UCLA QCBio

5th Annual Retreat

Wednesday, September 25, 2019

Annenberg Community Beach House

Welcome

The QCB Community extends a warm welcome to all. Our annual retreat is an opportunity to welcome new faculty, postdocs, and students, but also to take stock of where we are, to set goals, and define agendas.

We are in the middle of a Biosciences revolution. Whereas <5% research effort in Biosciences was computational in the year 2000, by some estimates we have now reached 25%, and will reach 50% by the 2030's. The consequences of this revolution cannot be overstated. Leading research universities such as UCLA must transform their research training programs at postdoctoral, graduate and undergraduate levels, as well as the undergraduate education programs offered by Life Sciences, Physical Sciences, and the Engineering Sciences.

Quantitative and Computational Biosciences encompasses diverse approaches ranging from data processing and analysis, to data-driven and knowledge-based, statistical and mathematical modeling, to achieve prospective prediction and gain insight about the emergent properties of complex biological systems. They are applied to a broad swath of biosciences research ranging from genetics and epigenetics, electronic medical records and precision medicine, to molecular, cellular, or organismal networks whose dynamics define cell function, host-pathogen interactions, populations, or ecosystems.

QCB was initiated to support those who develop novel computational algorithms or mathematical models. Our research certainly relies on collaborating with data generators. However, excellence in research also requires continued interactions among quantitative and computational bioscientists, where the mathematics and the computational work are the focus, not only the biological question. As an interdepartmental institute, QCB has a unique role in supporting interdepartmental research excellence, research training and educational programs.

The big news at UCLA is the transformation of the Biomath department to Computational Medicine, implementing a plan first recommended by a campus-wide task force almost 4 years ago. With Eleazar Eskin chairing the Department, computational genetics will continue to thrive, and medical data sciences and math modeling will be strengthened. We welcome Noah Zaitlen (Computational Medicine and Neurology) and Paul Boutros (Human Genetics and Urology). UCLA's leadership here is evident in the wildly impactful, NIH-funded Computational Genetics Summer Institute (CGSI).

At the same time, Life Sciences departments have expanded their footprint in computational biology with the recruitment of Nandita Garud (EEB), Pavak Shah (MCDB), Eric Deeds (IBP), and Frank Alber (MIMG). We're also excited about new computational biology faculty in Engineering, including the most recent addition, Elisa Franco (MAE). These further strengthen the rapidly expanding undergraduate major in Computational & Systems Biology (Van Savage), which is in part fueled by the immensely successful new freshman Math program (Alan Garfinkel, Eric Deeds).

Indeed, QCB's integration of multiple strands of computational biology makes UCLA unique among peer institutions. That includes strong coordination between Bioinformatics (Grace Xiao), the newly established home area in Medical informatics (Alex Bui), and Genetics & Genomics (Paivi Pajukanta). Similarly, potential partnerships between Life and Physical Sciences (Roy Wollman) and Biomathematics (Eric Sobel, Tom Chou) are being explored.

At the postdoctoral and undergraduate research level the benefits of the integrated approach are also evident: The QCB Collaboratory (Matteo Pellegrini) continues to broaden its impact among UCLA's biomedical research community, graduate students, and beyond, via UCLA extension, while providing valuable postdoctoral training to QCB Collaboratory Fellows. The BIG Summer program hosted 50 highly motivated, high achieving students this year undertaking research in a variety of strands of computational biology, from genomics to dynamic modeling.

I look forward to an exciting day – I invite everyone to contribute ideas for initiatives, plans and agenda's for the coming academic year. QCB is here to support you!

Special thanks to Caroline Baron and our MBI office partners for organizing the Retreat this year!
Alexander Hoffmann

Lunch/Breakout Sessions

The QCB retreat lunch is a time to engage with other community members around topics of mutual interest and generate some recommendations to share with all. The goal is to collect ideas that we can implement in the coming year for enhancing quantitative and computational biosciences at UCLA.

Each lunch table has been assigned a specific topic:

- **Table 1:** Ideas for supporting the goals of postdoctoral fellows
- **Table 2:** Ideas for symposia/workshops
- **Table 3:** Ideas for supporting the goals of graduate students
- **Table 4:** Ideas for supporting the goals of graduate students
- **Table 5:** Ideas for improving faculty mentoring, the trainee experience, including mental health
- **Table 6:** Ideas for seminar series and speakers
- **Table 7:** Ideas for improving the diversity/inclusivity of our academic community
- **Table 8:** Ideas for symposia/workshops
- **Table 9 -12:** Ideas for any of the above topics or any other topic.

