

How A Biochemist Grew in Big Data Health Science Training

5th National Big Data Health Science Conference
February 3, 2024 | Columbia, SC

Jie Li, Ph.D.

Assistant Professor, Department of Chemistry and Biochemistry, USC



Big Data Health Science Center

UNIVERSITY OF SOUTH CAROLINA

Outline

- **My Unique Experience as A “R25 Fellow”**
- **Significant Outcomes**
- **Future Plans**



My Background

- **B.S., M.S., and Ph.D. in Pharmaceutical Sciences**
- **A biochemist aiming to convert microbial (especially human microbiota) genes into the compounds they encode for improving human health**



My “R25 Fellow” Application Process

- Five different information sources
- Accommodations received from USC
BDHSC
 - Deadline, mentor



My “R25 Fellow” Learning Process

- **Monthly “R25 Fellow” Meeting**
- **USC BDHSC Seminars/Workshops**
- **Courses taken online (Coursera etc.)**
 - Machine Learning, Statistical Analysis of Microbiome Data
- **Literature exchanges with my mentors**



My “R25 Fellow” Mentoring Process

- **My Mentors**

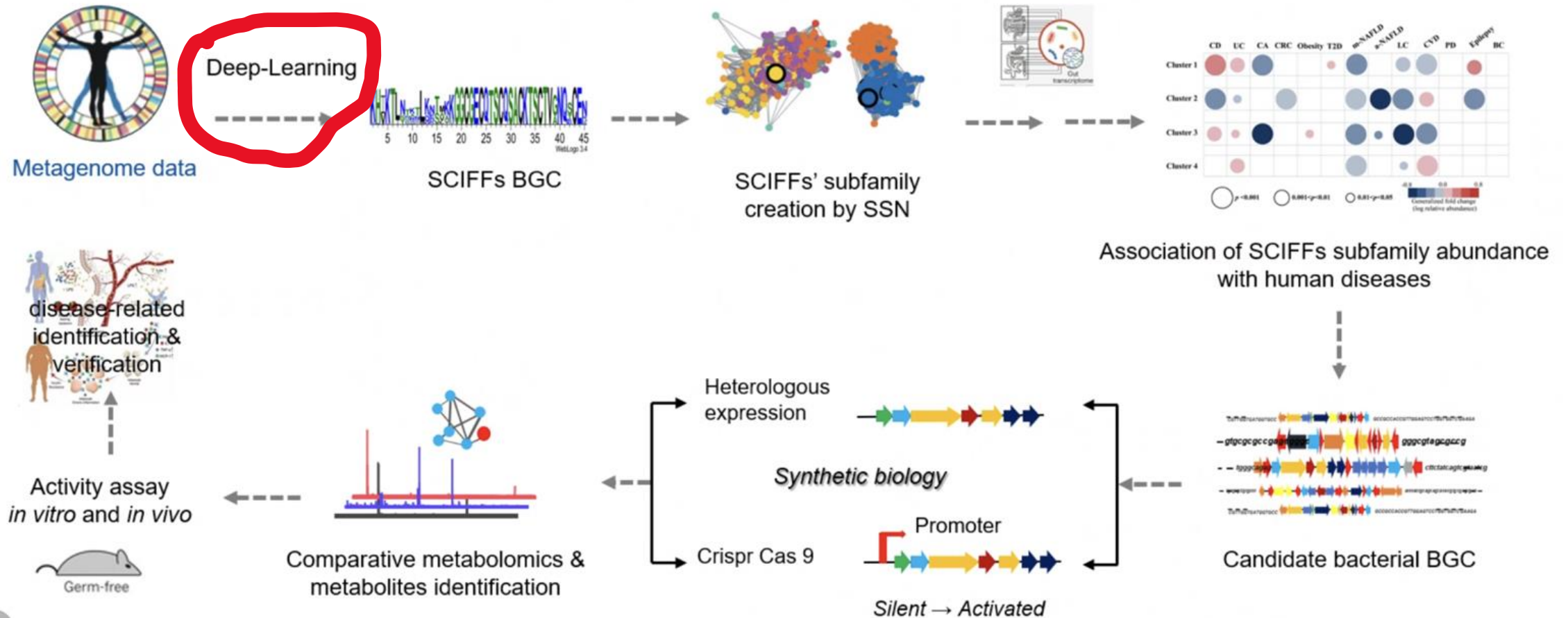
- Dr. Christopher Sutton (Department of Chemistry and Biochemistry, USC; machine learning)
- Dr. Qi Wang (Department of Mathematics, USC; mathematical modeling, machine learning)

