NANYANG
TECHNOLOGICAL
UNIVERSITY

School of Biological Sciences

Reg. No. 200604393R

**Research Theme: Computational Biology; Functional Genomics; Cancer Biology**

# Research Project Title: Biomarker prediction of cancer based on big data and network analytics

**Principal Investigator/Supervisor: Wilson Wen Bin Goh**

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

## Project Description

### a) Background:

Cancer is a heterogeneous disease that comprises of diverse subtypes harboring different genetic dependencies. Different cancers rely on distinct sets of genes that can interact with or function together to drive the aberrant activity of various signaling pathways essential for their survival and proliferation. Recent advances in whole-genome sequencing have led to the identification of multiple genetic abnormalities associated with distinct cancer types. These mutations frequently occur in non-coding, regulatory DNA including gene promoters, enhancers, silencers and insulators, which regulate the expression of cancer driver genes via transcriptional or epigenetic mechanisms.

High-throughput biology, in the form of the various -omics platforms, have given rise to biological big data. Yet, big data is of little value without robust analytical approaches. Approaches utilizing biological context in the form of biological networks is a promising approach for integrating the various -omics platforms (multi-omics), improving signal identification and also analytical reproducibility.

This project aims to identify and validate predicted markers based on network-based approaches. In addition, we wish to use these networks to characterize the molecular drivers responsible for disease progression.

### b) Proposed work:

This project requires an in-depth understanding and application of computational biology (*in silico*).

You will begin with analytical pipeline development for understanding irreproducibility and random signature superiority issues in cancer study. By optimizing and developing a logical set of normalization, cleaning and statistical procedures, we will come to a consensus on the conserved gene or subnetwork set.

You will then validate these predictions using molecular biology techniques in both cell lines and cancer models.

### c) Preferred skills:

Highly motivated candidates with strong computing skills. Knowledge and proven competence in

| |
|---|
| machine learning, data modelling and/or biomedical science is highly desired. |
| **Supervisor contact:**<br>**If you have questions regarding this project, please email the Principal Investigator:** |
| **SBS contact and how to apply:**<br>Associate Chair-Biological Sciences (Graduate Studies) : AC-SBS-GS@ntu.edu.sg<br>Please apply at the following:<br>http://admissions.ntu.edu.sg/graduate/R-Programs/R-WhenYouApply/Pages/R-ApplyOnline.aspx |