You have been assigned to a table. Look for the table # on your badge

All table discussions will present their slides in the plenary session at 4:20 p.m.

Instructions:

- Pick a scribe
- Summarize your thoughts on a few PowerPoint slides
- Send slides to uclaqcbio@gmail.com – indicate your table number
- Pick a presenter for the 4:20 p.m. session

Agenda

8:00 a.m. **Breakfast buffet in the Terrace Lounge**

9:00 a.m. **WELCOME**
SESSION I
STATUS REPORTS
- **Alexander Hoffmann**, Director, QCBio, BIG Summer
- **Matteo Pellegrini**, Director, QCBio Collaboratory
- **Eleazar Eskin**, Director of Bioinformatics Minor, Computational Genetics Summer Institute
- **Van Savage**, Director of Computational and Systems Biology Major
- **Eric Deeds**, Associate Director of Life Science Core, Freshman Math

9:40 a.m. **INVITED TALK**
- **Eric Deeds**, Associate Professor, Department of Integrative Biology and Physiology

10:00 a.m. **SELECTED TALKS**
- **Robert Foreman**, QCBio Collaboratory Fellow, Wollman Lab
- **Ha Vu**, Bioinformatics PhD student, Ernst Lab

10:30 a.m. **Coffee Break**

11:00 a.m. **SESSION II**
STATUS REPORTS
- **Grace Xiao**, Director, Bioinformatics Interdepartmental Ph.D. Program
- **Alex Bui**, Director, Medical Informatics Ph.D. Program Home Area
- **Paivi Pajukanta**, Director, Genetic & Genomics, Ph.D. Program
- **Eric Sobel**, Director, Biomathematics, Ph.D. Program

11:30 a.m. **INVITED TALK**
- **Paul Boutros**, Director of Cancer Data Science for the Cancer Center, Associate Director of Cancer Informatics at the Institute for Precision Health, and Professor of Urology and Human Genetics

12:00 pm. **SELECTED TALKS**
- **Shuo Li**, Bioinformatics PhD student, Zhou Lab
- **Igor Mandric,** QCBio Collaboratory Fellow, Halperin Lab

12.30 p.m. **Breakout Session over Lunch**
Free Time: Volleyball, Beach, Networking, Brainstorming

2:30 p.m. **SESSION III**
KEYNOTE
- **Marc Suchard,** Professor in Computational Medicine, and Biostatistics

3:10 p.m. **INVITED TALKS**
- **Pavak Shah**, Assistant Professor, Department of Molecular, Cellular, and Developmental Biology
- **Nandita Garud**, Assistant Professor, Department of Ecology and Evolutionary Biology

3:50 p.m. **SELECTED TALKS**
- **Ning Wang**, Bioinformatics PhD student, Hoffmann Lab
- **Lorenzo Boninsegna**, Postdoc, Alber Lab

4:20 p.m. **BREAKOUT SESSION REPORTS**

5:00 p.m. **POSTER SESSION in Sand & Sea**

7:00 p.m. **CONCLUDING REMARKS**

Keynote Speaker

Marc A. Suchard, M.D., Ph.D.

Professor, David Geffen School of Medicine at UCLA Departments of Biomathematics, Biostatistics and Human Genetics

Marc Suchard is helping to develop the nascent field of evolutionary medicine. This field harnesses the power of methods and theory from evolutionary biology to advance our understanding of human disease processes. Just as phylogenetic approaches have stimulated the field of evolution at large, they posses the potential to revolutionize evolutionary medicine, particularly in the study of rapidly evolving pathogens. To bridge the gap between phylogenetics and human-pathogen biology, Dr Suchard's interests focus on the development of novel reconstruction methods drawing heavily on statistical, mathematical and computation techniques. Some of his current projects involve jointly estimating alignments and phylogenies from molecular sequence data and mapping recombination hot-spots in the HIV genome.

