

Department of Systems and Computational Biology

Albert Einstein College of Medicine



Albert Einstein College of Medicine

Montefiore

*Faculty Research Interests
2019-2020*

Introduction

The Department of Systems and Computational Biology, through innovative research and education, focuses its efforts on advancing the understanding of living systems as a whole by promoting a new approach to biology. Our research combines theoretical and experimental approaches aimed at explaining how the higher-level properties of complex biological systems materialize from the interactions among their parts.

The faculty develop research and education programs that embrace engineering, computational, mathematical and physical sciences as an integral part of the Biological and Biomedical sciences, leading to the foundation of a Systems and Computational Biology discipline. We seek to form an academic environment that benefits from and respects the value of these existing disciplines. An important part of this fusion entails serious, sustained and multidisciplinary research programs in pursuit of fundamental questions in Biology, ranging from the function of biological systems to an understanding of the evolution of life's diversity.

The main challenges currently facing Biology involve the understanding of biological phenomena at a holistic, systems level without neglecting the reductionist, detailed information approach related to its individual components. This is in contrast to the "massively parallel reductionism" view. That is, despite the availability of a vast amount of information, the focus, nonetheless, remains on the individual components. The academic community has recognized, once again, the need for integrated, systems-level strategies that combine computational and experimental tools, as well as evolutionary-based inferences in addressing these challenges. It is this integration that defines the department's activity.

Faculty, Research Fellows, Research Trainees, Graduate Students, and Bioinformatic Analysts

Faculty

Aviv Bergman, Ph.D., Professor and Founding Chairman
Ruben Coen-Cagli, Ph.D., Assistant Professor
Eduardo Fajardo, Ph.D., Research Assistant Professor
Andras Fiser, Ph.D., Professor
Libusha Kelly, Ph.D., Assistant Professor
Adam Kohn, Ph.D.** , Professor
Parsa Mirajhi, M.D., Ph.D., Research Associate Professor
Ian Willis, Ph.D.* , Professor
Yinghao Wu, Ph.D., Associate Professor
Qian K. Ye, Ph.D.* , Associate Professor

Research Fellows

William Chang, Ph.D. (Kelly Lab)
Dylan Festa, Ph.D. (Coen-Cagli Lab)
Yehonatan Sella, Ph.D. (Bergman Lab)
Rojan Shrestha, Ph.D. (Fiser Lab)
Zhaoqian Su, Ph.D. (Wu Lab)
Sarah Wolfson, Ph.D. (Kelly Lab)
Jonathan Vacher, Ph.D. (Coen-Cagli Lab)

Research Trainees

Sacha Sokoloski (Coen-Cagli Lab)
Daniel Herrera (Pre-doctoral Research Trainee in the Coen-Cagli Lab)

Graduate Students

Reese Hitchings (Advisor: Kelly)
Ruth Hauptman, MSTP Student (Advisor: Kelly)
Saad Khan, MSTP Student (Advisor: Kelly)
Maryl Lambros (Advisor: Bergman)
Oren Weiss, MSTP Student (Advisor: Coen-Cagli)

Bioinformatic Analysts

Carlos Madrid-Aliste (Fiser Lab)
Al Tucker (Fiser Lab)

**Secondary appointment in the Department of Systems and Computational Biology*

***Tertiary appointment in the Department of Systems and Computational Biology*

Requirements for a Ph.D. in Systems and Computational Biology

A. Courses

Students who wish to pursue a Ph.D. in Systems and Computational Biology must successfully complete seven courses. Introduction to Systems Biology: Theory and Case Studies and the Systems Biology Seminar must both be successfully completed. The other five courses will be determined with the help of the student's mentor and advisory committee, although Graduate Biochemistry and either Molecular Genetics or Gene Expression are strongly recommended courses for all students.

B. Advisors

By the end of the first year, all graduate students choose a faculty member to serve as their primary research advisor, as well as an advisory committee made up of three to five faculty members (all will not necessarily be from the Department of Systems and Computational Biology).

C. Qualifying Examination

Each candidate for the Ph.D. degree must satisfactorily complete a qualifying examination, typically taken after the second year of study. By this time, the student should have fulfilled the bulk of his or her coursework. The exam includes both an oral presentation and written proposal on the student's proposed thesis project, and will be assessed by a Qualifying Exam Committee.

D. Thesis Dissertation

1. First Author Manuscript: The Graduate School requires that at least one first-author manuscript must be published before a student can defend his or her thesis. If a manuscript is not published, then the final draft of a Submitted, In Revision, or In Press first-author manuscript must be appended to the thesis.
2. Thesis Defense Committee: According to designated criteria established by the Graduate Committee, all Thesis Defense Committees must be approved by the Assistant Dean. The Committee must include a minimum of five faculty, one of which will be designated as the Committee Chair and two of which must be from the Department of Systems and Computational Biology. The thesis must be presented to this committee at least 3 weeks prior to the thesis defense.
3. Public Seminar: The presentation of a public seminar at the College of Medicine is required for successful completion of the Ph.D. degree.
4. Thesis Defense: All candidates for the Ph.D. degree must submit a satisfactory dissertation and pass an oral examination (defense) of their thesis.

E. Other Requirements

The department Journal Club/Works-in-Progress meets once a month. Seminars are held monthly during the academic year. Additional seminars in specific areas are also held throughout the semester. Attendance at these activities is required.

Please see <www.einstein.yu.edu/phd> for a more detailed guide to requirements and policies.

Aviv Bergman, Ph.D.
Professor and Founding Chairman, Department of Systems and Computational Biology
Professor, Department of Pathology
Professor, Dominick P. Purpura Department of Neuroscience
Price Center, Room 153C
aviv@einstein.yu.edu; (718) 678-1063

Selective Research Topics from the Bergman Lab:

Topology of biological networks

We study the relationship between the topology of biological networks and their functional (e.g. robustness) and evolutionary (e.g. polymorphism and divergence) properties. It has been conjectured that genes with a large number of downstream targets are more highly conserved, and when compromised, will tend to have a larger effect on network functioning than sparsely connected genes. However, we have shown that 'topdown' inferences of biological properties based on simple measures such as number of targets, are of limited utility. We argue that such lack of predictive power is the result of a composite effect in which certain sub-networks obeying a strong correlation between biological function and simple measures, coexist with other sub-networks having no correlation at all. We have demonstrated that more detailed information, e.g., dynamic gene-expression data, and the specifics of the genetic background, are needed to make meaningful functional and evolutionary inferences.

Investigations with an evolutionary perspective, such as these, can also be extended to biomedical research of phenotypic traits resulting from complex genetic interactions, including Cancer, Diabetes, Hypertension and Aging, as well as mechanistic models of the immune system. Indeed, we have successfully applied methodologies adopted from evolutionary theory to identify genes associated with extreme longevity as well as their targets, age-related disease genes.

Computational Immunology and somatic hypermutation

Somatic hypermutation (SHM) is a key process in the generation of antibody diversity that normally operates in antibody-forming B cells by introducing point mutations into the variable regions of immunoglobulin (Ig) heavy and light chain genes. SHM is initiated when the highly mutagenic enzyme activation-induced deaminase (AID) generates C→U mutations by deaminating cytosines preferentially at WRC hotspots motifs (where W=A/T, R=G/A and C is the mutated base). In collaboration with Matthew Scharff (Department of Cell Biology, Albert Einstein College of Medicine), we use computational and statistical methods together with relevant experimental data to improve our understanding of the molecular mechanisms underlying SHM.

