

Applying for, Operating, and Renewing a Research Experience for Undergraduates Site Program in Computational Biology

Hong QIN

Department of Computer Science and Engineering
Department of Biology, Geology, and Environmental Sciences
University of Tennessee at Chattanooga
Chattanooga, Tennessee, 37403, U.S.A.

ABSTRACT¹

We reflected on our experience in applying for, operating, and renewing a Research Experience for Undergraduates Site program in computational biology from 2016 to 2021 at a metropolitan public university, funded by the National Science Foundation of the U.S.A. Our mistakes during the application process were avoidable in retrospect. Our Interdisciplinary Computational Biology, iCompBio, REU site was a 10-week summer research program that included R and Python coding bootcamps at the beginning and short workshops on various topics interspersed throughout the program. ICompBio was in-person in 2019 and virtual in 2020 and 2021. ICompBio was fortunate to have a team of faculty mentors that were committed to undergraduate research training and recognized the importance of diversity in computational biology. Many students and faculty mentors turned the challenges in virtual REU programs into learning opportunities and had fulfilling research experiences. ICompBio participants reported learning gains higher than average in a 2020 external evaluation for virtual REU sites, particularly in the aspects of Research Self-Efficacy, Science Identity, and Life Utility. Our lessons and experiences may be helpful to other faculty and researchers that are interested in similar programs at their own institutions.

Keywords: Research Experience for Undergraduates, Computational Biology .

1. INTRODUCTION

The Research Experience for Undergraduates (REU) is a program funded by the National Science Foundation to offer research opportunities for undergraduates in the United States. An REU site is typically located in a host institution. Each REU participant is expected to work closely with a faculty mentor and other researchers on a specific research project. REU participants typically receive stipends and assistance for housing, travel, and food. The NSF REU program supports both REU Sites and REU Supplements. The former is typically a summer research program with a cohort of students, whereas the latter often supports a few students for a research project

funded separately by the NSF. Many NSF units support the REU program. We submitted our REU Site application on inter-disciplinary Computational Biology (iCompBio) to the NSF Division of Biological Infrastructure (DBI) in the Directorate of Biological Science.

In recent years, the deadlines for REU Site applications were in late August, and applicants are expected to receive reviews and funding decisions within six months. Intellectual Merit and Broad Impacts are two of the NSF's merit review criteria. External reviewers are often used in the NSF Merit Review process, either as a group known as a panel or as individuals known as ad hoc reviewers. For each proposal, reviewers can give it an Excellent, Very Good, Good, Fair, or Poor rating. Three separate reviewers are usually assigned to each proposal. The review panel typically meets as a group to discuss all proposals and make a recommendation to each one.

2. THE APPLICATION PROCESS IN 2016-2018

A beginner's mistake in 2016.

Our initial application for the REU program was submitted in August 2016, and received a non-competitive rating from an NSF review panel. Naively, we proposed several research projects on disease and human health, which was raised as a major concern by the review panel. Traditionally, disease and human health research are funded by the National Institutes of Health, and thus is not a funding priority area of the NSF, particularly for the bio-related funding programs.

A medium priority rating in 2017.

Our second attempt was submitted in August 2017. Our major revision was the removal of research projects on diseases and human health. Our proposal received a "medium priority" recommendation from the panel based on two "Very Good" and one "Good" review.

Based on three reviews and the panel summary, the strength in Intellectual Merit of our proposal included the interdisciplinary nature of computing and biology, the externally funded research projects, diverse expertise of the faculty mentors, our active YouTube education channel [1], research training workshops, and our letters of collaborations from minority-serving institutions

¹ We thank Dr. Yingfeng Wang for editing this manuscript.

One major critique was that our proposed external participants at 50% were too low. Another major concern was about our proposed two-week bootcamps at the beginning of the iCompBio, as well as the two-week writing and presentation session at the end. The panel constructively suggested us replace these two-week bootcamps and workshops with interspersed short versions throughout the summer program. Because we removed the research component of disease and human health-related projects, we had only 8 faculty mentors and 8 research projects, which was a concern raised in the review process. The review panel was also concerned about the coordination of joint mentoring of an experimental mentor and a computational mentor.

For the Broader Impacts, the panel highlighted our target recruitment plan with minority-serving institutions and the interdisciplinary nature of the program. No concerns were raised about the Broader Impacts in our proposal.

A highly meritorious rating in 2018.