Large-scale evidence generation across a network of databases (LEGEND) for hypertension: real-world, reliable and reproducible

Concerns over reproducibility in science extend to research using existing healthcare data; many observational studies investigating the same topic produce conflicting results, even when using the same data. To address this problem, we propose a paradigm shift. The current paradigm centers on generating one estimate at a time using a unique study design with unknown reliability and publishing (or not) one estimate at a time. The new paradigm advocates for high-throughput observational studies using consistent and standardized methods, allowing evaluation, calibration, and unbiased dissemination to generate a more reliable and complete evidence base. We demonstrate this new paradigm by comparing all hypertension treatments for a set of effectiveness and safety outcomes, producing 587,020 hazard ratios, each using methodology on par with state-of-the-art studies. We furthermore include control hypotheses to evaluate and calibrate our evidence generation process. Results agree with the limited number of randomized trials. The distribution of effect size estimates reported in literature reveals an absence of small or null effects, with a sharp cutoff at p = 0.05. No such phenomena were observed in our results, suggesting more complete and more reliable evidence.

Invited Talks

Dr. Paul C. Boutros
Director of Cancer Data Science for the Cancer Center, Associate Director of Cancer Informatics at the Institute for Precision Health, and Professor of Urology and Human Genetics, UCLA

Dr. Boutros uses big data to help optimize treatment for patients, further advancing UCLA's reputation as a world leader in the field of genomics-based cancer research. He is an integral member of the Institute of Urologic Oncology and the Eli and Edythe Broad Center for Regenerative Medicine and Stem Cell Research at UCLA. Dr. Paul Boutros, who has a doctorate in Medical Biophysics, joined UCLA from the University of Toronto, where he served as an Associate Professor in Pharmacology and Toxicology and Medical Biophysics. He was also a Principal Investigator in the Informatics and Biocomputing Program for the Ontario Institute for Cancer Research, and led Canada's national prostate cancer genomics program. His work focuses on the development of clinically useful biomarkers using genomic and data science techniques, such as next-generation sequencing, clinical and cellular imaging, machine-learning, crowd-sourcing and cloud-computing.

Variability in Tumour Presentation: The Influence of the Germline on Somatic Evolution

In essentially every tumour type, both the clinical presentation and genomic features of localized tumours vary dramatically at initial presentation. Tumours follow different evolutionary paths to reach this clinical starting-point, and subsequently evolve in different ways in response to the selective pressures of treatment. Much of this variability is stochastic, occurring through incremental mutational processes. But these mutational processes are not unbiased – certain types of mutations occur more frequently in some populations than in others. We outline here the influence of patient germline on the molecular features of a tumour at initial presentation. We start by looking at patient sex and BRCA2-carrier status, then move to study the influences of racio-ethnic effects and ultimately common germline variants in shaping tumour evolution and presentation. These data suggest there may be benefit from direct incorporation of germline features into biomarkers for early detection of aggressive tumours.

Eric Deeds
Associate Professor, Department of Integrative Biology and Physiology, Institute for Quantitative and Computational Biosciences, UCLA

Prof. Eric Deeds is a member of the Department of Integrative Biology and Physiology and the Institute for Quantitative Biology and Physiology at UCLA. Research in the Deeds lab is aimed at understanding the dynamics of complex networks in living systems. This includes understanding the self-assembly kinetics of complex macromolecular structures, the dynamics of gene regulatory networks and cell signaling networks, and cell fate decision making. The Deeds lab also develops new tools for data analysis in order to facilitate comparison of theoretical predictions with experimental results.
Prof. Deeds received his undergraduate degrees from Case Western Reserve University. His graduate work in the lab of Eugene Shakhnovich focused on computational biophysics, and he received his Ph.D. from Harvard University in 2005. He did his postdoctoral studies with Walter Fontana in the Department of Systems Biology at Harvard Medical School, where his research focused on understanding the dynamics of complex protein interaction networks. In 2010 he started as an Assistant Professor at the University of Kansas, and he was promoted to Associate professor at KU in 2016. Prof. Deeds moved his lab to UCLA in 2019.