How does the target sequence affect AID activity? To study the behavior of AID and the role of the target sequence, we have used computational methods to compare mutated sequences from three different models of AID activity: (a) an *in vivo* mouse model, (b) an *in vitro* model which captures essential biochemical activity of AID on DNA, and (c) an *in silico* model which simulates only hotspot targeting. This analysis suggests that there is considerably more complexity involved in the mutation process than can be described by simple WRC hotspot motifs. We have also found strong differences between the two strands (transcribed and non-transcribed) in terms of the similarity between the models. A potential clue comes from differences in the profile of inter-mutational distances between the two strands, which suggest the existence of a complex interplay between the enzyme structure and the sequence.

Evolution of gene regulatory networks

There is little doubt that plasticity in gene regulatory networks plays a key role in evolution, particularly in developmental networks. We use computational and mathematical models of gene networks to investigate key evolutionary questions and generate novel hypotheses. Where possible we also use relevant biological data to confirm theoretical findings.

How does degeneracy in transcription factor binding motifs affect evolution of *cis*-regulatory regions? In collaboration with Andras Fiser (SCB, Albert Einstein College of Medicine) we are developing structural models of transcription factor – DNA interactions in which we predict binding affinities for all possible interactions. The predicted binding affinities have been integrated with existing evolutionary models, enabling us to address questions concerning the evolution of regulatory motifs. Turnover of transcription factor binding sites is widespread in both insects and mammals, yet is poorly understood. Using our modeling framework we aim to understand what factors (e.g. motif degeneracy or selection) influence turnover rates.

What is fate of duplicated genes in networks? Several explanations have been proposed to explain the unexpectedly high retention of duplicate genes. One popular theory is the duplication-degeneration-complementation (DDC) model, which proposes that following gene duplication the two gene copies degenerate to perform complementary functions that jointly match that of the single ancestral gene, a process also known as subfunctionalization. However, the DDC model is gene-centric, and does not take into account the network context. Using computational models of evolving gene networks we have analyzed the fate of duplicate genes and found that network plasticity undermines the relevance of subfunctionalization, and that neofunctionalization (recruitment of novel interactions) plays a more predominant role than was previously thought.

How did sexual reproduction evolve? The prevalence of sexual reproduction, as opposed to asexual reproduction, remains one of the most perplexing phenomena in evolutionary biology. We have used computational modeling to explore the role played by epistasis, a condition in which mutations cause a greater change in fitness when combined than would be expected from their individual effects.

Caloric Restriction

Caloric restriction (CR) is a major intervention conclusively shown to extend lifespan in many organisms including mammals, birds, nematodes, flies and even unicellular species. Besides extending lifespan, CR also has been shown to prevent age-associated diseases and keeps organisms in a relatively youthful and healthy state compared to the ad libitum fed counterparts. These observations suggest that the somatic maintenance functions (e.g., cellular error-checking and damage repair) may be up-regulated in animals under CR conditions.

We try to derive a general quantitative and predictive theory, from physical energetic viewpoints, for understanding CR's effects on retarding aging and maintenance. We hypothesize that the longevity of an organism is correlated to biological pathways of maintenance of its integrities (e.g., repairing damage and error-checking), which are energetically costly, and that CR, counter-intuitively, by suppressing organisms' caloric energy supply and biosynthesis, alters the organisms' energy allocation strategy and channels additional energy/resource to the maintenance pathways, therefore retarding aging and extending the lifespan.

Based on principles of mass and energy balance and allometric scaling of metabolism and biosynthesis, we have developed an empirically grounded theoretical model that correctly predicts how organisms allocate energy between the synthesis of new biomass and the maintenance of existing biomass with normal food supply. We then try to extend the model and apply it to animals under CR. During growth, organisms need to do override work (indirect metabolic work) to store energy in new biomass. Our preliminary study suggests that because CR suppresses the energy storage in biomass, organisms do not devote as much metabolic work to do this storage as with normal food supply, therefore this amount of metabolic work can possibly be channeled to maintenance. We take the elongated lifespan as a measurement of enhanced maintenance during CR, and focus on four longstanding questions regarding CR's effects on maintaining organisms' integrity. (1) How does body temperature drop in CR animals influence CR's effect? (2) What is the relationship between intensity of CR and its effect? (3) With the same intensity and period of CR, how is an organism's adult body size correlated to CR's effect? (4) How does the age at which CR begins influence the CR's effect?

Besides the theoretical development, we also test hypotheses by a meta-analysis of empirical data on a diverse set of species from published literatures.

References:

Kunisaki, Y., Bruns, I., Scheiermann, C., Ahmed, J., Pinho¹, S., Zhang, D., Mizoguchi, T., Wei¹, Q., Lucas, D., Ito, K., Mar, J, C., Bergman, A., Frenette, P. S. Arteriolar niches maintain haematopoietic stem cell quiescence. *Nature*. 2013 Oct 31;502(7473):637–43. doi: 10.1038/nature12612

Pujato, M., MacCarthy, T., Fiser, A., and Bergman, A. The underlying molecular and network level mechanisms in the evolution of robustness in gene regulatory networks. *PLOS Computational Biology*. 2013 Jan. 9(1):e1002865

Conrad, T.M., Lewis, N.E. & Palsson, B.Ø. Microbial laboratory evolution in the era of genome-scale science. *Molecular Systems Biology* 7, (2011).

MacCarthy T, Kalis SL, Roa S, Pham P, Goodman MF, Scharff MD, Bergman A. V-region mutation in vitro, in vivo, and in silico reveal the importance of the enzymatic properties of AID and the sequence environment. *Proc Natl Acad Sci USA* 106(21): 8629-8634, 2009.

MacCarthy T, Roa S, Scharff, MD, Bergman A. SHMTool: a webserver for comparative analysis of somatic hypermutation datasets. *DNA Repair* 8(1): 137-141, 2009.

Buckling A, Craig Maclean R, Brockhurst MA, Colegrave N. The Beagle in a bottle. *Nature* 457, 824-9(2009).

West GB, Bergman A. Toward a systems biology framework for understanding aging and health span. *J Gerontol A Biol Sci* 64(2):205-208, 2009.

Roa S, Avdievich E, Peled JU, MacCarthy T, Werling U, Kuang FL, Kan R, Zhao C, Bergman A, Cohen PE, Edelman W, Scharff MD. Ubiquitylated PCNA plays a role in somatic hypermutation and class-switch recombination and is required for meiotic progression. *Proc Natl Acad Sci USA* 105(42): 16248-53, 2008.

MacCarthy T, Bergman A. The limits of subfunctionalization. *BMC Evol Biol* 7: 213, 2007.

Bergman A, Atzmon G, Ye K, MacCarthy T, Barzilai N. Buffering mechanisms in aging: a systems approach toward uncovering the genetic component of aging. *PLoS Comput Biol* 3(8): e170, 2007.

MacCarthy T, Bergman A. Coevolution of robustness, epistasis, and recombination favors asexual reproduction. *Proc Natl Acad Sci USA* 104(31), 12801-12806, 2007.

