Bioinformatics Institute
BII

SYSTEMS BIOLOGY AND DATA SCIENCE FOR BIOLOGICAL DIGITAL TWIN EFFORTS (?)

Kumar Selvarajoo, PhD

Senior Principal Investigator
Computational Biology & Omics Lab

12 MAY 2023

CREATING GROWTH, ENHANCING LIVES

WHAT IS A DIGITAL TWIN?

Originally described in 2002 by Michael Grieves:

The digital twin is a set of virtual information constructs that fully describes a potential or actual physical manufactured product from the micro atomic level to the macro geometrical level

The main idea is to predict undesirable or optimal outcome(s) under diverse operating conditions before actual manufacture

AEROSPACE EXAMPLE

FIRST SUCCESSFUL FLIGHT BY WRIGHT BROTHERS IN 1903; TRIAL & ERROR FOR BIPLANE DEVELOPMENT

LASTED 59s COVERING 256M

FORMALISING THEORIES GOVERING FLUID DYNAMICS

CREATING GROWTH, ENHANCING LIVES

is governed by the three fundamental laws of conservation:

i) ***mass***,
ii) ***energy*** &
iii) ***momentum***

Resulting in Navier–Stokes equations

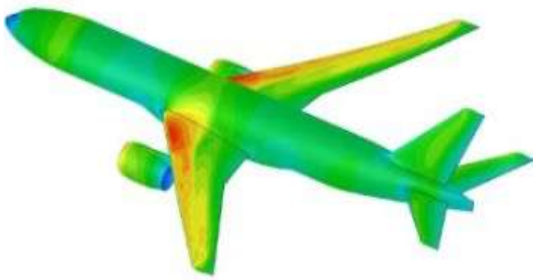
$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}.\nabla \mathbf{v}\right) = -\nabla p + \nabla.\mathrm{T} + f$$

v : flow velocity, ρ : fluid density, *p* : pressure, T : stress tensor, f : body forces/volume

CREATING GROWTH, ENHANCING LIVES

CFD DESIGNED BOEING 777

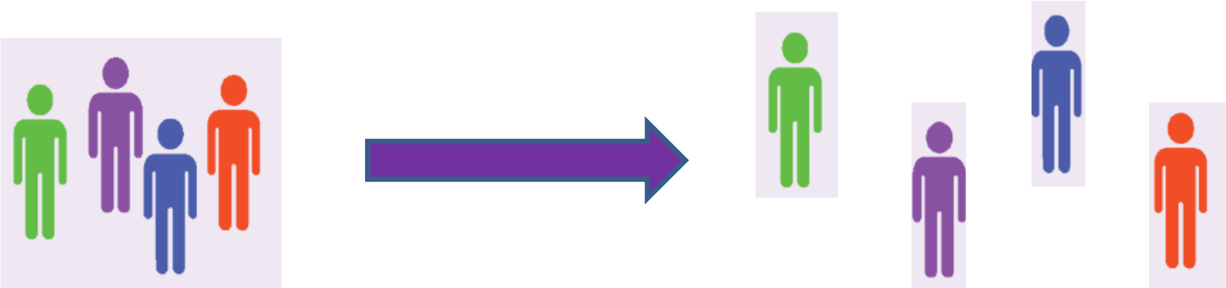
First commercial airline to be modeled and tested *in silico* using computational fluid dynamics (CFD) before production in 1991



The Impact of CFD on the Airplane Design Process: Today and Tomorrow

Ronald L. Bengelink and Paul E. Rubbert
Boeing Commercial Airplanes

CAN WE APPLY DIGITAL TWIN CONCEPT FOR BIOLOGY/MEDICINE?