In August 2018, we resubmitted our application with revisions to address the previous reviewers' concerns. We increased the number of external participants of the proposed REU to 80%. We shortened the beginning coding bootcamps. We proposed interspersed short workshops on computing, writing, and presentation, and soft-skill training throughout the program. We eliminated the rotation between experimental and computational labs, and let REU participants concentrate their research in one host lab. We recruited additional faculty mentors and proposed ten host labs and research projects. The review panel recognized the cross-disciplinary nature between computing and biology, the interdisciplinary background of the principal investigator, and agreed that our proposed research projects can provide proper cross-disciplinary training in biocomputing to undergraduate researchers. One concern was raised that a few research projects required substantial effort on data collection, and might not expose students to enough training in computational analysis or modeling. Fortunately, this issue appeared to be a minor one. Our proposal received a "Highly Meritorious" recommendation from the panel, based on two "Excellent" and one "Very Good" reviews.

It is worth noting that the REU proposal emphasizes cost effectiveness. In our case, we were expected to allocate at least 90% of the budget for the REU participants.

3. THE NORMAL IN-PERSON REU IN 2019

We started our recruitment effort in the early spring of 2019 as soon as we learned that our REU proposal would be funded. We emailed advertisement flyers to colleagues at dozens of minority-serving institutions and liberal arts colleges. We used Google Forms to handle the application process. Working with the IT department of our university, we set up a dedicated email address to handle all REU applications including letters of recommendation. A

dedicated website, <http://utc.edu/icompbio>, was set up to host our REU project information. We received 41 applications, including 10 Caucasians, 19 African Americans, and 7 Asians, and 5 Hispanics. After the reviews, we gave out 10 offers, but three of the applicants declined or did not respond in time. We then gave out additional three offers.

For logistics, we contacted the relevant university offices to request guest student IDs for email and network access, and swipe cards. With the help of departmental and university administrative assistants, we prepared the paperwork for stipend payments during their summer research. We contacted University Housing to host the incoming 10 REU students over the summer. We learned to schedule the REU student arrival and check-in with the housing staff work schedule. We eventually came up with a to-do list of relevant procedures for future reference.

We scheduled an orientation for the REU mentees and mentors. For research ethics training, we adopt the Collaborative Institutional Training Initiative (CITI) materials. We shared a master schedule using a Google Sheet with all REU mentees and mentors. All REU students went through a two-day R coding bootcamp, followed by a two-day Python coding bootcamp.

We were lucky to have committed faculty mentors and responsible REU participants in the first year of our REU program. In addition to the intensive research training, we organized field trips. REU students attended local research meetings, conferences, and professional workshops. REU students and some faculty often had informal lunch gatherings in the cafeteria. All of the participants in our iCompBio REU program gave poster presentations on their research projects during the last week of the program. One participant eventually had a first-authored publication based on his REU research project.

4. VIRTUAL REU IN 2020

Many universities switched to online education in the early spring of 2020. The University of Tennessee at Chattanooga switched to online education in the middle of March 2020. After contacting the Program Officer at NSF who was overseeing our REU grant, we submitted a "Change of Scope" request to change our in-person REU to a virtual REU for 2020. We revised our previous research projects to accommodate the remote research experience. We received 71 applications, made 17 offers, and received 12 positive responses. We provided laptops and internet connection support to these remote REU participants.

We offered online R and Python coding bootcamps not only to our iCompBio REU students but also to students from other REU programs. We recruited teaching assistants who worked with particular students in ZOOM break-out rooms to help them troubleshoot.

To build the cohort experience, REU students self-organized some online social events. We were fortunate to have an REU student that was a natural leader and motivated many students to participate in the online social events. Informal coffee hours were organized. Some research groups organized joint group meetings. At least three students contributed to the NSF REU blogs.

Our revised REU research projects include several projects on COVID-19, including an association study of virus transmission with mobility and weather conditions, social media analysis on COVID-19 related topics, predicting daily cases using machine learning models, predicting seasonality of COVID-19, and epidemiological modeling of COVID-19. The students enthusiastically pursued these projects. During the concluding online presentations of the summer research, some students invited their friends and

family members to join the online meetings. Many students stated that working on pandemic-related projects and comparing their computational and modeling results with the unfolding pandemics over the summer of 2020 had been a very fulfilling experience for them.

There are many unexpected challenges in the 2020 virtual REU. We did not anticipate the backlog of orders, and the laptops arrived after the mid-point of the 10-week program. We did not anticipate delayed responses from the university IT service during the pandemics. It took an unusually long time to resolve log-in issues when students need to access high-performance computing clusters through VPN. A few students also experienced software licensing issues due to working remotely. Some students experienced stress due to the lack of partition between work and personal space.

