NANYANG TECHNOLOGICAL UNIVERSITY
SINGAPORE

School of Biological Sciences
College of Science

Reg. No. 200604393R

Research Theme: Computational Biology; Functional Genomics

MSc Research Project Title: **Expression Quantitative Trait Loci (eQTLs) prediction optimisation and multi-data integration with machine learning to target novel cis-regulatory relationships in cancer**

Principal Investigator/Supervisor: **Asst/Prof Li Yinghui**

Co-supervisor/ Collaborator(s) (if any):

Project Description

a) Background:

The accumulation of mutations throughout the genome is responsible for igniting major diseases like cancer. Mutations within the protein coding part of the genome are well known to affect key genes acting as tumour suppressors and oncogenes. However, most of the regulatory mechanisms influencing the expression of these genes is encoded in the non-coding portion of the genome. Therefore, mutations occurring at regulatory elements several tens of thousands of bases away from their target genes can influence one or more cancer relevant genes through diverse regulatory relationships. As the world progresses towards personalised medicine, many genome wide studies identifying non-coding DNA variants in cancer patients provide valuable resources which can be combined with expression data to detect correlative trends to help identify and characterise the non-coding mutations disrupting oncogenic regulatory relationships.

b) Proposed work:

The successful candidate will apply sophisticated computational and statistical methods to identify critical links between genotypic and transcriptional variations to help elucidate the epigenetic regulatory mechanisms (ERMs) underlying the progression of myeloma and lymphoma cancers. Large processed DNA and RNA-seq datasets will be available that can be used for expression Quantitative Trait Loci (eQTLs) analyses. The results will be further refined and integrated with an array of epigenomics data generated from cell line models or patient derived cells including eRNA transcription datasets using machine learning methods. This will provide an exciting opportunity to consolidate candidate QTLs through orthogonal data sources and better characterise their nature in coordination with senior colleagues.

c) Preferred skills:

To progress at a desirable pace the candidate will be expected to be familiar with R and possess solid statistical foundations. Experience with Python or R machine learning techniques is also beneficial as well as general command line usage. Good organisational skills are essential as you will be required to record and present your research regularly, ideally using Jupyter notebooks.

Supervisor contact:

If you have questions regarding this project, please email the Principal Investigator:
{ **liyh@ntu.edu.sg** }

SBS contact and how to apply:

Associate Chair-Biological Sciences (Graduate Studies) : AC-SBS-GS@ntu.edu.sg

Please apply at the following:

Application portal:

https://venus.wis.ntu.edu.sg/GOAL/OnlineApplicationModule/frmOnlineApplication.ASPX