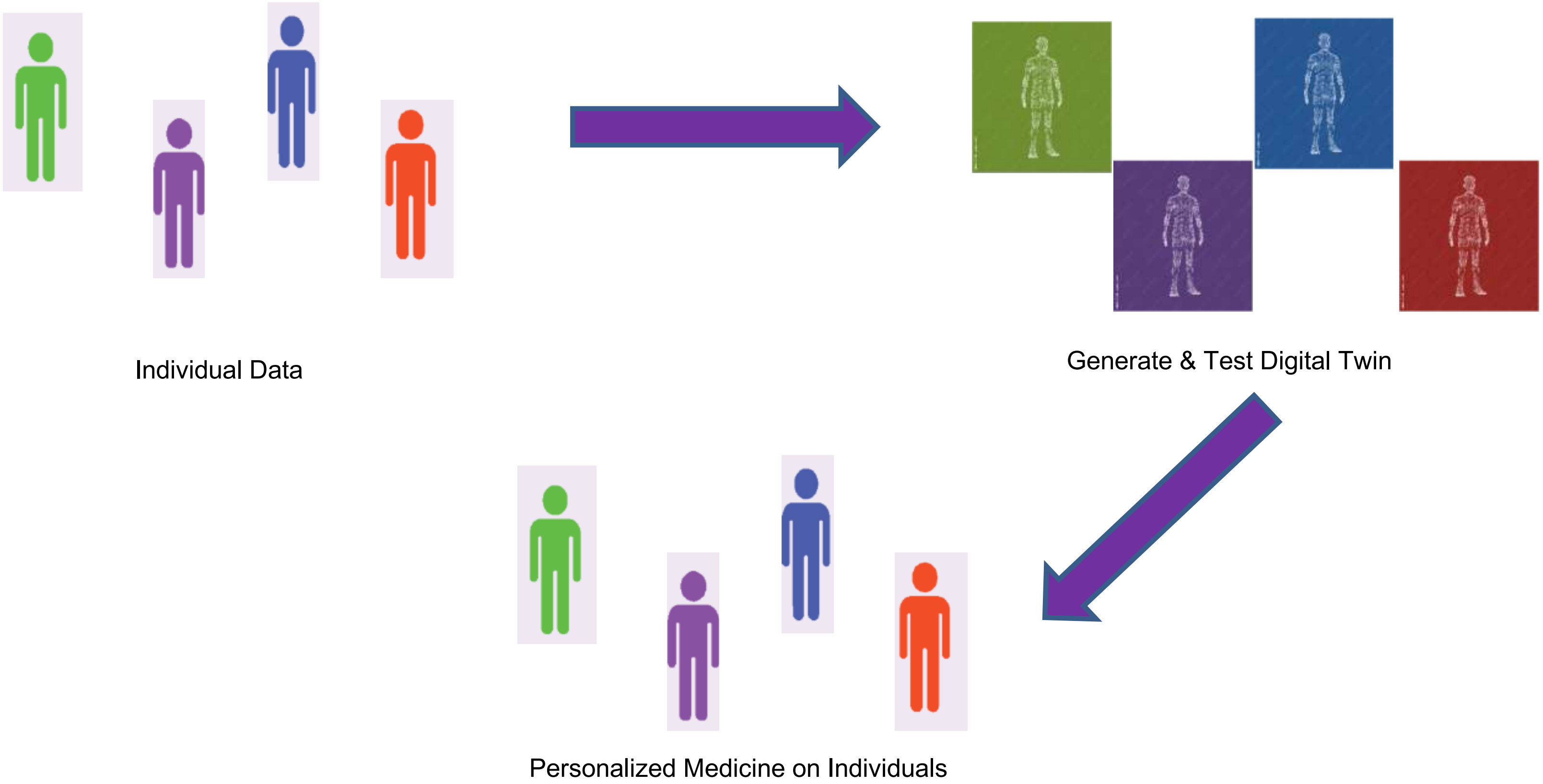


Traditional Medicine on Cohort

Personalized Medicine on Individuals

CREATING GROWTH, ENHANCING LIVES

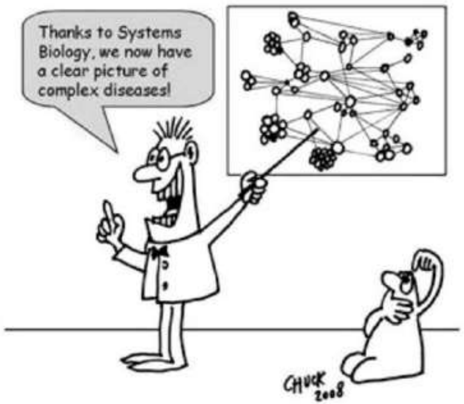
CAN WE APPLY DIGITAL TWIN CONCEPT FOR BIOLOGY/MEDICINE?



“Systems biology and medicine – not only in the lab but in the everyday lives of people – challenges the imagination and will transform the 21st Century.”

–Lee Hood (pioneer of systems biology & personalised medicine)

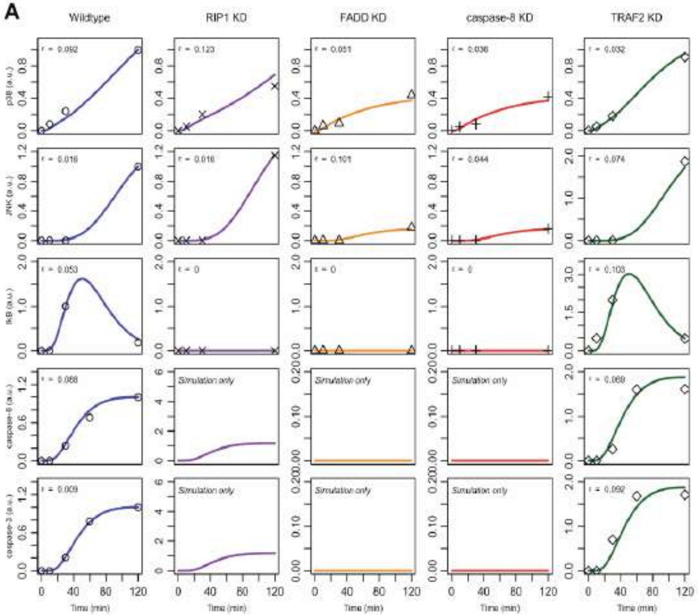
WHAT SYSTEMS BIOLOGY AND DATA SCIENCE CAN DO FOR DIGITAL TWINS?



