

Title: Request for Compute Resources to Support the Workshop in Molecular Evolution at the Marine Biological Laboratory at Woods Hole, MA, USA, 2020

Field of Science: Molecular Evolution and Phylogenetics (best description, but not a choice)

Principal Investigator: Paul O. Lewis (CV attached)

Principal Co-Investigator: Peter Beerli (CV attached)

Abstract: The Workshop in Molecular Evolution (MOLE) has been held every summer at the Marine Biological Laboratory (MBL), Woods Hole, Massachusetts, USA, since it began in 1988. MBL records indicate that 1884 scientists have been trained in this workshop to date, which focusses on the theory and computational methods in the fields of molecular evolution, comparative genomics, population genetics, and phylogenetics. The scale of modern genomic data sets demands that biologists become comfortable using more powerful remote computing resources. Many participants experience their first exposure to bash scripting and command line interfaces in the MOLE workshop, and our ability to help students over this hurdle is increasingly hampered by the aging and speed-heterogeneous cluster housed at MBL (which has been graciously provided for our use in past years by a research lab at MBL - it is not an MBL-wide resource). We are proposing to use Jetstream instances provided by Xsede in the workshop this coming year. Having a reliable and homogeneous computational environment will make our job much easier, and will make 60 participants (mostly senior graduate students and postdoctoral researchers) aware of the existence and usefulness of Xsede. Scientific computing in the future will be increasingly cloud-based, and learning about how to loan a cluster for your computing needs instead of purchasing/maintaining your own cluster is thus an important additional lesson that we can add to this venerable workshop. All lectures and computer lab tutorials presented at this workshop are openly available via the course web site.

Syllabus:

Schedule from the 2019 Workshop in Molecular Evolution:

<https://molevol.mbl.edu/index.php/Schedule>

Tentative schedule for the 2020 Workshop in Molecular Evolution (new web site):

<https://molevolworkshop.github.io/schedule/>

A PDF copy of the above 2020 workshop webpage is attached in case the online syllabus is not acceptable.

Resource Justification:

We are requesting 120 m1.quad instances (each with a unique IP address, 4 cores, 10 GB RAM, and 20 GB local storage) for 18 days, i.e. 207,360 service units:

$207,360 \text{ SUs} = (120 \text{ VMs}) \times (4 \text{ cores/VM}) \times (24 \text{ hours/day}) \times (18 \text{ days}) \text{ core-hours}$

There are 60 participants and 26 faculty (including teaching assistants), and the workshop duration is 11 days, May 31 through June 10, 2020. Although the directors and teaching assistants are familiar with high performance computing, none of us have used cloud-based computing for this workshop previously. We are thus requesting the equivalent of 1 extra week than actually needed during the workshop to allow the directors and TAs to fully test the workshop materials ahead of time and become comfortable with how the system works.

Requested End Date:

June 18, 2020