

*September 27, 2020***GRAIL**1525 O'BRIEN DRIVE
MENLO PARK, CA 94025

Dear Hiring Manager,

I am an first year MSc. student in Computational Biology at Carnegie Mellon University applying for the bioinformatics summer internship.

Cancer is to vast and complex a beast to try and detect, let alone detect early and robustly, you need equally vast and complex data to study it. I encountered this first hand during an internship in the Michael Hoffman lab at the University of Toronto under the supervision of Post-Doc and placenta expert Samantha Wilson. She conceptualized our project of comparing transcriptome profiles of various cancers and the placenta, trying to find what the placenta was regulating "correctly" that cancers cells were not. Specifically I collected and cleaned the data, preformed batch correction, differential gene expression, gene ontology enrichment, survival analysis for over 20,000 cancer and placenta samples. Doing all of these things required reading up about many statistical methods as well as applying and visualizing them using bioconductor, ggplot and some stand alone tools. I also spent much of my time visualizing my results and presenting them to experts and non-experts where the placenta and cancers differ and what that could imply for future cancer research.

I also have some experience with pipeline development from my undergraduate thesis. I was designing an end-to-end pipeline to explore the relationship between rates of horizontal gene transfer and the presence of CRISPR-Cas systems in bacteria. This project involved both significant software development and statistical considerations for the final analysis and write-up. I wrote scripts to retrieve the data for all complete bacterial genomes on NCBI, produce gene and species trees for each species and preform network analysis. Since I was working with several thousand genomes I also needed to make sure that the code was relatively efficient and produced clear progress reports and organized output files. Finially I presented my results and thier biological implications in the form a a manuscript.

I know my educational background in genetics, experience characterizing cancers and enthusiasm for intricate biological problems makes me especially suited for this internship.

Thank you for your time and consideration.

Sincerely,

Siddharth ReedMSc. Student in Computational Biology
Carnegie Mellon University