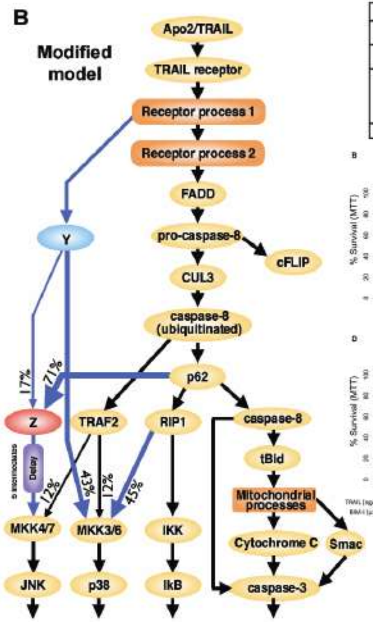
CREATING GROWTH, ENHANCING LIVES

NETWORK MODELING PREDICTS ENHANCEMENT OF APOPTOSIS IN SKIN CANCER BY PKC-Δ TARGET SUPPRESSION

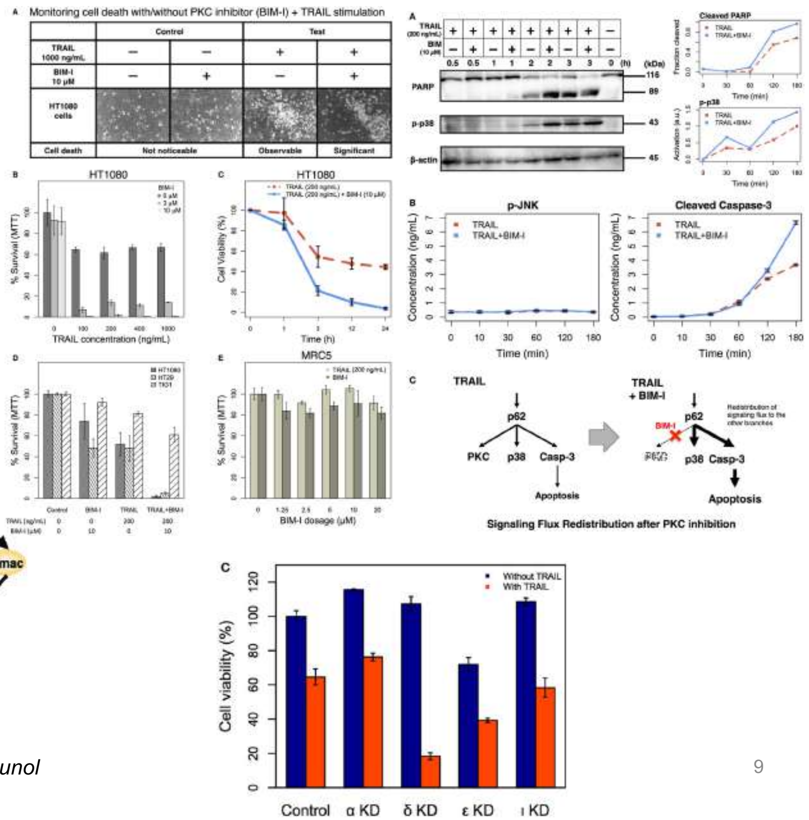
Data & Model Generation



Model Testing & Simulations



Validation of Model Prediction



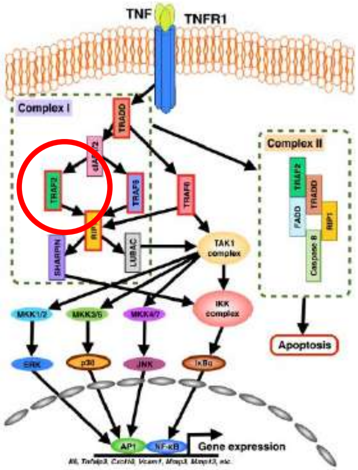
Piras V,…., Selvarajoo K (2011) *Sci Rep*

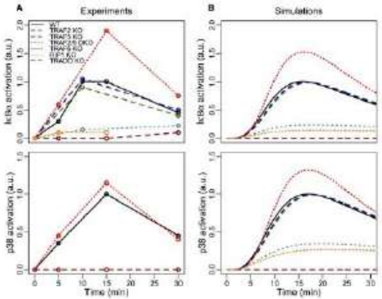
Hayashi K,…., Selvarajoo K (2015) *Front Immunol*