Siegal ML, Promislow DE, Bergman A. Functional and evolutionary inference in gene networks: does topology matter? *Genetica* 129(1): 83-103, 2006.

MacCarthy T, Pomiankowski A, Seymour R. Using large-scale perturbations in gene network reconstruction, *BMC Bioinformatics*, 6:11, 2005.

Masal J, Bergman A. The evolution of the evolvability properties of the yeast prion [PSI+]. *Evolution* 57(7): 1498-1512, 2003.

Bergman A, Siegal ML. Evolutionary capacitance as a general feature of complex gene networks. *Nature* 424(6948): 549-552, 2003.

MacCarthy T, Seymour R, Pomiankowski A. The evolutionary potential of the *Drosophila* sex determination gene network, *J Theor Biol* 225: 461-468, 2003.

Elena, S.F. & Lenski, R.E. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics* 4, 457-69 (2003).

Siegal ML, Bergman A. Waddington's canalization revisited: developmental stability and evolution. *Proc Natl Acad Sci USA* 99(16): 10528-10532, 2002.

Mac Lane, S. *Categories for the Working Mathematician (Graduate Texts in Mathematics)*. 326(Springer: 1998).4.

Gould, S. *Full House: the spread of excellence from Plato to Darwin*. (Three Rivers Press: 1997).

Carroll, S.B. Chance and necessity: the evolution of morphological complexity and diversity. *Nature* 409, 1102-9(2001).3. Buss, L.W. *The evolution of individuality*. (Princeton University Press: Princeton, 1987).

Buss, L.W. *The evolution of individuality*. (Princeton University Press: Princeton, 1987).

Ruben Coen-Cagli, Ph.D.
Assistant Professor, Department of Systems and Computational Biology
Assistant Professor, Dominick P. Purpura Department of Neuroscience
Price Center, Room 353B
Ruben.coen-cagli@einstein.yu.edu; (718) 678-1150

WHAT? My lab studies neural computation with the broader goal of explaining our perceptual experience. A central function of the visual system is to produce correct interpretations of sensory signals, to guide appropriate behavioral responses. However, the surrounding environment is in general ambiguous (e.g. different objects can produce similar retinal images) and computationally intractable (e.g. the same object can produce countless different retinal images). To solve these problems, the brain must evaluate how probable different interpretations of the sensory input are. Understanding such probabilistic inference in natural sensory processing will be central to understanding perception, and much of the computation realized by cortical neurons.

HOW? Our lab follows a hypothesis-driven approach to understanding cortical processing of natural images and linking it to visual perception. Computer vision and machine learning provide insights into the complex structure of natural signals and how they could be processed efficiently. Probabilistic neural coding provides the theoretical framework to understand how veridical perception is achieved in face of abundant sensory noise and image ambiguities. We combine advances in both fields to generate novel hypotheses about cortical computation in natural vision, and test them experimentally with psychophysics in the lab and electrophysiology through collaborations.

WHY? Explaining how the human visual system achieves its impressive feats – from fast and accurate recognition of people and their actions, to the appreciation of Picasso’s Guernica – is a major goal of neuroscience, and more generally biology and medicine. My lab’s research aims to contribute a substantial step forward to this endeavor, by taking a principled approach to studying the visual system in its natural operation mode. In the longer run, we hope this research will contribute to elucidating how the brain produces the vivid, coherent, stable percepts we experience in everyday life; to advancing technologies that could restore impaired vision and enhance normal vision; and to deciphering the neural basis of human visual creativity.

RECENT AND ONGOING RESEARCH

Natural image processing in visual cortex

Signals from the natural environment are processed by neuronal populations in the cortex. Understanding the relationship between those signals and cortical activity is central to understanding normal cortical function and how it is impaired in psychiatric and neurodevelopmental disorders. Substantial progress has been made in elucidating cortical

processing of simple, parametric stimuli, and computational technology is improving descriptions of neural responses to naturalistic stimuli. However, how cortical populations encode the complex, natural inputs received during every day perceptual experience is largely unknown. This project aims to elucidate how natural visual inputs are represented by neuronal populations in primary visual cortex (V1). Progress to date has been limited primarily by two factors. First, during natural vision, the inputs to V1 neurons are always embedded in a spatial and temporal context, but how V1 integrates this contextual information in natural visual inputs is poorly understood. Second, prior work focused almost exclusively on single-neuron firing rate, but to understand cortical representations one must consider the structure of population activity—the substantial trial-to-trial variability that is shared among neurons and evolves dynamically—as this structure influences population information and perception. The central hypothesis of this project is that cortical response structure is modulated by visual context to approximate an optimal representation of natural visual inputs. To test the hypothesis, this project combines machine learning to quantify the statistical properties of natural visual inputs, with a theory of how cortical populations should encode those images to achieve an optimal representation, to arrive at concrete, falsifiable predictions for V1 response structure. The predictions will be tested with measurements of population activity in V1 of awake monkeys viewing natural images and movies. This project will provide the first test of a unified functional theory of contextual modulation in V1 encoding of natural visual inputs, and shed light on key aspects of natural vision that have been neglected to date.

- [1] R. Coen-Cagli, A. Kohn*, O. Schwartz*, (2015) Flexible Gating of Contextual Modulation During Natural Vision. *Nature Neuroscience* 18: 1648–1655
- [2] R. Coen-Cagli, P. Dayan, O. Schwartz (2012) Cortical Surround Interactions and Perceptual Saliency Via Natural Scene Statistics. *PLoS Computational Biology*, 8(3): e1002405.
- [3] R. Coen-Cagli, O. Schwartz, (2013) The impact on mid-level vision of statistically optimal divisive normalization in V1. *Journal of Vision*, 13(8):13
- [4] O. Schwartz, R. Coen-Cagli, (2013) Visual Attention and Flexible Normalization Pools. *Journal of Vision*, 13(1):25
- [5] M. Snow, R. Coen-Cagli, O. Schwartz (2016) Specificity and timescales of cortical adaptation as inferences about natural movie statistics. *Journal of Vision* October 2016, Vol.16, 1.

Probabilistic models of perceptual grouping and segmentation in natural vision

To understand and navigate the sensory environment, sensory systems must solve simultaneously two competing and challenging tasks: the segmentation of a sensory scene into individual objects and the grouping of elementary sensory features to build these objects. Understanding perceptual grouping and segmentation is therefore a major goal of sensory neuroscience, and it is central to advancing artificial perceptual systems. To make progress in understanding image segmentation and improving algorithms, we are pursuing an approach that combines two key components. First, we are developing a new

experimental paradigm that allows for well-controlled measurements of perceptual segmentation of natural images. This addresses a major limitation of existing data that are either restricted to simple, artificial stimuli, or, for natural images, rely on manual labeling and conflate perceptual, motor, and cognitive factors. Second, we are extending and testing a computational framework that accommodates bottom-up information about image statistics and top-down information about objects and behavioral goals. This is in contrast with the paradigmatic view of visual processing as a feedforward cascade of feature detectors, that has long dominated computer vision algorithms and our understanding of visual processing. Our approach builds instead on the influential theory that perception requires probabilistic inference to extract meaning from ambiguous sensory inputs. Segmentation is a prime example of inference on ambiguous inputs: the pixels of an image often cannot be labeled with certainty as grouped or segmented. We will test the central hypothesis that human visual segmentation is a process of hierarchical probabilistic inference.

