tutorial if I am to remain competitive for certain kinds of funding.

Yes, at the very least knowing about HPC in R and HPC versions of common phylogenetics programs would be good...

Please use this space for any additional comments:

1. It has remained unclear to me what the learning goals for the workshop were exactly. 2. I noticed that the hands-on tutorials were sometimes confusing because of the different operating systems used by the various participants. A possible solution for this might be to let everyone work with the same (usb-installed?) Linux distribution.

Everything was just good, people were very nice with the participants, the instructors were very good at what they do, they are famous for this so it was an honor to be there as a participant.

I really appreciated the tutorial--thanks very much!

I really enjoy the tutorial

It would be a great service to many of my colleagues if this tutorial were offered again...to those who were not selected to participate the first time. I think that the influence of this tutorial was wide and deep, because it is very likely that almost all participants will return to their institutions, share what they learned, and change the way they are communicating/sharing their data, as well as analyzing their data. The content was very up to date...and visionary in some presentations. Although the 1st day was especially frustrating (learning PERL), things finally clicked on the 2nd day. This is a great use of NSF funds...and a superb example of large awards, such as for centers, having potentially far more impact than dozens of small awards made to single PIs. I say this, even though I am often distraught by the low funding rates in the research programs.

Kudos to Jennifer for making my travel arrangements! Same to Eric who dealt with the numerous wireless issues. Stamatakis, Wilgenbusch and Thomson were exceptionally helpful. O'Meara put on a damn good tutorial, although there is certainly room for improvement, especially in terms of some of the subject matter covered. Stanzione was energetic, exceedingly knowledgeable and very helpful.

Thank you for having me, I enjoyed my visit and learned a lot from the course.

Thank you for the opportunity to participate! I am more than ever motivated to learn more Unix and PERL and dive into larger data set analyses. I hope to get to the point where I can teach this to our undergraduate students (I know no one else of my colleagues will do). I think it would be great (possibly for a grant?) to be paired with a computer science/math faculty or computational biology faculty who could co-teach a course in bioinformatics/genomics/phylogenetics etc. Probably with rotating locations. Most small, primarily undergrad schools don't often have faculty in all areas of expertise, limiting collaborations of that kind.

Thank you so much to all those who organized and to the lecturers of this tutorial. This is very very useful to me because we never had someone teach me how to use HPC for phylogenetic analyses - I have to learn everything by myself...and I still have a lot to learn! It would be nice if the tutorial is longer so we will have more time to get comfortable using it. A part2 would be nice! Again, thank you NIMBioS!

Thank you!!!! This was such a wonderful experience.

Thanks a lot for an awesome learning experience!

Thanks!

The tutorial was a great opportunity for me to learn new techniques about phylogenetic analysis.

This tutorial was overall quite useful, but for me, most of sessions were pretty evenly divided between "things I know very well and use every day already" and "things that are kind of over my head." I completely understand the necessity of getting everyone up to speed on command line use, but as someone who uses shell scripts every day, that part was a little slow for me. (The Perl help was great, though). On the other hand, some of the more detailed sessions on computing architectures, etc., were a bit beyond my computer science expertise. If many of the folks in the room were new to Unix/Linux, I can't imagine that they found that part to be any easier. Would it be possible to split the group for some sessions?