CREATING GROWTH, ENHANCING LIVES

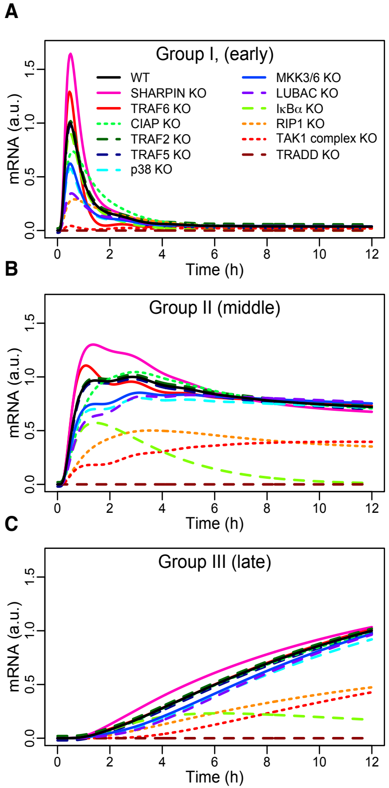
NETWORK MODELING PREDICTS TARGET *IN SILICO* TO SUPPRESS TNF-INDUCED PROINFLAMMATORY GENE EXPRESSIONS VIA RIP1

Data & Model Generation

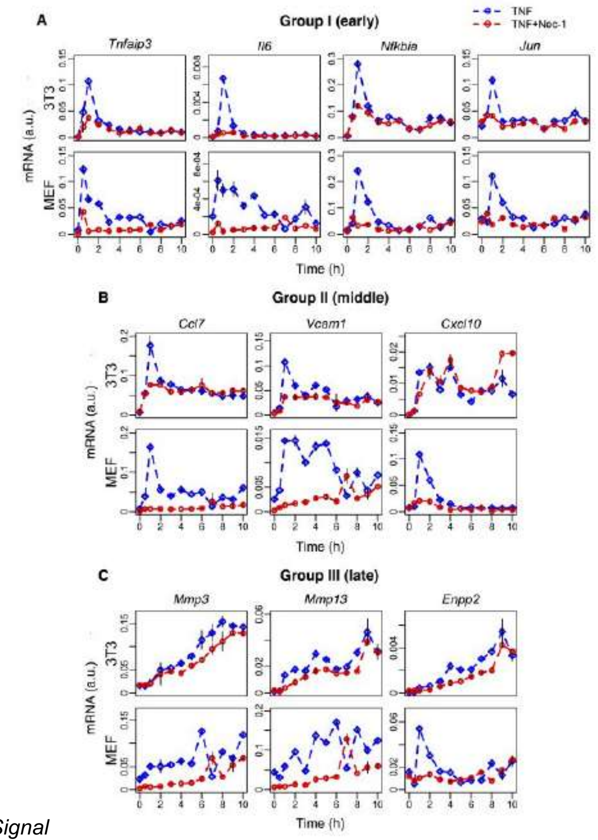




Model Testing & Simulations

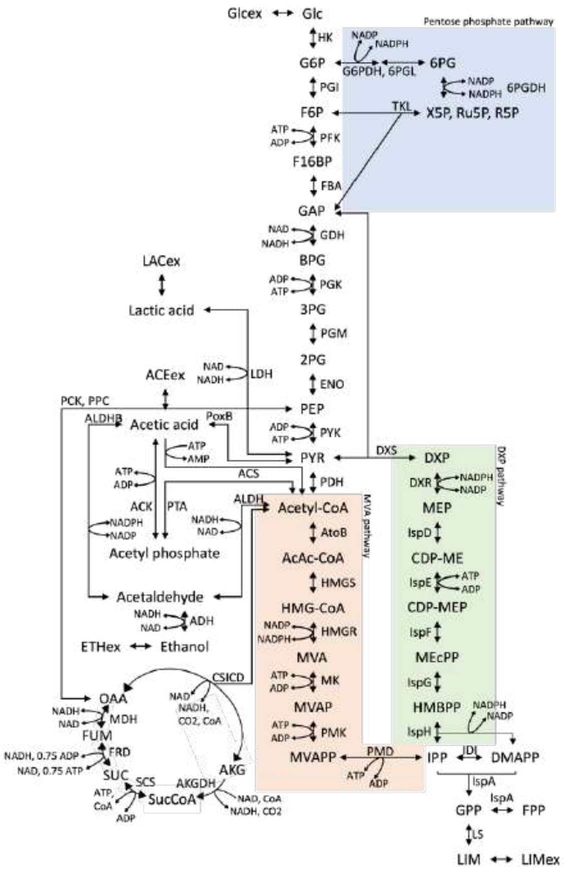


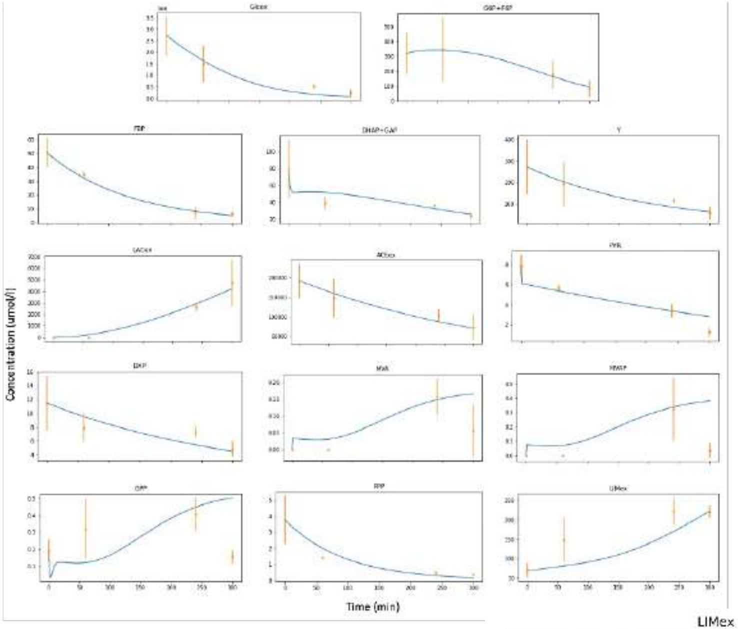
