**BIOC4966 Biochemistry Internship Report**

**Research Project Title: LINE1 type transposase domain containing 1 (L1TD1) gene expression in colorectal (CRC) adenoma**

By RHO, Yerim

**Supervisor:** Dr. Wong, Jason Wing Hon | **Co-supervisor:** Dr. Chung, Clive Yik Sham

**Acknowledgements**

First and foremost, I would like to express my sincere appreciation to my supervisor, Dr Jason Wong, for his continuous support and guidance throughout the internship project duration. The opportunities he has provided me to conduct bioinformatics projects allowed me to expand my understanding and appreciation of cancer genomics research. I am immensely grateful for his encouragement and advice, not only in the project but also in my academic pursuits. I am also sincerely thankful to my co-supervisor, Dr Clive Chung, for allowing me to perform research outside the Biochemistry department and to broaden my research interest. My utmost appreciation goes to my mentor, Sojung Sandy Lee, for her assistance, kindness, and patience. I would not have successfully carried out this project without her support and I will be grateful evermore. I would also like to thank all the members of Jason lab, for their kindness and inspiration during my time in the lab. Last but not least, thank you to my family and friends for endless support, love, and presence.

RHO, Yerim

**Introduction**

The internship was held in Dr Jason Wong’s lab at the School of Biomedical Science at The University of Hong Kong. Jason’s laboratory specializes in cancer genomics research to investigate the effect of mutations on the way genes are regulated. My project focused on gene expression of LINE 1 type transposase domain containing 1 (L1TD1) in colorectal (CRC) adenoma. During the internship, I utilized bioinformatics and computational techniques to analyze existing large genomics datasets mainly from Gene Expression Omnibus (GEO). The techniques included bulk RNA-sequencing (RNA-seq), single-cell RNA-sequencing(scRNA-seq), motif analysis, gene set enrichment analysis (GSEA), and DNA methylation analysis. The majority of the results of analysis on bulk RNA-seq revealed upregulation of L1TD1 gene expression distinctly in adenoma of colorectal cancer. Thus, various hypothesis was formulated to interfere the molecular mechanism behind the regulation of gene expression of L1TD1 using bioinformatics tools.

**Link to science & society**

Cancer genomics research is immensely significance for its positive impact on the role of science in society. Academic cancer research can lead to differences in the way of health care for the human body to maintain a better quality of life. By understanding the alteration of the development of cancer at the genomic level, relevant genes and pathways involved in tumorigenesis can be identified. Not only applicable to clinically treat cancer patients but also prognosis prediction would be developed when advanced understanding of fundamental genetic processes.

Throughout the internship, I got a new insight into the field of bioinformatics on such efforts to detect useful information from large genomics datasets. The comprehensive bioinformatics analysis to identify potential prognostic biomarkers includes the availability of big omics data. Traditionally, developing a pipeline for new drugs was a lengthy process. The large-scale dataset contains potentially diverse forms of information, some are yet to be revealed. Bioinformatics analysis, like integration and machine learning algorithms, can shorten the time to discern meaningful associations.

Despite advancing the science, cancer genomics research in academia may raise ethical issues for society due to their perceived moral implications. However, the findings in the research laboratory can suggest other clinical approaches to the genes of cancer patients, such genetic editing techniques may be viewed as ethically controversial.

**Opportunities for problem solving**

During my internship, the greatest challenge I faced was discerning the relevant dataset from the vast pool of available big data to align with my research goal. I felt inadequate when I spent a weighty amount of time searching for datasets, thinking I needed to engage more actively in the analysis skills I had learned. However, upon obtaining the results, I realized the significance of the time invested in preparing the details before conducting the analysis. As a result, I learned that the preparation phase was valuable and contributed to my insight in identifying relevant datasets for future research.

In addition, I became aware of the bias I developed toward my research due to spending excessive time on the same topic. In one of my meetings with Dr Jason Wong, he emphasized the challenges of formulating and testing hypotheses. Recognizing the need to break free from my bias, I realized the importance of actively seeking diverse perspectives through establishing biological foundations or engaging in communications with others in the lab. In the future, I aim to enhance my skills by dedicating more time to establishing solid research foundations, formulating hypotheses, and developing communication skills.

**Personal Reflection**

I approached the internship with a learner’s mindset when I joined Jason’s lab. My primary goal was to determine if the dry lab environment suited me, considering my lack of experience not only in research labs but also in the field of bioinformatics. To effectively achieve this goal, I utilized applications such as Notion and GitHub to organize written records of details. This proved invaluable when I encountered challenges in applying the skills I had acquired during my time in the lab. Especially, despite my lack of prior coding knowledge, my written records served as a valuable resource whenever I encountered coding obstacles. It allowed me to discern patterns by referring back to past code.

I discovered a genuine enjoyment in this learning experience when I could visualize the outcomes of my analyses and independently construct code. This enjoyment inspired me to adopt a mindset of curiosity, always seeking to ask questions rather than provide answers. Also, I came to recognize the importance of continuously enhancing my knowledge rather than focusing on teaching others. This realization has motivated me to embrace the role of a researcher and continuity in the bioinformatics field.

Upon reflection, I have recognized a particular weakness in my performance during the internship. I noticed that I tended to rely heavily on my mentor, constantly seeking her instructions. Although I understood it was my first experience in a research lab environment, I believed I could have shown more confidence in my creative thinking process. For instance, I often found myself preoccupied with completing tasks my mentor and professor instructed, wanting to showcase my progress and dedication. In the future working environment, I will prioritize the development of my creative thinking skills, ensuring that I take ownership of the project and recognize the significant role I play in its success.

**Conclusion**

In conclusion, my internship was transformative, providing me with a clear picture and deepening my desire to continue expanding my knowledge in the field of cancer genomics as a researcher. I found joy in visualizing outcomes and coding independently, embracing curiosity through continuous learning. I want to continue contributing to the scientific community, reaffirming my passion for bioinformatics and commitment to ongoing research. I am grateful to have this internship experience during my undergraduate studies.