

Selection of Sponsor and Institution:

Selection of Sponsor:

I rotated in Peter Kharchenko's lab the summer prior to the start of my graduate program. I worked closely with Dr. Kharchenko to build upon his previous method for single cell differential expression, published in Nature Methods, in order to identify cell subpopulations based on gene expression. This was my first introduction to single cell technologies and I was quickly enthralled with the statistical and computational methods behind the single cell analysis and enticed by the opportunities presented in this emerging field. As his first, and currently only graduate student, I receive a lot of personalized attention and hands-on mentorship. For example, Dr. Kharchenko takes time to sit down and debug with me. I have already improved immensely as a programmer simply by observing Dr. Kharchenko and programming with him. Dr. Kharchenko and I have also attended conferences, written grants, and reviewed manuscripts together. I am confident that Dr. Kharchenko will be able to provide the mentorship and guidance I need to accomplish the goals in this proposal.

Selection of Co-Sponsor:

Dr. Kharchenko completed his graduate research in George Church's lab. I have taken an ethics course with Dr. Church and find him to be an inspiring visionary with ethical character. I believe Dr. Church will provide guidance to both myself and Dr. Kharchenko in the execution of this proposal.

Selection of Institution:

The Bioinformatics and Integrative Genomics (BIG) program at Harvard University is one of the premier places to train as a bioinformatics graduate student. Based at the Center for Biomedical Informatics (CBMI) at Harvard Medical School (HMS), BIG provides interdisciplinary training in biological as well as quantitative methods. HMS offers additional opportunities for training outside the classroom, ranging from seminars to career panels and workshops. The Harvard medical area fosters a collaborative atmosphere between doctors and researchers and provides ample access to high quality datasets.

In addition, HMS provides access to significant computing resources in the shared research cluster ("Orchestra"), which was recently expanded with a three million dollar ARRA grant. The cluster comprises more than 400 Linux compute nodes and 4500 cores, provides access to 10TB of disk space on a high performance Isilon storage cluster for ongoing computational tasks, and allows for multiple job submissions via Platform Computing's LSF resource management system, thereby allowing for powerful computations to be completed quickly. Multiple core facilities in the Harvard Medical Area including the Broad Institute provide access to next-generation sequencing platforms including the Fluidigm C1 machine for singlecell RNA sequencing.