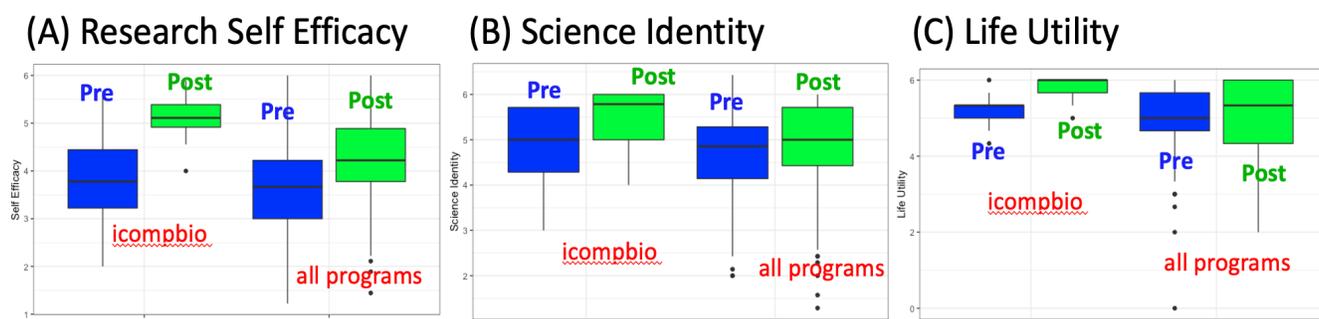


Figure 1. External 2020 virtual REU evaluation show iCompBio-2020 students reported higher gains than average, especially in (A) Research Self Efficacy, (B) Science Identity, and (C) Life Utility when compared with all students in the 23 remote REU programs. (These results were shared with us by Dr. Dolan’s research group [2]).

The 2020 virtual iCompBio participated in the NSF sponsored external evaluation for 23 remote REU programs, with an aggregated report in a peer-reviewed publication [2]. Focus group interviews were conducted in early- and post-program time points. One early-program survey and one post-survey were administered to seek student self-evaluation on learning gains of research self-efficacy, science identity, science value alignment, career intentions, graduate school intentions, intrinsic motivation, life utility, person importance, social utility, and person cost for research success. The iCompBio students reported gains in all 10 categories and are often higher than average. The gains of iCompBio students in Research Self Efficacy, Science Identity, Life Utility are particularly much higher than the average of all students in the 23 Remote REU programs (Fig. 1).

5. VIRTUAL REU IN 2021

After submitting a formal request to the funding agency again in the spring of 2021, we offered another virtual REU to 15 students in the summer of 2021. We actually gave 21 offers out of the 75 applications.

Based on our previous experience and student responses, we planned ahead to address potential IT issues and computer support. We offered tips to both mentors and mentees to schedule work routines, and intentionally partition work and personal space and time. We scheduled interspersed online workshops to help participants with basic computer usages, computational works, writing, presentations, and soft-skill training. All students gave presentations of their research projects at the end of the summer research.

We noticed that the fraction of under-represented minority applicants in 2021 was 20%, much lower than those in previous years (50% in 2019 and 34% in 2020). This might suggest education inequity was amplified during the pandemics, but a large-scale study would be needed to formally address this issue.

6. ASSESSMENT AND RENEWAL

In all three summer REU programs, we used the SALG pre- and post-REU surveys, which have been used as evaluation tools for many REU programs. Our 2019-2021 pre- and post-REU SALG instruments contained the same set of multiple-choice questions to ask students self-

evaluation on their understanding of concepts, skills, attitudes, integration of multi-disciplinary knowledge relevant to computational biology research, and their overall confidence as scientists and researchers. There were also short-answer questions to gather students' feedback and levels of satisfaction with the REU program structure and logistics.

For long-term career tracking, we requested REU participants to setup LinkedIn profiles and provided alternative emails. We were able to track the career status of all of the past 37 participants. At the time of our renewal application, 14 students have completed their undergraduate degrees. One of the 2020 REU participants received the prestigious NSF GRFP award. One of the 2019 REU participants received a prestigious Goldwater scholarship. Six of these students are in graduate training programs, and others are in various jobs in research and technology fields.

Our REU participants co-authored 8 peer-reviewed publications and 1 accepted manuscript [3-10], including two first-authored publications [3, 6]. One African American female student is the first author of a paper in the *Journal of Theoretical Biology* [6]. All REU students presented their projects in person in 2019 and online in 2020 and 2021. Five REU students, including 1 African American male and 1 African American female, presented their research in 9 national and regional conferences.

Many REU participants gave consent to release the recorded presentations of their summer research projects. We hosted these videos on a YouTube educational channel [1], and then cited these student presentations videos on our iCompBio website and our renewal proposal. We shared our R and Python coding bootcamps as public GitHub repositories [11, 12]. Many recorded tutorial videos of the R and Python bootcamps were also released on our YouTube educational channel.