A novel metric reveals previously unrecognized distortion in dimensionality reduction of scRNA-Seq data

High-dimensional data are becoming increasingly common in nearly all areas of science. Developing approaches to analyze these data and understand their meaning is a pressing issue. This is particularly true for the rapidly growing field of single-cell RNA-Seq (scRNA-Seq), a technique that simultaneously measures the expression of tens of thousands of genes in thousands to millions of single cells. The emerging consensus for analysis workflows reduces the dimensionality of the dataset before performing downstream analysis, such as assignment of cell types. One problem with this approach is that dimensionality reduction can introduce substantial distortion into the data; consider the familiar example of trying to represent the three-dimensional earth as a two-dimensional map. It is currently unclear if such distortion affects analysis of scRNA-Seq data sets. In this work, we introduced a straightforward approach to quantifying this distortion by comparing the local neighborhoods of points before and after dimensionality reduction. We found that popular techniques like t-SNE and UMAP introduce significant distortion even for relatively simple geometries such as simulated hyperspheres. For scRNA-Seq data, we found the distortion in local neighborhoods was greater than 95% in the reduced-dimensional spaces typically used for downstream analysis. This high level of distortion can readily introduce important errors into cell type identification, pseudotime ordering, and other analyses that rely on local relationships. We found that principal component analysis can generate accurate embeddings of the data, but only when using dimensionalities that are much higher than typically used in scRNA-Seq analysis. We suggest approaches to take these findings into account and call for a new generation of dimensional reduction algorithms that can accurately embed high dimensional data in its true latent dimension.

Nandita Garud, PhD
Assistant Professor, Department of Ecology and Evolutionary Biology, UCLA

Dr. Nandita Garud is an assistant professor in the Ecology and Evolutionary Biology department at UCLA. She leads a computational group studying how natural populations evolve and has been focusing on bacteria in the human microbiome and Drosophila melanogaster. Nandita completed her M.S. in Statistics and Ph.D. in Genetics at Stanford University where she developed a new statistical method to detect signatures of rapid adaptation in Drosophila melanogaster population genomic data. Nandita completed he rpostdoctoral work at the Gladstone Institute at UCSF studying the evolution of bacteria in the human microbiome.

Evolution in the Human Gut Microbiome

The human microbiome experiences a plethora of new mutations daily, and thus has the potential to evolve rapidly. This genetic dynamism is both an opportunity (enabling digestion of new foods) and a challenge (the evolution of drug resistance). To understand how the human gut microbiome evolves over time, we quantified the evolutionary dynamics of roughly 40 prevalent species of gut bacteria. We found that gut bacteria can evolve in humans in the space of just six months, but that over our lifetimes, the bacteria inside us are completely replaced. These results suggest that gut bacteria can evolve on timescales relevant to our health, but that they do not become so personalized that they cannot be replaced. Simply uncovering evidence for evolution in the microbiome is just the start to fully characterizing the extent and limitations of evolution in this complex community. With a bounty of genetic variation, evolutionary inquiry in the human microbiome is an exciting frontier for future research, with important biomedical implications.

Pavak Shah, Ph.D.

Pavak Shah earned his BS in Biomedical Engineering at NC State University in 2009 where he developed low cost imaging systems for infectious disease diagnostics. He earned his PhD in BME in 2014 in the lab of Nancy Allbritton at UNC Chapel Hill developing microdevices and automated imaging systems for single cell analysis and sorting. During his postdoc in Zhirong Bao's lab at Memorial Sloan Kettering Cancer Center, he studied neural morphogenesis in the C. elegans embryo and developed a real-time image analysis system for optically manipulating single cells in developing embryos and living tissues.

An Embryo's First Thought: From Form to Function in the Developing Nervous System

The origins of neuronal function during embryonic development is poorly understood. Essentially all animal embryos begin exhibiting signs of behavior and neuronal function at fairly early stages of embryonic development. How are these early neuronal circuits patterned and refined? What mechanisms define how circuits are built over time to produce defined output and perform useful computation? We aim to answer these questions by examining the development of the C. elegans nervous system. A microscope nematode, the adult nervous system of the C. elegans hermaphrodite contains just 300 terminally differentiated neurons. Its entire development is complete in ~12 hours in a transparent egg that can be easily interrogated by light microscopy. Using a combination of classical genetics, modern optogenetics, 4D light microscopy, and automated image analysis methods, we are studying the emergence of the first patterned behavior exhibited by the embryo. Ultimately, we hope to understand the emergence of functional circuits across the entire nervous system and to decode how individual circuits are patterned and then coupled together to give rise to useful behavior over the course of development.