Validation of Model Prediction

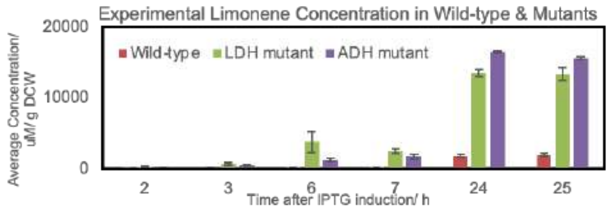


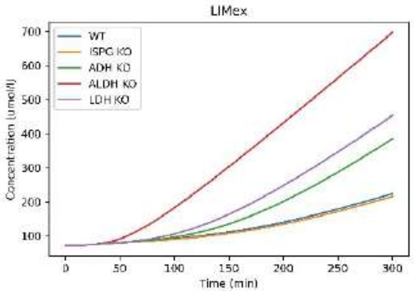
Hayashi,…,Selvarajoo (2013) *Cell Commun Signal*

RE-ENGINEERING CELLS: IDENTIFYING INTRACELLULAR TARGETS FOR LIMONENE YIELD ENHANCEMENT









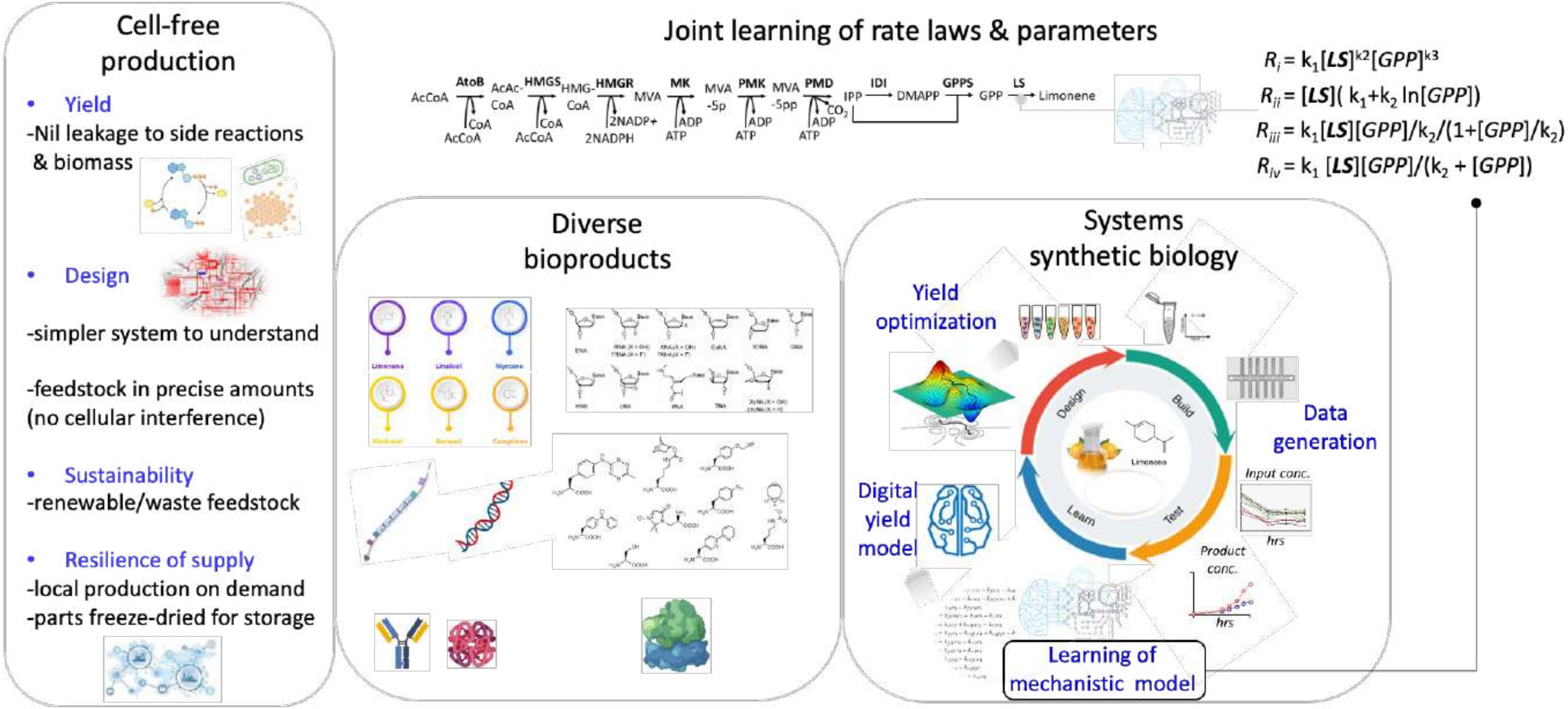
Khanijou, …., Selvarajoo K (in prep)