- **One-on-one meetings with mentors**



My "R25 Fellow" Mentoring Process

- Attending mentor's lab group meeting



My “R25 Fellow” Mentoring Process

- Planned a workshop for USC BDHSC integrating mutual expertise

08:30-10:00	Lecture
10:00-10:15	Break
10:15-11:45	Lecture
11:45-13:15	Lunch
13:15-14:15	Group Discussion
14:15-14:30	Break
14:30-16:00	Lab/Hands-on activities
16:00 - 16:15	Break
16:15-17:30	Lab/Hands-on activities

The mini syllabus may include:

- Learning Objective
- Topics covered for the [session](#)
- Textbook /reference (reference: [Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases](#))
- Discussion Activities
- Lab/Hands-on activities
- Homework/ Reading materials (reference: [Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases](#))
- Evaluation criteria for the learning

Morning lecture 8:30 – 10:00 am

Overview of the lecture / Topics that will be discussed:

1. Introduction of the Central dogma of biology: DNA to RNA to proteins.
2. Introduction of key metabolites like ATP in the cell are entirely description of the cell [function](#)
3. Introduction to Metabolomics
 - a. What is metabolomics?
 - b. Importance of metabolite identification
 - c. Overview of the metabolome
 - d. Metabolite identification
4. Connection of the structure-properties in proteins, from structural data information on the 3D structures of proteins, often obtained through techniques like X-ray crystallography)
 - a. Protein structures are determined using experimental methods and using ML methods such as [AlphaFold](#)
5. Overview of Ligand-protein interactions

Late-morning lecture 10:15-11:45

1. Analytical Techniques for metabolite quantification and challenges
2. Data Acquisition and Processing
3. Metabolite Identification Strategies
 - a. Use of databases

My “R25 Fellow” Significant Outcomes

- One manuscript in preparation
- NIH R35 MIRA (awarded)
 - “Accessing and expanding natural products chemical diversity by **big-data analysis** and biosynthetic investigation”
- One NIH R41 proposal submitted (pending)



Future Plans

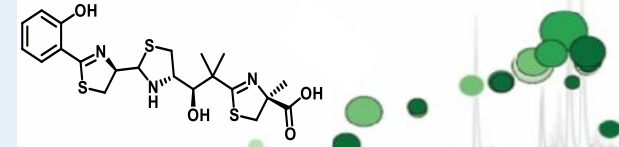
- **Innovative research directions**



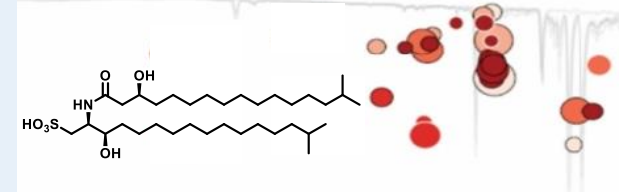
Discovery of functional metabolites produced by human microbiota