[1] J. Vacher, R. Coen-Cagli. Combining mixture models with linear mixing updates: multilayer image segmentation and synthesis. arxiv:1905.10629.

[2] J. Vacher, P. Mamassian, R. Coen-Cagli. Probabilistic model of visual segmentation. arxiv:1806.00111.

The origins of neural and behavioral variability.

Perception and behavior are often variable, even when sensory stimuli are held constant and regardless of how hard one tries. Such variability is a consequence of the variable activity of populations of neurons. Our recent research has focused on the origins of such variability. We developed a modeling framework that takes into account realistic, high-dimensional, noisy sensory inputs, such as the signals transmitted by the retina to the thalamus and visual cortex. Within this framework, we derived new analytical results showing that feedforward filtering of noisy visual stimuli induces realistic cortical variability. This variability is not independent across neurons, consistent with electrophysiological data, and such noise correlations limit coding accuracy even when the filtering is optimal, accounting well for behavioral thresholds in orientation discrimination [1,2]. We are currently extending the framework to auditory inputs. Our initial results show that it explains behavioral variability and peculiar biases in sound-source localization, while capturing neural response properties in the barn owl auditory pathway [3]. Furthermore, we are interested in directly relating cortical variability measured experimentally to perceptual variability. This usually involves decoding neuronal activity to quantify how much information neurons carry about the visual stimulus. We have developed a new analytical method for quantifying coding accuracy from experimental data, that is more reliable, and much faster, than state-of-the-art decoding algorithms [4].

[1] I. Kanitscheider*, R. Coen-Cagli*, A. Pouget (2015) The origin of information-limiting noise correlations. PNAS 112(50): E6973-E6982.

[2] A. Kohn, R. Coen-Cagli, I. Kanitscheider, A. Pouget (2016) Correlations and neuronal population information. Annual Reviews of Neuroscience. 39:237-256.

[3] R. Coen-Cagli, G. Dehaene, A. Pouget, Uncertainty coding in a model of auditory localization. *Cosyne* 2016.

[4] I. Kanitscheider*, R. Coen-Cagli*, A. Kohn, A. Pouget (2015) Measuring Fisher Information Accurately in Correlated Neural Populations. *PLoS Computational Biology*, 11(6): e1004218

[5] R. Coen-Cagli, S. S. Solomon (2019) Neural response variability and divisive normalization. *Journal of Neuroscience* (in press)

Andras Fiser, Ph.D.
Professor, Department of Systems and Computational Biology
Professor, Department of Biochemistry
Price Center, Room 453A
andras.fiser@einstein.yu.edu, (718) 678-1068

Computational protein structure modeling and design

The main interest of our group is to study the evolution of protein structures from a modular perspective and to develop theoretical methods based on these observations to model and design protein structures and functions. We are developing methods to design new molecular shapes, either by redesigning existing proteins into biologically more viable shapes or to explore thermodynamically possible molecular architectures. In terms of modeling we are developing techniques in hybrid modeling, specifically, using NMR chemical shift information for de novo prediction of protein structures.

T-cell costimulatory proteins

We are interested to understand how co-stimulation is enacted on a molecular level in the immunological synapse. We are exploring the molecular classification, recognition, regulation of Immunoglobulin Superfamily proteins playing role in T-cell recognition, and modulating the immune response. We are developing techniques that combine molecular design with docking to establish cognate partners within this essential set of proteins. This work is conducted within the Protein Structure Initiative.

Glioblastoma multiforme stem cells

Despite the advances in surgical, radiation and chemotherapy treatments of human glioblastoma (GBM) the prognosis remains poor with fewer than 3% of glioma patients alive 5-years after diagnosis. Molecular targeted therapies hold the promise of providing new anticancer treatments that are more effective than traditional therapies. In a four-way collaboration with brain surgeons, molecular biologists and a high throughput sequencing facility, we are exploring molecular signatures of GBM stem cells from the perspective of gene expression, RNA editing, RNA methylation.

References:

Vallat B, Madrid-Aliste C, Fiser A
Modularity of Protein Folds as a Tool for Template-Free Modeling of Structures.
PLoS Comput Biol (2015) 11(8): e1004419

Pujato M, Kieken F, Skiles AA, Tapinos N, Fiser A
Prediction of DNA binding motifs from 3D models of transcription factors; identifying TLX3

regulated genes.

Nucleic Acids Res (2014) 42(22): 13500-12

Khafizov K, Madrid-Aliste C, Almo SC, Fiser A. Trends in structural coverage of the protein universe and the impact of the Protein Structure Initiative. *Proc Natl Acad Sci U S A* (2014) **111**(10): 3733-8

Yap EH, Rosche T, Almo S, Fiser A. Functional clustering of immunoglobulin superfamily proteins with protein-protein interaction information calibrated hidden markov model sequence profiles. *J Mol Biol* (2014) **426**(4): 945-61

Menon V, Vallat BK, Dybas JM, Fiser A. Modeling proteins using a super-secondary structure library and NMR chemical shift information. *Structure* (2013) **21**(6): 891-9

Rubinstein R, Ramagopal UA, Nathenson SG, Almo SC, Fiser A. Functional classification of immune regulatory proteins. *Structure* (2013) **21**(5): 766-76

Pujato M, MacCarthy T, Fiser A, Bergman A. The underlying molecular and network level mechanisms in the evolution of robustness in gene regulatory networks. *PLoS Comput Biol* (2013) **9**(1): e1002865

Libusha Kelly, Ph.D.
Assistant Professor, Department of Systems and Computational Biology
Assistant Professor, Department of Microbiology and Immunology
Price Center, Room 553B
Libusha.kelly@einstein.yu.edu; (718) 678-1148

The Kelly lab: understanding microbiome population dynamics to improve human health.

Microbial populations are dynamic, transactional social networks of paramount importance to human health. The collection of microbes living in and on our bodies (the ‘microbiome’) is influenced by top-down and bottom-up regulation and these processes are poorly understood. For example, the availability of resources produced by other microbes (‘public goods’) can inhibit or support the growth of particular microbes (bottom-up). Administering a drug or viral predation can remodel microbial community structure (top-down). Our research program focuses on **how microbial populations in the human body respond to perturbations such as diet, disease, and drugs, and how these responses are linked to health outcomes**. The overall goal of our work is to empower patients to improve their health via targeted control of their microbiome. Towards this end we study how microbial communities are influenced by top-down and bottom-up regulation with a focus on three main biological questions: **1)** How does microbial metabolism of drugs influence treatment outcomes in patients? Here we develop pretherapy analysis protocols to identify patients with high-risk microbiomes; we propose novel prebiotic approaches to influence microbial drug metabolism; and we predict new microbiome/drug interactions. **2)** How do interactions between bacteria and archaea at the bottom of the microbial food chain influence access to dietary substrates? We endeavor to increase access to beneficial dietary compounds and we study how aging influences food processing in the gut. **3)** How do virus/host interactions influence population-level metabolism in the gut? We predict viral influences on microbiome health and function and we identify new viruses that infect microbial populations in the human body.

Our approaches borrow from many fields and include microbiology, ecology, metabolomics, high throughput genomics, artificial intelligence, synthetic chemistry, and imaging. Taken together, our research program forms the foundation of a new field of targeted microbiome manipulation for personalized health care.