CREATING GROWTH, ENHANCING LIVES

CREATING GROWTH, ENHANCING LIVES

MACHINE LEARNING OF MECHANISTIC SYSTEM MODEL:

predicting optimal yield of cell factories



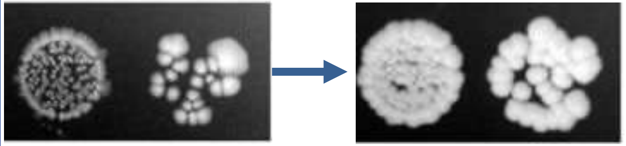
Yeo & Selvarajoo (2022) *Brief Bioinform*

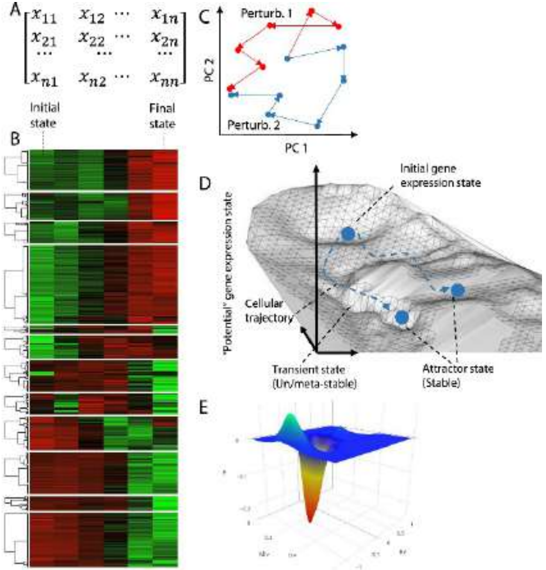
CREATING GROWTH, ENHANCING LIVES

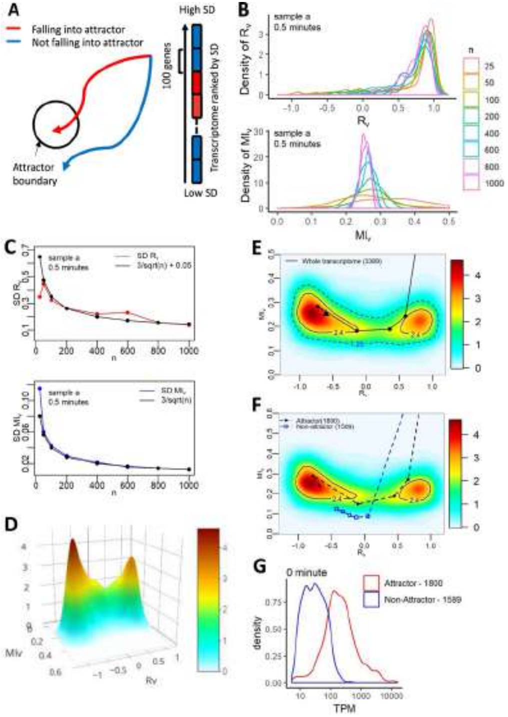
TRANSCRIPTOME-WIDE ATTRACTOR DATA ANALYTICS OF *E.COLI* STATE TRANSITION

Understanding high-throughput gene expression response of *E.Coli* under aerobosis