Differential Metabolites

Health



Diseases

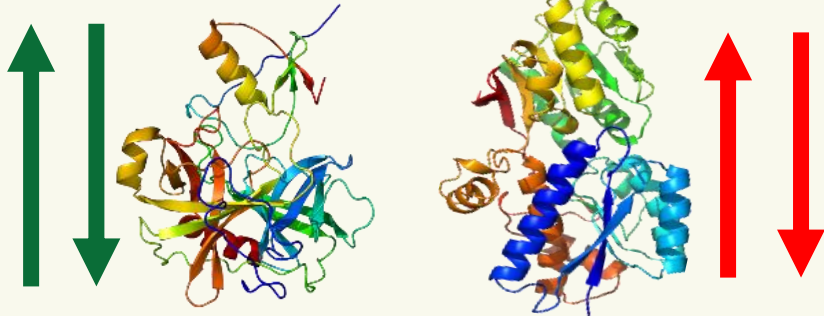


Discovery of functional metabolites produced by human microbiota

Differential Biosynthetic Enzymes

Health

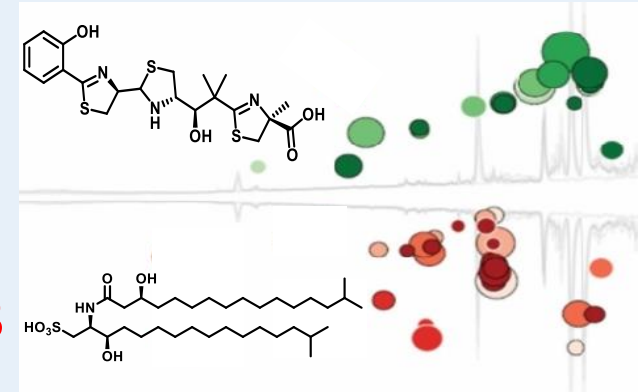
Diseases



Differential Metabolites

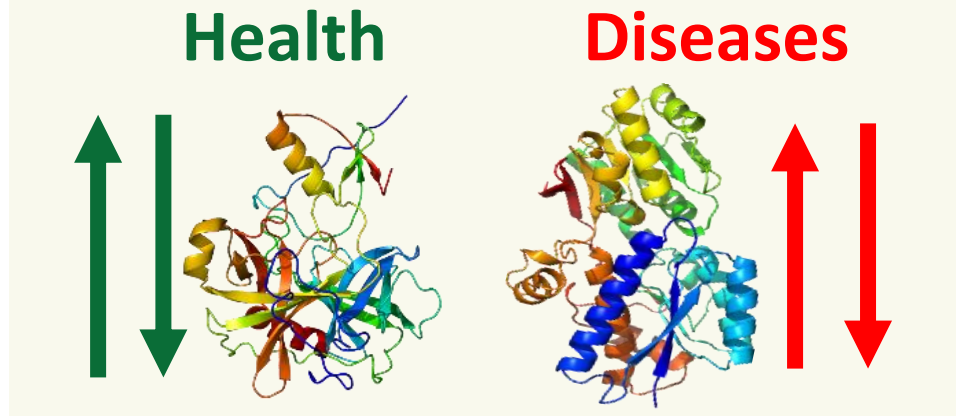
Health

Diseases



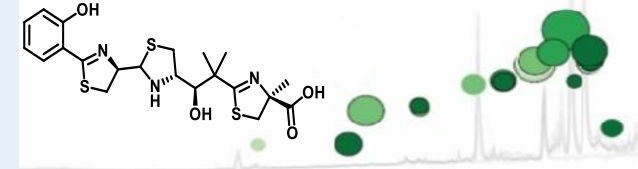
Discovery of functional metabolites produced by human microbiota

Differential Biosynthetic Enzymes



Differential Metabolites

Health



Diseases



High-quality
sequencing data



IBD Multi Omics Database (IBDMDB)

.....

Future Plans

- **Grant applications** (big data analysis + experimental verification: connecting human microbial metabolites to health/diseases; funding opportunities?)



Future Plans

- **Publications**
- **On both career and personal levels:
continue improvement in big data health
sciences**



THANKS!

QUESTIONS?



Big Data Health Science Center

UNIVERSITY OF SOUTH CAROLINA