Selected Talks

Driven Population Modeling of 3D Genome Architecture: The Integrated Genome Modeling (IGM) platform

Lorenzo Boninsegna[1,2], Asli Yildirim[1,2], Guido Polles[3], Nan Hua[3], Frank Alber[1,2,3]

[1] Institute for Quantitative and Computational Biosciences
[2] Department of Microbiology, Immunology and Molecular Genetics
University of California Los Angeles, Los Angeles (CA), USA
[3] Quantitative and Computational Biology, University of Southern California, Los Angeles (CA), USA

Accurate modeling of chromosome three-dimensional organization and location with respect to nuclear bodies is crucial to elucidate the relationship between genomic structure and biological functions. In this sense, data – driven approaches have become increasingly popular, since model resolution and accuracy both benefit from convoluting and cross-validating data from a range of different technologies. Among others, population-based deconvolution methods provide an ideal framework for systematically integrating diverse data sets into genome structure maps. We propose a novel modeling protocol, in which experimental information is incorporated by introducing additional energy terms to the model force field. Extensive calculations indicate that the equilibrated structures fully recapitulate the experimental input data. The modeling pipeline is carried out by a new platform called Integrated Genome Modeling (IGM), which is intended to be able to handle different genome types (diploid, haploid, phased, modified), nuclear shapes and data sources like Hi-C, DamID, FISH and SPRITE datasets. Currently, the implementation is being extended to enable integration of volumetric data from imaging and single-cell information as well.

Quantifying gene expression sources of calcium signaling heterogeneity

Robert Foreman[1,2], Evan Maltz[1], Gregory Johnson[3], and Roy Wollman[1]

[1] Department of Chemistry and Biochemistry, UC Los Angeles
[2] Systems Biology and Bioinformatics, UC San Diego
[3] Animated Cell, Allen Institute

Recent advances in single-cell technology are rapidly revealing cellular states and their dynamics. However, not all differences at the gene expression level are necessarily related to cell states. Differences between cells can arise from noisy effects such as transcriptional bursting. Similarly, not all phenotypic variability necessarily arises from systematic differences in gene expression. Variability can be explained by post transcriptional regulation and/or intrinsic fluctuations in components of the signaling networks. In order to clarify the sources of variability in calcium signaling, we sought to use *in situ* sequential hybridization smFISH in order to obtain highly accurate single-cell expression counts paired to measurements of a complex phenotype, calcium signaling dynamics, an emergent property of gene expression plus post-transcriptional regulation. We identify an upper bound for how much gene expression variability could arise from allele specific transcriptional bursting, and show that ~30% of calcium signaling variability is explained by gene expression.

Sensitive detection of tumor mutations from blood and its application to immunotherapy prognosis

Shuo Li[1, 2], Zorawar Noor[1], Weihua Zeng[1], Xiaohui Ni[1], Zuyang Yuan[1], Wenyuan Li[1], Edward B. Garon[1], Xianghong Jasmine Zhou[1]

[1]David Geffen School of Medicine, UCLA
[2]Bioinformatics Interdepartmental Graduate Program, UCLA

Liquid biopsy holds great promise to transform cancer care in various aspects. Sensitive somatic SNV detection from cfDNA to clinical applications is fundamental and faces unique challenges. We developed a Bayesian probabilistic framework called *cfSNV*, to infer somatic SNVs from cfDNA. Our method stratifies somatic SNV signals in cfDNA by combining estimated tumor fraction and an iterative candidate screening process, and ensures quality of variants by site-level and read-level sequencing error filtration. Tumor fractions estimated from our method are highly correlated with ground truth or independent experiments. Our method showed higher sensitivity at comparable specificity and a higher overall confirmation rate on sequential plasma and tumor samples was achieved using our method (median = 96%) than MuTect and Strelka2. Subsequently, we applied our method to plasma samples from non-small-cell lung cancer patients with immunotherapy. By incorporating clonality, we showed that our truncal tumor mutational burden provides better prognosis prediction.

Optimal design of single-cell RNA sequencing for cell-type-specific eQTL analysis

Igor Mandric[1], Tommer Schwarz[2], Arunabha Majumdar[5], Bogdan Pasaniuc[2,4,5,6]*, Eran Halperin[1,2,3,4]*

[1]Department of Computer Science, University of California Los Angeles, Los Angeles, CA, USA
[2]Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA
[3]Department of Anesthesiology and Perioperative Medicine, University of California Los Angeles, Los Angeles, CA, USA
[4]Department of Computational Medicine, University of California Los Angeles, Los Angeles, CA, USA
[5]Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA
[6]Bioinformatics Interdepartmental Program, University of California Los Angeles, Los Angeles, CA, USA
*These authors jointly supervised this work.