Selected References and Products:

† Indicates that authors contributed equally to the publication

Drug and food metabolism by the human microbiome

Guthrie L, Wolfson S, **Kelly L**. The human gut chemical landscape predicts microbe-mediated biotransformation of foods and drugs. *Elife*. 2019 Jun 11;;8.

Hitchings R, **Kelly L**. Predicting and Understanding the Human Microbiome's Impact on Pharmacology. *Trends Pharmacol Sci*. 2019 Jul;;40(7):495-505.

Guthrie L, **Kelly L**. Bringing microbiome-drug interaction research into the clinic. *EBioMedicine*. 2019 Jun;;44:708-715.

Guthrie L, Gupta S, Daily J, **Kelly L**. Human microbiome signatures of differential colorectal cancer drug metabolism. *NPJ Biofilms Microbiomes*. 2017 Nov 1;;3:27.

Ian A. Downs, Olga Aroniadis, **Libusha Kelly**, and Lawrence J. Brandt. "Postinfection irritable bowel syndrome: the links between gastroenteritis, inflammation, the microbiome, and functional disease". *Journal of Clinical Gastroenterology*. 2017 Nov/Dec;51(10):869-877.

Wallace BD, Roberts AB, Pollet RM, Ingle JD, Biernat KA, Pellock SJ, Venkatesh MK, Guthrie L, O'Neal SK, Robinson SJ, Dollinger M, Figueroa E, McShane SR, Cohen RD, Jin J, Frye SV, Zamboni WC, Pepe-Ranney C, Mani S, **Kelly L**, Redinbo MR, "Structure and Inhibition of Microbiome β -Glucuronidases Essential to the Alleviation of Cancer Drug Toxicity." *Chem Biol*. 2015 Sep 17;22(9):1238-49. doi: 10.1016/j.chembiol.2015.08.005.

Microbiome-based medicine

Patent submitted to U.S. Patent Office. International Patent Application Number PCT/US2016/52742. Title: Endogenous microbial hyperswarmers protect against intestinal inflammation. **Libusha Kelly** and Sridhar Mani.

Phage discovery and characterization

Kauffman KM, Hussain FA, Yang J, Arevalo P, Brown JM, Chang WK, VanInsberghe D, Elsherbini J, Sharma RS, Cutler MB, **Kelly L***, Polz MF*. "A major lineage of non-tailed dsDNA viruses as unrecognized killers of marine bacteria." In revision, *Nature*. 2018 Feb 1;554(7690):118-122. (*=co-corresponding author)

Kelly L, Ding H, Huang KH, Osburne MS, Chisholm SW. Genetic diversity in cultured and wild marine cyanomyoviruses reveals phosphorus stress as a strong selective agent. *ISME J*. 2013 May 9. doi: 10.1038/ismej.2013.58. [Epub ahead of print]

Kelly L†, Huang KH†, Ding H, Chisholm SW. ProPortal: A resource for integrated systems biology of *Prochlorococcus* and its phage. *Nucleic Acids Res*. 2012 Jan;40(Database issue):D632-40.

Microbial genomics and ecology

Malmstrom RR†, Rodrigue S†, Huang KH, **Kelly L**, Kern SE, Thompson A, Roggensack S, Berube PM, Henn MR, Chisholm SW. Ecology of uncultured *Prochlorococcus* clades revealed through single-cell genomics and biogeographic analysis. *ISME J*. 2012 Aug 16. [Epub ahead of print].

Thompson LR†, Zeng Q†, **Kelly L**, Huang KH, Singer AU, Stubbe J, Chisholm SW. Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proc Natl Acad Sci U S A*. 2011 Sep 27;108(39):E757-64.

Li B†, Sher D†, **Kelly L**, Shi Y, Huang K, Knerr PJ, Joewono I, Rusch D, Chisholm SW, van der Donk WA. Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. *Proc Natl Acad Sci U S A*. 2010 Jun 8;107(23):10430-5.

For a complete list of Dr. Kelly's publications see:

https://scholar.google.com/citations?hl=en&user=sg7-rm4AAAAJ&view_op=list_works&sortby=pubdate

Adam Kohn, Ph.D.
Professor, Dominick P. Purpura Department of Neuroscience
Professor, Department of Ophthalmology and Visual Sciences
Professor, Department of Systems and Computational Biology
Kennedy Center, Room 822
adam.kohn@einstein.yu.edu; (718) 430-2417

My lab investigates how visual information is encoded and processed by populations of cortical neurons, and how this processing is affected by recent stimulus history, or adaptation. Our work thus addresses issues of neural coding, cortical plasticity, corticocortical signaling, and the neuronal basis of visual perception. Our approach involves multielectrode recordings in early and midlevel visual areas of anesthetized and awake, behaving macaque monkeys. We use computational methods to interpret and understand physiological data, and psychophysical methods to relate our findings to human perception. We have also begun using optogenetic tools, to further our understanding of corticocortical signaling in primate cortex.

Specific areas of focus include:

1. Understanding the coordination of spiking activity in visual cortex: The coordination of spiking activity can strongly affect encoding of sensory information by neuronal populations (Cohen and Kohn, 2011). Despite its potential importance, the properties of spiking correlations and the mechanisms that determine their strength are poorly understood. We performed one of the first quantitative studies of spiking correlations in cortex, reconciling a literature focused on spike timing synchrony with work on spike count correlations (Kohn and Smith, 2005). We extended this work using planar microelectrode arrays, elucidating the structure of coordinated activity across physical and functional space (Smith and Kohn, 2008). More recently, we have shown that the temporal coordination of spiking activity strongly influences signaling to direct downstream target network (Zandvakili and Kohn, 2015). These studies have helped clarify the structure of neuronal population responses, and begun to elucidate how dependencies among spiking activity of neurons contribute to basic cortical functions.

References

- A) Kohn A, Smith MA (2005) Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *Journal of Neuroscience* 25: 3661-3673.
- B) Smith MA, Kohn A (2008) Spatial and temporal scales of neuronal correlation in primary visual cortex. *Journal of Neuroscience* 28: 12591-12603. PMID: PMC2656500
- C) Cohen MC, Kohn A (2011) Measuring and interpreting neuronal correlations. *Nature Neuroscience* 14: 811-819. PMID: PMC3586814

D) Zandvakili A, Kohn A (2015) Coordinated neuronal activity enhances corticocortical communication. *Neuron* 87: 827-839. PMID: PMC4545497

2. Elucidating the properties and function of gamma activity: Gamma band activity (typically 30-50 Hz) has received considerable attention as a proposed coding and signaling mechanism in the brain (Jia and Kohn, 2011a). The disruption of gamma has been associated with a number of disorders, including schizophrenia and autism. Wishing to understand the potential role of gamma in cortical function, we tested the spatial and functional properties of gamma components of the LFP and found that it can share a common orientation preference across millimeters of cortex (Jia et al., 2011b), suggesting it reflects the biased representation of orientation in a cortical region rather than the functional properties of nearby neurons. We built on these findings to show that gamma amplitude and peak frequency are modulated by a number of stimulus features, and we proposed a simple computational model that could account for these features (Jia et al., 2013a). Finally, we showed that elevated LFP gamma power was associated with weak but consistent modulation of spiking activity, enhancing synchrony and coupling to downstream neurons in V2 (Jia et al., 2013b). However, inconsistent with the 'Communication through Coherence' hypothesis, we found that the downstream gamma rhythm had little effect on modulating the efficacy of input from V1. Our studies have thus provided a deeper understanding of gamma and its role as a cortical signaling mechanism.