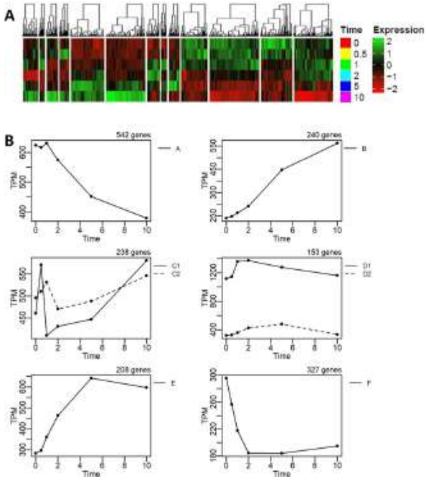
ANAERBOIC TO AEROBIC TRANSITION

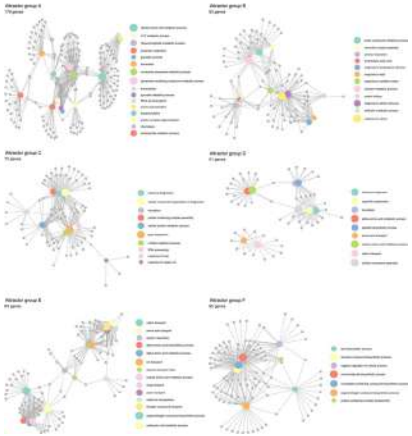






Bui & Selvarajoo (2020) *Sci Rep*



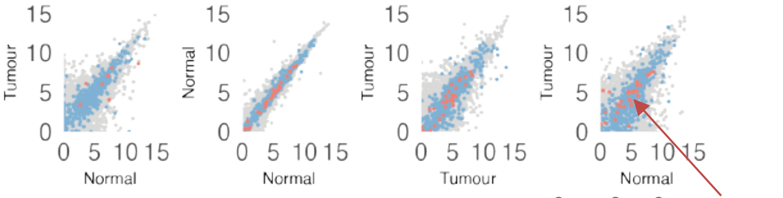


CREATING GROWTH, ENHANCING LIVES

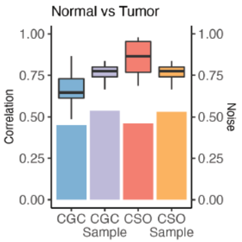
CANCER TRANSCRIPTOMIC ANALYSIS:

BULK (INVARIANT) VS SINGLE CELLS (VARIANT) ONCOGENES

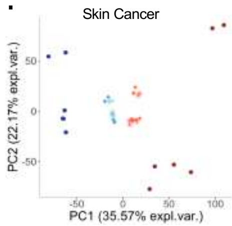
1 Using Cancer Gene Census (CGC) database to identify oncogenes for analysis



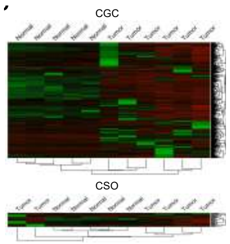
2 Various statistical/ML tools applied on bulk RNA-seq data reveal invariance of oncogene expression



Noise and correlation analysis

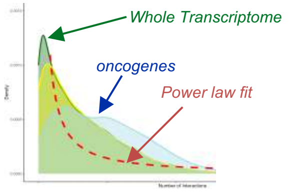


Dimensionality reduction analysis

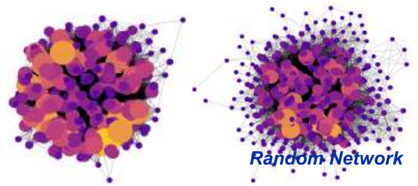


Sample Similarity Analysis

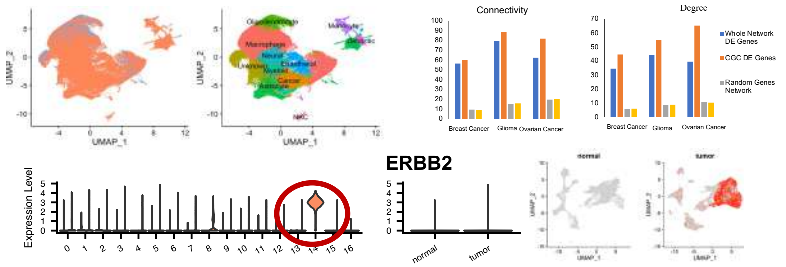
3 PPI Connectivity of oncogenes does not follow Power Law



4 Oncogenes have higher connectivity



5 scRNA-seq data reveals variance of oncogene expression at tissue-specific level



Sirbu… Selvarajoo (under revision)

CREATING GROWTH, ENHANCING LIVES

MULTI-OMICS ML INTEGRATION FOR PREDICTING INDIVIDUAL DRUG RESPONSE

Data acquisition

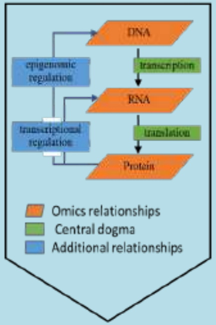
- WES
  - Somatic mutations
  - TP53 …
- RNA-seq
  - DEG
  - ESR1 …
  - Molecular markers
  - Swanton. Paclitaxel Score …
  - Microenv
  - Mast cell …

