**Preface**

*Me, a Biomedical Engineer*

I don’t label myself an engineer in the traditional sense; I don’t design things to be manufactured. I also dislike being called a scientist because of the visual it may provoke; I don’t spend most of my time in a long white coat at a bench. Though others might think me one, I don’t consider myself a statistician, computer scientist, or biologist either.

I am a scientist in that I’m curious about nature, especially health and medicine. I’m an engineer in that I’m curious about translating scientific understanding to ideas and tools that will improve others’ quality of life. To do this, I use and invent tools in mathematics, statistics, computer science, and the biological sciences. There is no formal and universally accepted definition of a “biomedical engineer” and likely never will be. However, the above description is the one I use and, I believe, close to the definition of many other self-proclaimed biomedical engineers.

*Me, a Computational Systems Biologist*

As a biomedical engineer in the information age, trillions of data points are available. In this dissertation alone, I gather data from dozens of expression states of the 3.3-billion letter human genome or 2.8-billion letter mouse genome. Tens of thousands of megapixel images with of individual cells are captured at different wavelengths of light. Tissue and blood samples from hundreds of mice under different stresses are analyzed microscopically and by molecular assays to quantify pathology, the proportion of different cells (e.g., epithelial, white blood cell, etc.), and the amounts of dozens of proteins. Many more data are presented in the following chapters.

However, these data independent of each other are of little use. Consider a mouse that ingests a toxin. We observe inflammation in the mouse’s intestine. What else is happening before this inflammation? We then find that a particular gene is expressed. Is this gene alone responsible for inflammation? Are there other genes whose regulation is linked to this gene of interest? How does expression of this gene translate the amount of its gene product, the protein that physically interacts with the cells’ environment? Is this inflammation only due to local events? What about the brain and the nervous system? What happens after inflammation? Are there any changes throughout the body from the local injury in the intestine (e.g., in the blood)? Where did the toxin go? The questions go on, but it is apparent that there are many levels of data and interactions, a **system**, that determine the apparent clinical manifestation, inflammation. **Systems Biology** aims to consider biological systems as a whole and answer how one or many changes affects the state or output of the entire system.

Systems biology requires computational tools just so data can be managed, so the term **Computational Systems Biology** is somewhat redundant. The term’s definition will change from person to person. As I use it in this dissertation, it is distinct in that computational tools are used to make insights that would not otherwise be available from simple comparative experiments (e.g., think null versus alternate hypotheses and p-values). However, computational systems biology does not exclude such experiments. It can go one step further by presenting the data with novel visualizations or novel summary statistics. Consider a metaphor to different classes of proteins or genes. We hypothesize that introverts have fewer friends than extraverts and through a survey find that this is indeed true. However, a network diagram of connections in a social network with more metadata may reveal richer, more specific interpretations (e.g., clusters of introverted individuals or people that are the hubs, the social organizers).

For instance, consider how a network diagram of hundreds of connections from a social media service

But I am computational

*My dissertation as Biomedical Engineer and Computational Systems Biologist*

The unifying element of is a common goal, a better un

The greatest challenge is to make practical use of this data

This data must be understandable and reproducible.

*Acknowledgements*

Colleagues

Lab

MARCE team

Committee

Jason

Friends

Family

**Abstract**

**Introduction**

Biological motivation

C. difficile background

Computational motivation

Great comedians tell us of everyday experiences in new, unexpected ways. Novelists and poets express our thoughts better than we could. Statisticians place abstract common sense into concrete writing. My job is to take things that were always there, but we didn’t know. To present and communicate in a way that allows us to do things in a novel way that lead to discoveries and ways of thinking.  Use many tools to reveal findings from data. Typically, models greatest usefulness are there ability to guide us, not to tell us the answers.

There is always a constant tug from “hypothesis generating” research and “hypothesis generated” research. Discovery based versus logical yes/no experiments. It’s a spectrum, but it’s how we think about it.

**Introduction**