References

A) Jia X, Kohn A (2011a) Gamma rhythms in the brain. *PLOS Biology* 9(4): e1001045. doi:10.1371/journal.pbio/1001045. PMID: PMC3084194

B) Jia X, Smith MA, Kohn A (2011b) Stimulus selectivity and spatial coherence of gamma components of the local field potential. *Journal of Neuroscience* 31: 9390-9403. PMID: PMC3133446

C) Jia X, Xing D, Kohn A (2013a) No consistent relationship between gamma power and peak frequency in macaque primary visual cortex. *Journal of Neuroscience* 33: 17-25. PMID: PMC3560843

D) Jia X, Tanabe S, Kohn A (2013b) Gamma and the coordination of spiking activity in early visual cortex. *Neuron* 77: 762-774. PMID: PMC3632874

3. Understanding neuronal population coding: Through collaboration we have explored the effect of spiking correlations on various proposed coding schemes. Using information theoretic measures, we showed that stimulus-dependent correlations contain information about stimulus identity which offsets the redundant signals provided by nearby neurons, yielding a code that is only weakly synergistic (i.e. nearly independent; Montani et al., 2007). We extended this work in Graf et al., 2011, showing that decoders that take correlation structure into account out-perform parametric decoders which assume independence. We have also quantified the information available in response latency and

proposed a simple decoder which could extract this information without knowledge of stimulus onset time (Shriki et al., 2012). Most recently, we have shown that fluctuations in population activity levels affect neuronal tuning in both a multiplicative and additive manner, in both anesthetized and awake macaque V1 (Arandia-Romero et al., 2016). This modulation, in term, differentially affects the encoding of sensory information by single neurons and small neuronal ensembles. Together, these studies have provided concrete assessments of neuronal population codes and the factors that influence them. We are building on these results to determine how noisy sensory representations are used to make perceptual decisions in awake monkeys performing discrimination tasks.

References

A) Montani F, Kohn A, Smith MA, Schultz SR (2007) The role of correlations in direction and contrast coding in the primary visual cortex. *Journal of Neuroscience* 27: 2338-2348.

B) Graf ABA, Kohn A, Jazayeri M, Movshon JA (2011) Decoding the activity of neuronal populations in macaque primary visual cortex. *Nature Neuroscience* 14: 239-245. PMID: PMC3081541

C) Shriki O, Kohn A, Shamir M (2012) Fast coding of orientation in primary visual cortex. *PLOS Computational Biology* 8(6): e1002536. PMID: PMC3375217

D) Arandia-Romero I, Tanabe S, Drugowitsch J, Kohn A, Moreno-Bote R (2016) Multiplicative and additive modulation of neuronal tuning with population activity affects encoded information. *Neuron* 89: 1305-1316.

E) Kohn A, Coen-Cagli R, Kantischeider I, Pouget A (2016) Correlations and neuronal population information. *Annu Rev Neurosci* 39: 237-256.

4. Understanding the effects of adaptation across the cortical hierarchy: The responses of most neurons are strongly influenced by recent stimulus history, or adaptation. Most studies of adaptation have focused on the the retina, LGN, and V1. We studied the adaptation properties of neurons in extrastriate area MT, and found that responsivity there was maintained for stimuli similar to the adapter and strongly reduced for different stimuli (Kohn and Movshon, 2004); in contrast, adaptation causes V1 responsivity to be most strongly reduced for stimuli that matched the adapter. We recently reconciled this apparent discrepancy, showing that adaptation effects in V1 are strongly modulated by surround suppression. Adaptation with stimuli that recruit substantial suppression results in disinhibition, yielding effects similar to those we discovered in MT (Patterson et al., 2013; Solomon and Kohn, 2014). We have used our understanding of adaptation effects in these two areas, and the circuitry linking them, to elucidate how plasticity effects cascade through the visual system. We first showed that contrast adaptation effects in MT can be explained largely by inheritance (Kohn and Movshon, 2003). However, the adaptation-induced changes in the representation of motion in V1 can derail computations performed within MT (Patterson et al., 2014). We are currently building on this work to understand how altered

population representations in the visual system give rise to altered perceptual decision making. This work is relevant for understanding brain plasticity, including recovery from traumatic injury and stroke.

References

A) Kohn A, Movshon JA (2003) Neuronal adaptation to visual motion in area MT of the macaque. *Neuron* 39: 681-691.

B) Kohn A, Movshon JA (2004) Adaptation changes the direction tuning of macaque MT neurons. *Nature Neuroscience* 7: 764-772.

C) Patterson CA, Wissig SC, Kohn A (2013) Distinct effects of brief and prolonged adaptation on orientation tuning in primary visual cortex. *Journal of Neuroscience* 33: 532-543. PMID: PMC3710132

D) Patterson CA, Wissig SC, Kohn A (2014) Adaptation disrupts motion integration in the primate dorsal stream. *Neuron* 81: 674-686. PMID: PMC3955163.

E) Solomon SG, Kohn A (2014) Moving sensory adaptation beyond suppressive effects in single neurons. *Curr Biology* 24: R1012-R1022. PMID: PMC4279707

5. Understanding visual processing of natural scenes: In recent work, we elucidated the manner in which spatial contextual effects modulate the responses of V1 neurons. Cortical neurons are known to be strongly influenced by stimuli placed outside the classical receptive field—in the surround. Descriptions of surround modulation have been based on measurements made with simple stimuli like bars and gratings. We found that existing models of surround contextual effects derived from such measurements fail to explain responses to natural scenes. Instead, surround modulation for natural scenes can be best explained by a model which considers the statistics of natural images and considers how such images can be represented most efficiently. The key insight of this framework is that surround modulation should be gated—it should be engaged for images which have similar statistical properties across space, and it should be muted when images do not. We found incorporating this insight substantially extended our ability to predict how V1 neurons respond to natural scenes.

A) Coen-Cagli R, Kohn A*, Schwartz O* (2015) Flexible gating of contextual influences in natural vision. *Nature Neuroscience*. 18: 1648-1655. PMID: PMC4624479

Ian Willis, Ph.D.
Professor, Department of Biochemistry
Professor, Department of Systems and Computational Biology
Forchheimer Building, Room 316
ian.willis@einstein.yu.edu; (718) 430-2839

Signaling pathways and transcriptional regulation in growth control and metabolism

Our laboratory is conducting basic research on the mechanisms of eukaryotic transcriptional regulation in response to nutrients and environmental and cellular stress. We are interested in defining the signaling pathways and the mechanisms that regulate transcription of ribosomal components and transfer RNAs since these processes are critically important for controlling cell growth. Deregulation of cell growth control is widely recognized as a key event in cell transformation and tumorigenesis and is relevant to a broad range of human diseases. In addition, as the synthesis of new protein synthetic machinery constitutes ~85% of nuclear gene transcription in growing cell populations, the tight coordinate control of this process, which involves all three nuclear RNA polymerases, is considered to be critical for metabolic economy. Our research programs span genetics, molecular biology, biochemistry and structural biology and utilize budding yeast, mammalian cells and mice as model experimental systems. Much of our current focus is on Maf1, a structurally and functionally novel protein that integrates the outputs of diverse signaling pathways and regulates transcription by all three nuclear RNA polymerases. The mechanisms of Maf1-dependent repression, the biological consequences of deleting Maf1 in the mouse and studies on novel downstream regulators and targets in the TOR signaling pathway are central areas of investigation in the lab. The conservation of Maf1 along with the signaling pathways that regulate Maf1 function enables the reciprocal translation of knowledge between yeast and mammalian systems and facilitates the discovery of new biology.