One of the main limitations of single-cell RNA-Seq is its high cost which prohibits population-scale analyses that aim to connect population level variation (e.g., genetics and/or disease status) with single-cell transcriptomics. We demonstrate that cell-type-specific gene expression can be accurately inferred with low-coverage single-cell RNA sequencing given enough cells and individuals. We show that taking into account all related costs such as the library preparation ones, using low-coverage single-cell sequencing can considerably decrease the cost of a cell-type-specific eQTL study without sacrificing its power. For example, we show that effective sample size of 50 can be achieved by sequencing 56 individuals with 33,000 reads per cell (2,750 cells per individual) which costs $50,000. With our proposed approach of using low-coverage sequencing, the same effective sample size can be achieved by sequencing 96 individuals at 1,500 reads per cell (2,500 cells per individual) and the total costs $25,000. We also provide a practical methodology on designing cell-type-specific eQTL studies which maximizes statistical power. Our results provide a clear pathway for the design of efficient cell-type-specific association studies that are scalable to large populations.

Annotating the human genome by integrating over a thousand epigenomic datasets

Ha Vu[1,2], Jason Ernst[1,2,3,4,5,6]

[1] Bioinformatics Interdepartmental Program, University of California, Los Angeles, CA, 90095, USA.
[2] Department of Biological Chemistry, University of California, Los Angeles, Los Angeles, CA 90095, USA
[3] Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research at University of California, Los Angeles, Los Angeles, CA 90095, USA
[4] Computer Science Department, University of California, Los Angeles, Los Angeles, CA 90095, USA
[5] Jonsson Comprehensive Cancer Center, University of California, Los Angeles, Los Angeles, CA 90095, USA
[6] Molecular Biology Institute, University of California, Los Angeles, Los Angeles, CA 90095, USA

Epigenomic marks, such as histone modifications/acetylations, open chromatin regions, etc. is a powerful source of information, revealing the dynamics and locations of DNA regulatory regions. In such context, ChromHMM was created to classify the chromatin combinatorial patterns into chromatin states and annotate genomic functional domains, hence assisting studies of the long-range chromatin interactions, nascent transcript, cellular reprogramming, etc. Previously, ChromHMM has been applied in cell-type-specific manners; creating separate functional annotations for different cell types. Currently, given the volume of epigenetic data in an ever-increasing number of cell types, we are interested in investigating the benefits a non-cell-type-specific ChromHMM state genome annotation by aggregating all available epigenetic signal tracks in different cell types. In this project, we trained a ChromHMM model using data of >1000 experiments of chromatin marks' signal profiled in 127 cell/tissue types from Roadmap Epigenetics project, to annotate the genome into 100 distinct functional states. We conducted various analysis to evaluate the benefits of using such a universal annotation in understanding numerous regulatory contexts. We showed that the increased complexity of 100-state genome annotation – afforded through augmenting data from all available cell types – results in better recovery of various functional domains and disease-associated variants, as compared to using cell-type-specfiic annotations. In particular, genomic bases prioritized by various scores—on the basis of predicted deleteriousness or conservation—are most enriched in states of promoter functionality across all the cell types. Cancer-associated somatic mutations are enriched in heterochromatin states, marked by H3K9me3. Full-stack ChromHMM states are a resource for analyzing genomes and genetic variants.

On the identifiability of gene regulatory strategies by combinations of signal-dependent transcription factors

Ning Wang[1,3], Anup Mazumber[2], Alexander Hoffmann[1,2,3]

[1]Interdepartmental Program in Bioinformatics, UCLA
[2]Department of Microbiology, Immunology, and Molecular Genetics, UCLA
[3]Institute for Quantitative and Computational Biosciences (QCBio), UCLA

Upon stimulation, cells decide which genes to activate and to what extent via a gene regulatory strategy (GRS) associated with each response gene. Prior studies have been restricted to identifying statistical correlations between SDTFs and mRNA abundance, but these approaches limit our ability to quantitatively understand the underlying regulatory mechanism. Here, we explored a quantitative modeling framework to determine the identifiability of all possible synergistic or non-synergistic GRSs involving three SDTFs based on stimulus-response (input-output) datasets. We show that in the absence of experimental measurement error all GRSs including relative regulations strengths of associated SDTFs are potentially identifiable; we report what the minimal and most informative datasets are that provide for their unambiguous identification. We then develop an error model appropriate for stimulus-response data; this model, unlike conventional counterparts, can maximize the identifiability of GRSs. Our study provides a framework for designing and quantitatively to identify underlying gene regulatory strategies.