Tumour
Primary tumour bed area
Fraction of invasive cancer

lymph node
Diameter of largest metastasis
Number of positive lymph nodes

pCR | RCB-I | RCB-II | RCB-III
Not tumour | Increasing residual disease

Feature Selection & Multi-Omics integration



BMSCCA: Biologically structured Multi-class SCCA) Learning with Feature Selection

Objective function:

$$\min_{U,V} \sum_{k=1}^{K}\sum_{m=1}^{M} -u_k^T X_k^T Y_{km} v_{km} + R(U) + R(V_m),$$

$$\text{s.t. } \|X_k u_k\|_2^2 = \|Y_k v_k\|_2^2 = 1, \ \forall k.$$

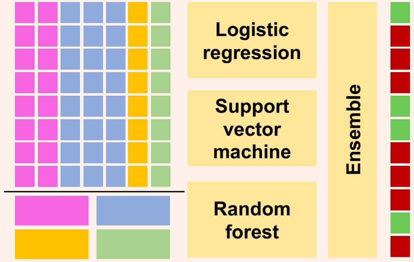
regularizations:

$$R(U) = \alpha_1\|u_k\|_{1,1} + \alpha_2\|U\|_{2,1}$$

$$R(V_m) = \beta_1\|v_{km}\|_{1,1} + \beta_2\|V\|_{2,1}$$

$k(1,\ldots,K)$: disease subtypes
$m(1,\ldots,M)$: different omics modality

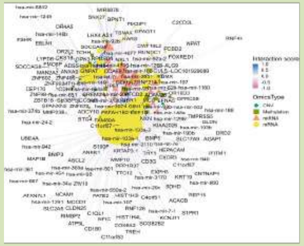
Machine learning



- Logistic regression
- Support vector machine
- Random forest

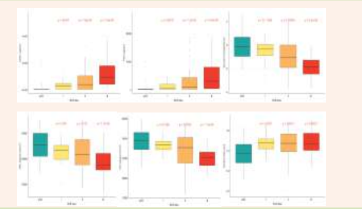
Ensemble

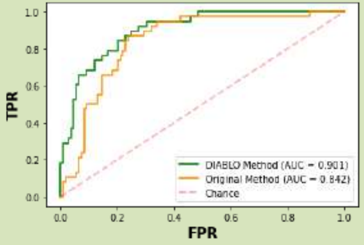
Disease and Patient Specific Results



Identification novel interactions

ZNF496/ ADSS
?
hsa-mir-92b





Rashid, Yuzhou, Selvarajoo (in preparation)

CREATING GROWTH, ENHANCING LIVES

SUMMARY

- Various *in silico* tools/methods (based on physical laws, rules and machine learning) have shown predictive results that are experimentally tested, verified or validated

- These, however, are at cellular levels, and mostly at single layer levels (e.g. transcriptomics, proteomics, metabolomics)

- Digital Twins approach requires a more holistic understanding of the interaction at many levels and layers (e.g. organ level, tumor microenvironment, multi-omics dynamic interactions)

- Big challenges ahead, thus, closer interaction between physicists, mathematicians, biologists, chemists and medical doctors are imminent with clearly defined objectives

CREATING GROWTH, ENHANCING LIVES

Current Members

Dr Hock Chuan Yeo (BII)
Dr Mamun Rashid (BII)
Olga Sirbu (BII)
Yan Ting Hee (BII)
Dr Jasmeet Kaur (SIFBI)
Shi Mun Lee (NUS/ASRL/BII)
Clarence Sim (NTU/BII)
Kamil Konrad Pabis (NUS/BII)
Gunjan Agarwal (SUTD/BII)
Archita Dev (Velore/BII)
Shawn Tay (NTU/BII)
Lucas Kumar (NUS/BII)
Nicole Chong (NUS/BII)
Yuzhou Chen (NTU/BII)

Bioinformatics Institute
BII

THANK YOU

www.a-star.edu.sg

Current Collaborators

Thomas Dawson (ARSL)
Jan Gruber (NUS)
Brian Kennedy (NUS)
Karen Crasta (NUS)
Yvonne Chow (SIFBI)
Wee Chew (SIFBI)
Derek Smith (SIFBI)
Yan Liu (ISCE)
Staffan Kjelleberg (NTU)
Vidu Regina R (NTU)
Alessandro Giuliani (ISS)
Mariano Bizzari (Rome)
Philip Day (Manchester)
Vincent Piras (Toulouse)
Floriant Bellvert (Toulouse)
SYNCTI NUS members
CNRS (Toulouse) members

Acknowledgement: EcoCTs CNRS-CREATE GRANT

https://www.cbio-kumar.org