Genetic arrays, gene networks and functional genomics

Synthetic genetic array analysis and other systematic genome-wide genetic approaches such as synthetic dosage lethality and suppression are being conducted by robotic pinning of high density arrays of yeast strains. This technology enables the mapping of genetic interaction networks, defines the function of genes and establishes functional relationships between biochemical pathways. These genetic array-based approaches are being used to interrogate a range of biological processes including transcriptional regulation as described above. The robot also serves as a resource to other researchers at Einstein and elsewhere who are working in yeast or in mammalian systems on genes that have homologs in yeast. The integration of genetic interaction data with other large scale datasets such as DNA microarray, RNA and ChIP-sequencing and protein-protein interaction data is used to inform testable hypotheses of the systems level behavior of genes and their products.

Selected References:

- Willis, I.M., Chua, G., Tong, A.H., Brost, R.L., Hughes, T.R., Boone, C., Moir, R.D. (2008) Genetic interactions of *MAF1* identify a role for Med20 in transcriptional repression of ribosomal protein genes. *PLoS Genet.* 4, e1000112.
- Short, M.K., Hallett, J.P., Tar, K., Dange, T., Schmidt, M., Moir, R.D., Willis, I.M. and Jubinsky, P.T. (2012) The Yeast Magma ortholog Pam16 has an essential function in fermentative growth that involves sphingolipid metabolism. *PLoS One* 7, e39428.
- Lee, J., Moir, R.D., McIntosh, K.B. and Willis, I.M. (2012) TOR Signaling Regulates Ribosome and tRNA Synthesis via LAMMER and GSK-3 Family Kinases. *Mol. Cell.* 45, 836-843.
- Moir, R.D., Lee, J., and Willis, I.M. (2012) Recovery of RNA polymerase III Transcription from the Glycerol-Repressed State: Revisiting the Role of Protein Kinase CK2 in Maf1 Phosphoregulation *J. Biol. Chem.* 287, 30833-30841.
- Moir, R.D., Gross, D.A., Silver, D.L. and Willis, I.M. (2012) *SCS3* and *YFT2* Link Transcription of Phospholipid Biosynthetic Genes to ER Stress and the UPR. *PLoS Genet.* 8, e1002890.
- Lee J, Moir RD and Willis IM. (2015) Differential Phosphorylation of RNA Polymerase III and the Initiation Factor TFIIB in *Saccharomyces cerevisiae*. *PLoS One* 10, e0127225.
- Bonhoure N, Byrnes A, Moir RD, Hodroj W, Preitner F et al., (2015) Loss of the RNA Polymerase III Repressor MAF1 Confers Obesity Resistance. *Genes & Dev.* 29, 934-947.
- Mange F, Praz V, Migliavacca E, Willis, IM, Schütz F, Hernandez N and the CycliX Consortium. (2017) Diurnal regulation of RNA polymerase III transcription is under the control of both the feeding-fasting response and the circadian clock. *Genome Res.* 27, 973-984.
- Willis IM and Moir RD. (2018) Signaling to and from the RNA polymerase III Transcription and Processing Machinery. *Ann. Rev. Biochem.* 87, 75-100.

Yinghao Wu, Ph.D.

Associate Professor, Department of Systems and Computational Biology

Price Center, Room 553A

Yinghao.wu@einstein.yu.edu; (718) 678-1232

The aggregation of membrane receptors during cell adhesion initiates the elaborate networks of signaling pathways. The complexities of the networks originate from the spatial-temporal interactions of their numerous cellular components. By integrating computational analysis with experimental measurements, our lab is focusing on developing a multi-scale modeling framework to understand the molecular mechanisms of protein interactions underlying the physics of cell adhesion, as well as their biological significance.

Method Development:

The development of a multi-scale modeling framework could lead to an integrative understanding of how extracellular signals regulate cell adhesion and downstream signaling pathways in various biological systems. By designing different simulation scenarios on **molecular level**, **sub-cellular level**, **systems level** and **multi-cellular level**, the framework could serve as a guide to reveal the molecular mechanism of specific disease-related problems.

Biological Applications:

Cadherin/Wnt Signaling The epithelial-mesenchymal transition (EMT), characterized by repression of cell adhesion, is the hallmark of both normal embryonic development and cancer metastasis. Wnt is one of the most important signaling pathways triggering EMT. The key players in Wnt signaling is β -catenin, which is involved in both intercellular adhesion and gene regulation. The binding of β -catenin to the cytoplasmic domain of E-cadherin results in the stabilization of adherens junctions. On the other hand, its association with the T-cell factor/lymphoid enhancer factor (TCF/LEF) DNA binding proteins changes the transcription of target genes, initializing the canonical Wnt pathway. The fate of β -catenin in adhesion and signaling is further regulated by Wnt activation and its downstream phosphorylations. As the functions of β -catenin have been studied separately in cadherin-based junction formation and in Wnt signaling pathway, relatively little has been done to connect these two systems. Our goal for this project is to quantitatively interrogate the interplay between cadherin-mediated junction formation and canonical Wnt signaling pathway by asking the direct question: How can competition of β -catenin between these two systems serve to integrate cell adhesion with gene expression?

Integrin Signaling During cell migration, large macromolecular assemblies form at focal adhesions to transmit mechanical force and regulatory signals across cell membranes. Integrins serve as the mechanical linkages to the extracellular matrix (ECM). Their clustering based on ligand binding provides a biochemical signaling hub to direct numerous signaling and adapter proteins such as talin and focal adhesion kinase (FAK). We study the molecular mechanism of integrin clustering and its impact on mechano-chemical coupling by multi-scale modeling. Our studies can be directly compared with

cellular imaging experiments, and will give insights into the dynamic coupling of integrin clustering with downstream signaling events, for instance, the recruitment of FAK.

T-cell Signaling T cells play a pivotal role in cell-mediated immunity. The spatiotemporal patterning between T-cells and antigen-presenting cells (APCs) leads to the maturation of the immunological synapse (IS). This process is highly correlated to T-cell activation. Although size of membrane receptors was suggested to drive synaptic patterning, detailed structural information has not been used to study such sub-cellular process. Combining knowledge from molecular and cellular levels, we are using multi-scale studies to understand why specific patterns can be formed on T cell surfaces and how they are related to the intracellular signaling.