Poster Session

1. **Biological significance of the oscillatory dynamics of NF-kB for target gene expression**
   [1]Minami Ando, [1]Shigeyuki Magi, [1]Mariko Okada

2. **Delineating the influence of tumor microenvironment on the evolution of cellular states and genetic alterations in glioblastoma**
   Nicholas A. Bayley[1,2], Henan Zhu[1], Christopher Tse[1], Lynn Baufeld[1], Laura Gosa[1], Weihong Yan[3], Timothy F. Cloughesy[1,4,5], Linda M. Liau[1,4,6], Thomas G. Graeber[1,4*], and David A. Nathanson[1,4*]
   *equal contribution from labs

3. **Ancient balancing selection maintains incompatible versions of a conserved metabolic pathway in yeast**
   James Boocock[1], Joshua Bloom[1], and Leonid Kruglyak[1]

4. **A taxonomy-aware method for background noise correction increases power to detect associations**
   Leah Briscoe[1], Brunilda Balliu[2], Liat Shenhav[3], Sriram Sankararaman[1,2,3,4], Eran Halperin[1,2,3,4,5]

5. **Regulatory-coding variant interactions shape genome structure and can be leveraged for gene prioritization of GWAS associations**
   Robert Brown[1], Arun Durvasula[2] and Sriram Sankararaman[1,2]

6. **Improving blood vessel tortuosity measurements via highly sampled numerical integration of the Frenet-Serret equations**
   Alexander Byers Brummer[1,2], Van M. Savage[1,2,3]

7. **Integrative Analysis of Oxidative Stress-Sensitive Post-translational Modifications in Cardiovascular Medicine using Machine Learning**
   Howard Choi[1,2,3], Bilal Mirza[1,2], Jie Wang[1,2], Dominic Ng[1,2], Neo Christopher Chung[1,4], Ding Wang[1,2], David A. Liem[1,2], Henning Hermjakob[1,5], Wei Wang[1,3,6], John R. Yates III[1,7], Peipei Ping[1,2,3,6]

8. **Global dysregulation of RNA editing in Schizophrenia**
   Mudra Choudhury[1], Dr. Xinshu (Grace) Xiao[1,2]

9. **Leveraging pleiotropy in genome-wide association studies across multiple traits with per trait interpretations**
   Kodi Collins[1] & Eleazar Eskin[1,2,3]

10. **Mapping global signaling state during switched RTK activation to identify the essential features of bypass resistance**
    Marc Creixell[1], Jacqueline Gerritsen[2], Song Yi Bae[1], Forest M. White[2] and Aaron S. Meyer[1]

11. **Named Entity Recognition in Spanish for Symptom-Level Phenotyping of Severe Mental Illness (SMI) from Electronic Health Records (EHR)**
    Juan De la Hoz[1], Loes Olde Loohuis[2], Mauricio Castaño[3], Janet Song[2], Susan Service[2], Terri Teshiba[2], Cristian Gallego[3], Chiara Sabatti[4], Javier Escobar[5], Victor Reus[6], Alex Bui[7], Carrie E. Bearden[2], Carlos Lopez-Jaramillo[8], Nelson Freimer[2]

12. **Developing a framework for scaling laws in neuronal branching**
    Paheli Desai-Chowdhry1, Alexander Brummer1,2, Van Savage1,2,3

13. **Modeling metabolic health in the METSIM cohort using targeted bisulfite sequencing.**
    Colin Farrell[1], Marco Morselli[2], Mila Rubbi[2], Sagi Snir[3], Matteo Pellegrini[2]

14. **Negative selection on complex traits limits genetic risk prediction accuracy between populations**
    Arun Durvasula[1], Kirk E. Lohmueller[1,2,3]

15. **Scalable prediction of polygenic risk from summary statistics in multiple traits**
    Lisa Gai[1], Sriram Sankararaman[1,2], Eleazar Eskin[1,2]

16. **Integrating data across scales to model within-host dynamics and assess zoonotic risk from novel viruses**
    Amandine Gamble[1], Jessica Y. Kasamoto[2,3], Natashia J. Benjamin[2,4], Christian T. Mason[1], Hector C. Aguilar[5], Vincent J. Munster[6], Raina K. Plowright[7], James O. Lloyd-Smith[1]

