Achal Neupane
Date: September 30, 2019

Dear search committee:

I am writing with the high interest and keenness for the advertised "Postdoctoral Associate" position. My dissertation work looking at both two-way and three-way interactions among host, fungal pathogen and mycoviruses, and my previous experiences in the US and Australian universities in the study of acute myeloid leukemia (blood cancer), including diseases of plants, insects and fungi have set me with highly diverse experiences in bioinformatics research. Additionally, my strong interdisciplinary academic background, and diverse research experiences prepare me with essential set of knowledge and skills as a suitable candidate for a bioinformatics research position.

I will be graduating with a PhD degree with specialization in bioinformatics from South Dakota State University (SDSU) in May, 2020. Currently, I have been working on my PhD dissertation (Dissecting RNA silencing pathways in white mold fungus *Sclerotinia sclerotiorum*) with Dr. Shinyi Marzano. I have been studying the roles of dicers (dcl-1 and dcl-2) as well as argonaute enzymes (agl-2 and agl-4) in small RNA (sRNA) metabolism in a plant pathogenic fungus *Sclerotinia sclerotiorum*. These gene disruption mutants were compared for changes in phenotype, virulence, viral susceptibility, gene expressions and sRNA profiles (miRNA, siRNA, tRNAs, including virus-derived sRNAs). Recently, I also confirmed the effectiveness of ds-RNA based pesticides targeting *S. sclerotiorum* RNA silencing pathway for the control of the economically important pathogen. Besides, I have also studied the diversity and evolution of mycoviruses identified from whole root metatranscriptomes from different plant hosts infected by arbuscular mycorrhizal fungi. Some of these studies have already been published in peer-reviewed journal articles. Additionally, I have been working on different microbiome projects to study the effects of crop rotation as well as soil salinity in soil microbiome diversity affecting the overall yield by analyzing 16S and ITS data.

Prior to my PhD, I earned an MS degree in Biology (Bioinformatics) from South Dakota State University (Brookings, SD, USA) in 2013 and a BS in Biological Science from University of DC (Washington DC, USA) in 2009. Since then, I have been working on several bioinformatics and statistical genetics focused projects. I later got an opportunity to work in Australia at the University of Queensland-Translational Research Institute (TRI), which is the third largest genotyping center behind the Broad and Sanger institutes. I decided to work there purely based on my interest to work on cancer genomes and genotyped SNP data. I also thought this would be a great opportunity to gain experience with Australian education/research system and learn highly technical skills required to handle human genomes/exomes and genotyped SNPs data. At UQ, I worked as a bioinformatics researcher in professor Matthew Brown's group where I gained key expertise in the analysis of large clinically focused datasets. The goals of these research were to characterize and understand the genetic origins of disease and to translate these findings into clinical practice. I was primarily involved in the characterization of the genetic landscape of Acute Myeloid Leukemia (AML) analyzing 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, I used sophisticated statistical genetics algorithms that identify 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.

Additionally, I have developed an algorithm and demonstrated a method to Quality Control Next Generation Sequencing (NGS) data to estimate pair-wise identity by descent (IBD) probability from high-density SNP data. Estimated IBD probability were demonstrated to effectively identify contaminated, related and distantly related samples, including to determine sample ethnicities. The implications of this method are in quality control of NGS genotype data, including those sequenced by targeted sequencing panels. I have developed several algorithms for these studies and the R codes are freely available on my github account.

Over the years, I have studied human cancer cells, and diseases affecting plants, insects and fungi analyzing NGS data (whole genome, whole exome, RNA-seq, SNPs, microbiome, etc.). During my PhD program, I have also taken several "data-heavy" graduate level courses from Statistics and Computer science departments to develop my skills for big-data analysis, algorithm development, scientific computing, including analysis of biological data. For my post-doctoral research, I would be interested in large-scale genome study of diseases (such as GWAS). I would be interested in joining a research group where I would also learn to use various statistical approaches using machine learning and artificial neural network analysis to study these diseases.

This cover letter only provides a brief synopsis of my background and experiences. I would be happy to meet and discuss my qualifications in detail for the position. I have enclosed my CV along with this letter.

I am looking forward to hearing from you. Thank you very much for your time and consideration.

Sincerely,

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