IITB-Monash Research Academy

Indian Institute of Technology Bombay | MONASH University

An Indian-Australian research partnership

Project Title:	**Application of Artificial Intelligence in human genomics**	
Project Number	**IMURA0806**	
Monash Main Supervisor (Name, Email Id, Phone)	Yuan Fang Li yuanfang.li@monash.edu	*Full name, Email*
Monash Co-supervisor(s) (Name, Email Id, Phone)		
Monash Head of Dept/Centre (Name,Email)	Alan Dorin Alan.Dorin@monash.edu	*Full name, email*
Monash Department:	Clayton school of IT	
Monash ADRT (Name,Email)	Sue Mckemmish sue.mckemmish@monash.edu	*Full name, email*
IITB Main Supervisor (Name, Email Id, Phone)	Ganesh Ramakrishnan (ganesh@cse.iitb.ac.in)	*Full name, Email*
IITB Co-supervisor(s) (Name, Email Id, Phone)		*Full name, Email*
IITB Head of Dept (Name, Email, Phone)	Uday Khedkar (head.cse@iitb.ac.in)	*Full name, email*
IITB Department:	Department of Computer Science and Engineering	

Research Clusters:

Highlight which of the Academy's CLUSTERS this project will address? *(Please nominate JUST **one.** For more information, see www.iitbmonash.org)*	
1	*Material Science/Engineering (including Nano, Metallurgy)*
2	*Energy, Green Chem, Chemistry, Catalysis, Reaction Eng*
3	*Math, CFD, Modelling, Manufacturing*
4	*CSE, IT, Optimisation, Data, Sensors, Systems, Signal Processing, Control*
5	*Earth Sciences and Civil Engineering (Geo, Water, Climate)*
6	*Bio, Stem Cells, Bio Chem, Pharma, Food*
7	*Semi-Conductors, Optics, Photonics, Networks, Telecomm, Power Eng*
8	*HSS, Design, Management*

Research Themes:

Highlight which of the Academy's Theme(s) this project will address? *(Feel free to nominate more than one. For more information, see www.iitbmonash.org)*	
1	*Advanced computational engineering, simulation and manufacture*
2	*Infrastructure Engineering*
3	*Clean Energy*
4	*Water*
5	*Nanotechnology*
6	*Biotechnology and Stem Cell Research*
7	*Humanities and social sciences*
8	Design

The research problem

Define the problem

Single nucleotide polymorphisms, called SNPs (snips), are the most common type of genetic variation among people. SNPs occur once in every 300 nucleotides on average in human genome. There are roughly 10 million SNPs in the human genome, and most commonly, these variations are found in the DNA between genes. In addition to snips, DNA-rearrangements, gene silencing, and satellites are also known to cause diseases. All of these can act as biological markers, helping scientists locate genes that are associated with disease.

Project aims

Define the aims of the project

- Catalogue of genes, SNPs, and DNA elements associated with high – probability of diseases (literature review, and data base construction).
- Developing specific DNA-marker database(s) with special emphasis on cancer (/diabetics/ cardiovascular diseases): Global and Indian population SNP map, and their comparison.
- GWAS to estimate disease risk: Development and application of genome-based strategies for the early detection, diagnosis, and treatment of disease.
- Development of robust bio-informatics pipeline to study genes and DNA elements on a large scale genomic data analysis.
- Integration of big-data analytics tools and artificial intelligence (AI) for human genome analysis.
- Use of AI and machine learning (ML) in early diagnosis of diseases, through advance data processing (viz. image data analysis), and correlation with that of genome data.

Expected outcomes

Highlight the expected outcomes of the project

- Novel biomarker discovery in human disease.
- Artificial Intelligence/Machine Learning methodologies in human genomics

How will the project address the Goals of the above Themes?

Describe how the project will address the goals of one or more of the 6 Themes listed above.

The project will combine the knowledge of Artificial Intelligence/Machine Learning and human genomics and integrate them to discover novel biomarkers of disease.

Capabilities and Degrees Required

List the ideal set of capabilities that a student should have for this project. Feel free to be as specific or as general as you like. These capabilities will be input into the online application form and students who opt for this project will be required to show that they can demonstrate these capabilities.

Artificial Intelligence/Machine Learning: preferred

Expertise in computer science and engineering
Knowledge of human biology and disease: preferred
Knowledge of genomics and biology: preferred
Competencies in computational biology: preferred

Potential Collaborators

Please visit the IITB website www.iitb.ac.in OR Monash Website www.monash.edu to highlight some potential collaborators that would be best suited for the area of research you are intending to float.