**Qingpeng Zhang**
*Department of Computer Science and Engineering*
*Michigan State University*
*East Lansing, MI 48824*
🕾 *+1 (678) 901 9911*
✉ *qingpeng@msu.edu*

**DOE Joint Genome Institute**                             November 06, 2014
*2800 Mitchell Drive*
*Walnut Creek, CA 94598*

Dear Search Committee,

I am excited to apply for the Simons Institute Postdoc- Bioinformaticist Postdoc Fellow position at JGI (JobID:80367). I am strongly interested in developing and applying computational methods to guide large scale efforts of using sequencing technologies as a tool to answer biological questions. I have more than ten years of experience in bioinformatics and have been working with metagenomic data as part of my PhD. I specialize in developing scalable tools for processing and analyzing complex shotgun metagenomes. Currently, I am a graduate student in Dr. C. Titus Brown's group at Michigan State University and expect to receive my doctorate this May.

In the past several years, I have helped starting and participated in an effort to develop an efficient k-mer counting software package based on a probabilistic data structure, Count-Min Sketch (`https://github.com/ged-lab/khmer`) with C. Titus Brown. We also put a lot of effort into making the khmer paper reproducible with an automatic pipeline (`https://github.com/ged-lab/2013-khmer-counting`). This software package has been applied to study the diversity of multiple metagenomic datasets, including the Great Prairie Grand Challenge, a collaboration with the JGI. Based on the work on khmer, I also helped developing digital normalization, a single-pass computational algorithm to discard redundant shotgun sequencing data to enable more efficient analysis. Specifically, I integrate efficient k-mer counting and read coverage analysis based on digital normalization to generate abundance profiles across multiple datasets, allowing for scalable diversity analysis of large, complex metagenomes without the need for assembly or reference sequences. I have evaluated this method on multiple metagenomes from a variety of environments (e.g., human gut, soil in collaboration with James Tiedje, ballast water viromes in collaboration with Joan Rose ). Going forward, I plan to extend the application of this method on binning approaches, functional annotation, and phylogenetic analysis for metagenomic analysis. Given the velocity in growth of sequencing data, I believe that this method is promising to highly diverse samples with relatively low computational requirements. Further, as they do not depend on reference genomes, these methods also provide opportunities to tackle the large amounts of unknown "dark matter" we find in metagenomic datasets.

I have an extensive history of working with sequencing data (since 2003). Previously, I worked at the Beijing Genomics Institute (BGI) where I participated in genome sequencing projects and later on systems biology such as regulatory network analysis and microRNA prediction. These experiences, combined with my PhD research and my dual major in computer science and quantitative biology, have given me a strong background in bioinformatics, molecular biology, and genomics. I enjoy interacting with a variety of collaborators, including biologists, statisticians, bioinformaticians, and software engineers.

My collaboration with the JGI has introduced me to the breadth of projects on-going, the diverse datasets to analyze, the broad biological questions to answer at the JGI. With the support of the Simons Fellowship and the participation of the semester-long program in Algorithmic Challenges in Genomics at Simons Institute, I will get exposed to the recent exciting progress of bioinformatics algorithms development and get the chance to work with many experts in different fields. I believe all these opportunities will facilitate my work at JGI considerably. All these reinforce my impression that working at JGI will offer me numerous opportunities to pursue my passion, which is to apply my skills and knowledge in computing to answer scientific questions. I am greatly enthusiastic about the possibility of contributing my skills and knowledge to facilitate the

scientific mission of the JGI as well as the Simons Institute. I have attached my CV for your review. I would greatly appreciate the opportunity to talk to you more about the position. Please do not hesitate to contact me with any questions.

Sincerely,


**Qingpeng Zhang**

*Attached: curriculum vitæ*