Achal Neupane
Date: October 1, 2019

Dear Search Committee:

I am writing to express my interest in the advertised Post Doctoral Research Associate - Computational Biology position at University of Nebraska-Lincoln. I am a PhD candidate at South Dakota State University (Brookings, SD, USA) and will be graduating with a specialization in bioinformatics in May, 2020. I have previously worked on a variety of bioinformatics projects at both US and Australian universities and have developed a strong interdisciplinary academic background. I believe these experiences have allowed me to acquire a diverse set of knowledge and skills that make me a suitable candidate for a bioinformatics research position.

I am currently working on my PhD dissertation, entitled Dissecting RNA Silencing Pathways in the White Mold Fungus, *Sclerotinia sclerotiorum* with Dr. Shinyi Marzano at South Dakota State University. Specifically, I have been using gene disruption mutants to identify the roles of dicers (dcl-1 and dcl-2) and argonaute enzymes (agl-2 and agl-4) in the small RNA (sRNA) metabolism of the plant pathogenic fungus, *Sclerotinia sclerotiorum*. I have also confirmed the effectiveness of ds-RNA based pesticides that target the *S. sclerotiorum* RNA silencing pathway and have analyzed the diversity and evolution of mycoviruses identified from the whole root metatranscriptomes of various plant host species infected by arbuscular mycorrhizal fungi. Several articles on this research have already been published in peer-reviewed journals.

Prior to my PhD, I earned a BS in Biological Science from University of DC (Washington DC, USA) in 2009 and an MS degree in Biology (Bioinformatics) from South Dakota State University in 2013. Following the completion of my master's degree, I worked on several projects focused on bioinformatics and statistical genetics. In 2015, I accepted an opportunity to work in Australia at the University of Queensland-Translational Research Institute (TRI), where I worked as a bioinformatics researcher in Dr. Matthew Brown's group. The goals of this group were to characterize the genetic origins of disease and translate these findings into clinical practice. I was primarily involved in the characterization of the genetic landscape of Acute Myeloid Leukemia (AML), and, to that end, I analyzed sequence genotyped SNP data from a cohort of 150 clinically characterized AML samples sequenced with whole exome and whole genome sequencing. I also analyzed a cohort of 900 control exomes and 600 whole genomes to validate and compare genotyping algorithms and sequencing technologies and to perform gene discovery for both somatic and germline risk variants. For this latter work, I used statistical genetics algorithms that identified both protective and deleterious variants. I then compared pathology, cytogenetic, and Sequenom genotyping data to refine and calibrate these algorithms. I also used several variant calling methods to characterize germline, copy number, and structural variants in AML.

While working at TRI, I also developed an algorithm to quality control Next Generation Sequencing (NGS) data in order to estimate pair-wise identity by descent (IBD) probabilities from high-density single nucleotide polymorphism (SNP) data. I demonstrated that the estimated IBD probabilities effectively identified contaminated, related, and distantly-related samples. This

method could be used to quality control NGS genotype data, including sequences produced by targeted sequencing panels. I have developed several algorithms for these purposes, and the R codes are available for free on my Github account.

I spent the initial years of my career studying human cancer cells, as well as diseases affecting plants, insects, and fungi, by analyzing NGS data (whole genome, whole exome, RNA-seq, SNPs, microbiome, etc.). Since the beginning my PhD program, I have taken several "data-heavy" graduate level courses from Statistics and Computer Science departments in order to develop my skills for big-data analysis, algorithm development, and scientific computing, including analysis of biological data. Additionally, I have been writing computer programs in R, python, Perl, JavaScript, and also have been working on Linux/Unix system with extensive use of shell (bash) and PBS scripts for nearly 10 years of my research career. Therefore, for my post-doctoral research, I would be interested in joining a research lab where I could use various statistical analyses, including machine learning, and artificial neural networks to study the genomics of various diseases.

This cover letter provides only a brief synopsis of my background and experiences. I would be glad to meet and further discuss my qualifications. I have enclosed my CV along with this letter.

Thank you very much for your time and consideration, and I look forward to hearing from you.

Sincerely,

Achal Neupane
PhD. Candidate
Department of Biology and Microbiology
South Dakota State University
Email: achalneupane@gmail.com
Phone: 2026298640
Website: achalneupane.github.io