The assessment effort and career path of past participants are required components in our REU Site renewal application. Student publications, student presentations, workshop materials, and tutorials are also tangible evidence of our activities and accomplishments. Thankfully, our renewal application received good reviews and is now officially funded.

7. EPILOGUE

To anyone interested in applying for and operating an REU Site program, it is important to recognize that the summer REU Site program is a service that would demand commitment and dedication not only from you, but also from a team of faculty mentors. The success of an REU Site would require the support and coordination of various units in the host institution. It would be useful to develop a to-do list for relevant procedures at the host institution. We hope that our experiences and lessons will be valuable

to others with similar interests and commitment to undergraduate research training.

8. ACKNOWLEDGEMENT

We thank the support of NSF award REU #1852042, CAREER #1453078 and #1720215, BD Spoke #1761839. We especially thank the support from the Office of Vice Chancellor for Research at the University of Tennessee at Chattanooga. We thank the SimCenter at the University of Tennessee at Chattanooga for providing high-performance computing facility and support of various events. We thank the commitment and hard work of all the student participants, faculty, and graduate student mentors of iCompBio from 2016 to 2021. We thank the administrative support from the Department of Computer Science and Engineering at the University of Tennessee at Chattanooga. We thank the Center for Student Success at the College of Engineering and Computer Science at the University of Tennessee at Chattanooga for supporting various events. We thank the Office of Undergraduate Research and Creative Endeavor at the University of Tennessee at Chattanooga for coordination and support of various events. We thank Dr. Mengjun Xie and Dr. Dalei Wu for reviewing this paper.

9. REFERENCES

- [1] H. Qin. "Hong Qin, YouTube educational channel." <http://youtube.com/qinstat/> (accessed June 1, 2022).
- [2] O. A. Erickson *et al.*, "'How Do We Do This at a Distance?'" A Descriptive Study of Remote Undergraduate Research Programs during COVID-19," *CBE Life Sci Educ*, vol. 21, no. 1, p. ar1, Mar 2022, doi: 10.1187/cbe.21-05-0125.
- [3] D. Ledesma, C. Powell, J. Shaw, and H. Qin, "Enabling automated herbarium sheet image post-processing through neural network models for color reference chart detection," *Applications in Plant Sciences*, p. Accepted. , 2020.
- [4] A. Boutchuen, D. Zimmerman, A. Arabshahi, J. Melnychuk, and S. Palchoudhury, "Understanding nanoparticle flow with a new in vitro experimental and computational approach using hydrogel channels," *Beilstein J Nanotechnol*, vol. 11, pp. 296-309, 2020, doi: 10.3762/bjnano.11.22.
- [5] S. M. Tareq *et al.*, "A Dynamic Light Scattering Approach for Detection of Nanomaterials in Tennessee River," *Water Resources Research*, vol. n/a, no. n/a, p. e2020WR028687, doi: <https://doi.org/10.1029/2020WR028687>.
- [6] Q. Baldwin and E. Panagiotou, "The local topological free energy of proteins," *J Theor Biol*, vol. 529, p. 110854, Nov 21 2021, doi: 10.1016/j.jtbi.2021.110854.
- [7] H. Qin *et al.*, "Cointegration of SARS-CoV-2 Transmission with Weather Conditions and Mobility during the First Year of the COVID-19 Pandemic in the United States," *Arxiv*, 2022.

- [8] J. Wang and E. Panagiotou, "The protein folding rate and the geometry and topology of the native state," *Sci Rep*, vol. 12, no. 1, p. 6384, Apr 16 2022, doi: 10.1038/s41598-022-09924-0.
- [9] C. Sun, V. K. Kumarasamy, Y. Liang, D. Wu, and Y. Wang, "Using a Layered Ensemble of Physics-Guided Graph Attention Networks to Predict COVID-19 Trends," *Applied Artificial Intelligence*, pp. 1-24, 2022, doi: 10.1080/08839514.2022.2055989.
- [10] M. H. Tsai, N. M. Ely, and Y. Wang, "Uncertainty Estimation for Twitter Inference," in *The 9th International Conference on Computational Science and Computational Intelligence (CSCI)*, Las Vegas, U.S.A., 2021 (accepted).
- [11] D. Ledesma, C. Powell, and H. Qin. "The iCompBio Python Coding Bootcamp." <https://github.com/ICompBioUTC/Python-Bootcamp> (accessed December 31, 2021).
- [12] H. Qin. "GitHub Public Repositories on R-programming in Computational Biology." <https://github.com/hongqin/RCompBio> (accessed May 1, 2021).