Penn Symposium on Mathematical & Computational Biology

May 22-23, 2017

Kevin Burrage – Queensland University of Technology and University of Oxford

Unlocking complex data sets by calibrating populations of models to data density: a study in atrial electrophysiology

Abstract: The understanding of complex physical or biological systems nearly always requires a characterisation of the variability that underpins these processes. In addition, the data used to calibrate such models may also often exhibit considerable variability. A recent approach to deal with these issues has been to calibrate populations of models (POMs), that is multiple copies of a single mathematical model but with different parameter values. To date this calibration has been limited to selecting models that produce outputs that fall within the ranges of the dataset, ignoring any trends that might be present in the data. We present here a novel and general methodology for calibrating POMs to the distributions of a set of measured values in a dataset. We demonstrate the benefits of our technique using a dataset from a cardiac atrial electrophysiology study based on the differences in atrial action potential readings between patients exhibiting sinus rhythm (SR) or chronic atrial fibrillation (cAF) and the Courtemanche-Ramirez-Nattel model for human atrial action potentials. Our approach accurately captures the variability inherent in the experimental population, and allows us to identify the differences underlying stratified data as well as the effects of drug block. (Joint work with B. Lawson, C. Drovandi, N. Cusimano, P. Burrage, B. Rodriguez)

Arup K. Chakraborty – Massachusetts Institute of Technology

How to hit HIV where it hurts

Abstract: No medical procedure has saved more lives than vaccination. But, today, some pathogens have evolved which have defied successful vaccination using the empirical paradigms pioneered by Pasteur and Jenner. I will describe how bringing together approaches from physics, engineering, and basic/clinical immunology is beginning to confront this challenge. One characteristic of many pathogens for which successful vaccines do not exist is that they present themselves in various guises. HIV is an extreme example because of its high mutability, and it continues to wreak havoc around the world, especially in developing countries. This highly mutable virus can evade natural or vaccine induced immune responses, often by mutating at multiple sites linked by compensatory interactions. If one wishes to define the mutational vulnerabilities of HIV, these collective compensatory pathways need to be identified so as to not target the involved sites by a vaccine induced immune response. Moreover, the combinations of mutations that the virus cannot make and still maintain viability need to be determined, so as to target the pertinent sites by vaccination. Thus, knowledge of the fitness landscape of HIV – fitness as a function of sequence with explicit account of coupling between mutations – is an important ingredient for rational design of vaccines that can confront this scourge. I will describe how we developed models to translate data on HIV protein sequences to knowledge of the HIV fitness landscape, and tested the results against in vitro and clinical data. Based on these results, a therapeutic T cell-based vaccine was designed, which is now being advanced to pre-clinical studies in monkeys. I will also describe work pertinent to how broadly neutralizing antibodies, which can neutralize diverse HIV strains, might be induced by vaccination.

Sarah Cobey – University of Chicago

Three time scales in the coevolution of influenza and human immunity

Abstract: The evolution of neutralizing antibodies from B cells is central to resisting infection by influenza and other pathogens, and the success of vaccines and other interventions hinge on their abilities to steer this evolution. What are the forces shaping B cell evolution? These forces vary across time scales. First, I review selective pressures on antibody repertoires during affinity maturation and then ask–as we might theoretically expect–whether similar patterns of selection are observable over hundreds of millions of years of immunoglobulin gene evolution. Across heavy chain immunoglobulins, there has been strong purifying selection and a decrease in mutability over time that appears to correlate with increased stability. I next investigate changes in the adaptability of long-lived B cell clones to identify how intrinsically high mutation rates, selection on protein structure, and incidental selection for adaptability drive changes in hotspot number over time. In these clones, there is no evidence of long-term purifying selection on hotspots, and hotspot loss does not inevitably result from presumed selection for protein structure: instead, hotspot loss can be predicted well by nearly neutral models of motif evolution. Finally, I review mounting evidence that competition between naive B cell populations and adaptive memory responses drives the phenomenon of "original antigenic sin" to influenza viruses and might explain not only patterns of antigenic evolution but also some unexpected effects of influenza vaccines.

Omer Dushek – University of Oxford

Control of T cell responses by accessory receptors

Abstract: T cells orchestrate immune responses crucial for the elimination of infections and cancers. They do this by initiating a diverse set of effector responses when their T cell surface receptors (TCRs) recognise these threats. It is now appreciated that a large number of other, "accessory", receptors shape these responses. Despite extensive research into the underlying biochemistry, we have yet to formulate canonical models of signalling that can predict how accessory receptors shape T cell responses. We recently showed that a minimal model of T cell signalling is sufficient to explain the T cell response to a 1-million fold variation in antigen dose and affinity. In this talk, we show our preliminary efforts to use this minimal model to identify the integration point(s) of several accessory receptors.