Publications:

- J. Chen, Z. Xie and Y. Wu , "Computational Modeling of the Interplay between Cadherin-Mediated Cell Adhesion and Wnt Signaling Pathway". *PLoS ONE* (2014) 9(6): e100702. doi:10.1371/journal.pone.0100702.
- J. Chen, Z. Xie and Y. Wu , "A multiscale model for simulating binding kinetics of proteins with flexible linkers". *Proteins: Structure, Function, and Bioinformatics* (2014) doi: 10.1002/prot.24614.
- Z. Xie, J. Chen and Y. Wu , "A coarse-grained model for the simulations of biomolecular interactions in cellular environments". *Journal of Chemical Physics* (2014) 140: 054112.
- Y. Wu, B. Honig and A. Ben-Shaul , "Theory and simulations of adhesion receptor dimerization on membrane surfaces". *Biophysical Journal* (2013) 104: 1221-1229.
- Y. Wu, J. Vendome, L. Shapiro, A. Ben-Shaul and B. Honig, "Transforming Binding Affinities from Three Dimension to Two With Application to Cadherin Clustering". *Nature* (2011) 475: 510-513.
- O. Harrison, X. Jin, S. Hong, F. Bahna¹, G. Ahlsen, J. Brasch, Y. Wu, J. Vendome, K. Felsovalyi, etc. "The extracellular architecture of adherens junctions revealed by crystal structures of type I cadherins". *Structure* (2011) 19:244-256.
- Y. Wu, X. Jin, O. Harrison, L. Shapiro, B. Honig and A. Ben-Shaul, "Cooperativity between trans and cis interactions in cadherin-mediated junction formation". *P.N.A.S.* (2010) 107 (41) 17592-17597.
- Y. Wu, AD. Dousis, M. Chen, J. Li, and J. Ma, "OPUS-Dom: Applying the Folding-Based Method VECFOLD to Determine Protein Domain Boundaries". *J. Mol. Biol.* (2009) 385, 1314-1329.
- M. Chen, AD. Dousis, Y. Wu, and J. Ma, "Predicting protein folding cores by empirical potential functions". *Archives of Biochemistry and Biophysics* (2009) 483, 16-22.
- Y. Wu, M. Lu, M. Chen, J. Li, and J. Ma, "OPUS-Ca: A knowledge-based potential function requiring only C-alpha Positions". *Protein Science* (2007) 16, 1449-1463.
- M. Chen, C.J. Wilson, Y. Wu, P. Wittung-Stafshede, and J. Ma, "Correlation between Protein Stability Cores and Protein Folding Kinetics: A Case Study on *Pseudomonas aeruginosa* Apo-Azurin". *Structure* (2006) 14, 1-10.
- Y. Wu, X. Tian, M. Lu, M. Chen, Q. Wang and J. Ma, "Folding of Small Helical Proteins Assisted by Small-Angle X-Ray Scattering Profiles". *Structure* (2005) 13, 1587-1597.
- Y. Yuan, Y. Wu and J. Zi, "Heat Capacities of Globular Proteins". *J. of Phys.: Condensed*

Matter, (2005) 17, 469-473.

Y. Wu, M. Chen, M. Lu, Q. Wang and J. Ma, "Determining Protein Topology from Skeletons of Secondary Structures". *J. Mol. Biol.* (2005) 350, 571-586.

Y. Wu and J. Ma, "Refinement of F-actin Model against Fiber Diffraction Data by Long-Range Normal Modes". *Biophysical J.* (2004) 86, 116-124.

Y. Wu and J. Ma, "Normal-mode-based Refinement of an F-actin Model Against Fiber Diffraction Data". *Fiber Diffraction Review*, (2004) 12, 25-28.

D. Ming, Y. Kong, Y. Wu and J. Ma, "Simulation of F-actin filaments of Several Microns". *Biophysical J.* (2003) 85, 27-35.

Y. Kong, D. Ming, Y. Wu, J. Stoops, H. Zhou and J. Ma, "Conformational Flexibility of Pyruvate Dehydrogenase Complexes: A Computational Analysis by QEDM". *J. Mol. Biol.* (2003) 330, 129-135.

D. Ming, Y. Kong, Y. Wu and J. Ma, "Substructure Synthesis Method for Simulating Large Molecular Complexes". *P.N.A.S.* (2003) 100, 104-109.

Y. Wu and J. Zi, "Universal Behavior of Localization of Residue Fluctuations in Globular Proteins". *Physical Reviews E.* (2003) 67, 041909-041912.

Qian K. Ye, Ph.D.

Associate Professor, Department of Epidemiology and Population Health
Associate Professor, Department of Systems and Computational Biology
Block, Room 310
kenny.ye@einstein.yu.edu; (718) 430-2590

Dr. Ye's research focuses on statistical design and analysis of studies in genetics and genomics. He conducts methodological research in statistical modeling and data mining with high dimensional data. He is developing new statistical and computational approaches for novel application of next generation sequencing data in biomedical research. He is also developing novel statistical approaches for quantifying genetic contributions to disease traits.

Selected Publications

Hu, J., Zhao, Z., Yalamanchili, H., Wang, J., Ye, K., and Fan, X., "Bayesian Detection of Embryonic Gene Expression Onset in *C. Elegans*", *Annals of Applied Statistics*, 2015, Vol. 9, No. 2, 950–968

Mirina A, Atzmon G, Ye K, Bergman A. "Gene size matters". *PLoS One*. 2012;7(11): e49093. doi: 10.1371/journal.pone.0049093. PubMed PMID: 23152854; PubMed Central PMCID: PMC3494661

Lin, C-Y, Lo, Y., and Ye, K., "Genotype Copy Number Variations using Gaussian Mixture Models: Theory and Algorithms", *Statistical Applications in Genetics and Molecular Biology*, 2012

Yoon S, Xuan Z, Makarov V, Ye K, Sebat J., "Sensitive and accurate detection of copy number variants using read depth of coverage", *Genome Research* 2009 Sep;19(9):1586-92. Epub 2009 Aug 5.

Sebat, J., B. Lakshmi, Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J., Leotta, A., Pai, D., Zhang, R., Lee, Y-H., Hicks, J., Spence, S.J., Lee, A.T., Puura, K., Lehtimki, T., Ledbetter, D., Gregersen, P.K., Bregman, J., Sutcliffe, J.S., Jobanputra, V., Chung, W., Warburton, D., King, M-C., Skuse, D., Geschwind, D.H., Gilliam, T.C., Ye, K., and Wigler, M., "Strong Association of De Novo Copy Number Mutations with Autism", *Science*, 316, 445-449, 2007

Zhao, X., Leotta, A., Kustanovich, V., Lajonchere, C., Geschwind, D. H., Law, K., Law, P., Qiu, S., Lord, C., Sebat, J., Ye, K., and Michael Wigler, "A unified genetic theory for sporadic and inherited autism", *PNAS*, 2007, 104(31), 12831-12836

Linkletter, C. Bingham, D., Hengartner, N., Higdon, D., and Ye, K.Q., "Variable Selection for Gaussian Process Models in Computer Experiments", *Technometrics*, 2006, 48, 478-490

Cheng, S-W, and Ye, K.Q., "Geometric Isomorphism And Minimum Aberration For Factorial Designs With Quantitative Factors", *Annals of Statistics*, 32, 2168-2185, 2004

Ye, K., Q., "Indicator Functions and Their Application in Two Level Factorial Designs", *Annals of Statistics*, 31, 984-994, 2003

Nair, V., Taam, W., and Ye, K.Q., "Analysis of Functional Response From Robust Design Study", *Journal of Quality Technology* 34, 355-370, 2002 (2003 Brumbaugh Award)