Albert Goldbeter – Université Libre de Bruxelles (ULB)

System Biology of Cellular Rhythms: Modeling the Dynamics of the Mammalian Cell Cycle

Abstract: Oscillations represent a phenomenon of temporal self-organization that arises from feedback processes in cellular regulatory networks. After providing an overview of oscillations in cellular systems and of their underlying mechanisms, I will focus on the cell cycle, which provides a major example of cellular rhythm. The mammalian cell cycle is driven by a network of cyclin-dependent kinases (CDKs). Computational models show how the regulatory design of the CDK network results in its temporal self-organization in the form of sustained oscillations that bring about the orderly progression along cell cycle phases. Multiple factors, intrinsic or extrinsic to the CDK network, control the balance between cell cycle arrest and cell proliferation, which is often deregulated in cancer. Among these factors are oncogenes, tumor suppressors, growth factors, as well as the extracellular matrix, and contact inhibition, which increases with cell density. Supra-threshold changes in the level of any of these factors can trigger a switch in the dynamical behavior of the CDK network between a stable steady state, associated with cell cycle arrest, and sustained CDK oscillations,

associated with cell proliferation. That different oscillatory processes may interact at the cellular level is illustrated by the coupling of the cell cycle to the circadian clock, which results in the synchronization of these two major cellular rhythms.

Shishi Luo – University of California, Berkeley
The evolution of the IGHV and TRBV gene families

Abstract: Despite the crucial role played by the immunoglobulin heavy variable (IGHV) and T cell beta variable (TRBV) loci in adaptive immune function, their genetic variation in the human population remains poorly characterized. Generated through a process of gene duplication/deletion and diversification, these loci can vary between individuals in gene copy number and contain genes that are highly similar. Such characteristics make the analysis of these loci technically challenging. Here, using a customized bioinformatic analysis, we quantify the copy number and single nucleotide variation in these two regions in a globally diverse sample of hundreds of individuals sequenced by the Simons Genome Diversity Project. We find that despite the shared molecular and functional characteristics between the IGHV and TRBV gene families, they exhibit starkly different patterns of variation. In particular, we find that the IGHV locus has a greater propensity for gene duplication and deletion, as evidenced by a high level of copy number variation, while the TRBV seems to have a greater propensity for nucleotide mutations, as evidenced by a moderately high level of allelic diversity. This pattern appears to be consistent across other species and suggests that the IGHV and TRBV loci have followed different evolutionary paths in their 500 million years or so of existence.

Erick Matsen – Fred Hutchinson Cancer Research Center
Learning about antibody affinity maturation from sequence data using probabilistic models

Abstract: Our survival in the face of pathogen exposure depends on a carefully orchestrated evolutionary process acting on our B lymphocytes. This process, called "affinity maturation," is driven by specialized machinery to mutate and enforce selection pressure on the B cell receptor (BCR) sequences that encode antibody specificity. It has recently become possible to sequence BCRs in high throughput. Although these sequences implicitly contain a wealth of information about both antigen exposure and the process by which we learn to resist pathogens, statistical methods and computer algorithms are needed to extract this information.

In this talk, I will describe two recent projects to develop model-based inferential tools for analyzing BCR sequences: first, a survival model framework to learn the nucleotide context sensitivity of the mutation process in affinity maturation, and second, a branching process model that allows us to improve phylogenetic estimation by using sequence abundances.

This work is joint with David Shaw and Will DeWitt (Fred Hutch), and Jean Feng, Vladimir Minin, and Noah Simon (University of Washington)

Thierry Mora – École Normale Supérieure

Optimal immune systems

Abstract: Immune systems protect us against a large diversity of pathogens, while being tolerant to our own molecules. What are the general design principles of immune systems that allow them to perform this complex task efficiently and reliably, with the physical constraints under which they operate? I will explore this question by considering optimal immune designs at three different scales of organisation. At the molecular scale, how do receptors and signaling network distinguish between foreign and self ligands? At the individual scale, how should the repertoire of specificities be organised to best detect infections? And at the population level, what are the optimal modes of immunity (adaptive, CRISPR, etc.) as a function of the pathogen statistics?

Eric Siggia – Rockefeller University

Computation to study insect evolution and temperature compensation in circadian clocks

Abstract: Computational evolution by mutation-selection is used to derive gene regulatory networks that perform a variety of functions. Recent examples formulate an analogy with the transition state in a chemical reaction to derive the anterior-posterior patterning in the last common ancestor between fly and mosquito. The circadian clock is built from temperature dependent components, yet has a temperature independent period. A possible architecture was derived computationally and was recently shown to have experimental support.

Christine Vogel – New York University

The ups and downs of protein expression regulation

Abstract: Gene expression is regulated by four major processes: transcription, translation, and RNA and protein degradation. These processes are adjusted, in different ways, when the cells respond to a stimulus. Many pathways are known, but their precise interaction over time is not well understood. In our lab, we use multiple time series datasets — on protein and mRNA expression changes and changes in the binding of ribosomes and other proteins — in combination with mass action models and other approaches to disentangle the contributions of the different levels of regulation and generate hypotheses on regulatory mechanisms. We focus on yeast and mammalian cells responding to stress of the endoplasmic reticulum, but have expanded these studies in a variety of directions.