17. **Genetic architecture of an island species, *Schmidtea mediterranea***
Longhua Guo[1], Marta Riutort[2], Joshua Bloom[1], Katarina Ho[1], Zain Kashif[1], Kirk Lohmueller[3], Alejandro Sánchez Alvarado[4], Leonid Kruglyak[1]

18. **Spatiotemporal Single-Cell Atlas of Corneal Wound Healing**
Zachary E Hemminger[1,2], Jennifer Oyler-Yaniv[1], Robert Foreman[1], Roy Wollman[1,2,3]

19. **Modeling proteasome assembly pathways in bacteria**
Pushpa Itagi[1,4] , Anupama Kante[2,4], Eric J. Deeds[3,4]

20. **B cells use mechanical energy to distinguish affinity and speed up adaptation**
Hongda Jiang[1], Shenshen Wang [1]

21. **An accurate and robust imputation method MBImpute for microbiome data**
Ruochen Jiang[1], Vivian Wei Li[1], Jessica Jingyi Li[1,2,3,]

22. **Kinetic Trapping and Robustness in Proteasome Assembly**
Anupama Kante [123] , Pushpa Itagi [123], Eric J. Deeds[12]

23. **Growth and adaptation in a fungal hydraulic network**
Bohyun Kim

24. **The landscape of 3' end modifications in extracellular microRNAs**
Kikuye Koyano[1], Hyun-Ik Jun[2], and Xinshu Xiao[1,2,3]

25. **Learning a human-mouse functional genomics conservation score**
Soo Bin Kwon[1,2], Jason Ernst[1,2,3,4,5,6]

26. **Dynamic fluctuations within an epigenetic landscape underlie gene expression variability**
Ryan Lannan[1,2], Alok Maity[1,2], and Roy Wollman[1,2]

27. **Identifying Causal Variants by Fine Mapping Across Multiple Studies**
Nathan LaPierre*, Kodi Collins*, Rosemary He, Xin Huang, and Eleazar Eskin
*These authors contributed equally to this work

28. **Tunable DNA Nanocalipers to Probe Structure and Dynamics of Chromatin**
J. V. Le[1], M.A. Darcy[2], Kyle Crocker[2], Dengke Zhao[2], Ralf A. Bundschuh[2, 3], Michael G. Poirier[2, 3], Carlos E. Castro[3, 4]

29. **Quantitative modeling of pre-mRNA degradation effects on gene expression**
Diane Lefaudeux[1], Emily Chen[1], Roberto Spreafico[1], Supriya Sen[1], Alexander Hoffmann[1]

30. **Identifying effects of rare variants on gene expression with likelihood ratio test**
Jiajin Li[1], Sungoo Hwang[2], Buhm Han[3], Jae Hoon Sul[2]

31. **Comparative assessment of automatic segmentation methods for whole cell cryo X-ray tomography analysis**
Yuhui Li[,1,2], Frank Alber[1,2]

32. **Developing a mathematical model of NFκB activity in single macrophages in response to pathogens and inflammatory cytokines**
Xiaofei Lin[1,4,5], Adewunmi Adelaja[2,4,5], Brooks Taylor[3], and Alexander Hoffmann[4,5]

33. **On the Concept of Epigenetic Temperature and Spatial Organization of Chromatin in Acute Myeloid Leukemia Development**
Davide Maestrini[1], S. Branciamore[2], M. Caselle[3], and R. Rockne[4]

34. **Decoding NF-κB Dynamics Using A High-Throughput, Information-Based Approach**
Evan Maltz, Robert Foreman, Oanh Huynh, Roy Wollman

35. **Allometric Scaling of Antibiotic Efficacy**
Shaili Mathur[1,2], Portia M. Mira[1], Pamela J. Yeh[1], Christopher P. Kempes[3], Van M. Savage[1,2,3]

36. **Leveraging the UK Biobank and human liver RNA-sequence and histology data to establish the causal impact of impaired liver health on heart disease**
Zong Miao[1,2], Kristina M. Garske[1], Arthur Ko[3], Dorota Kaminska[1,4,5], Janet S. Sinsheimer[2,6], Jussi Pihlajamäki[4,7], Päivi Pajukanta[1,2,8]

37. **Mathematical modeling of cell cycle phase-specific drug response in human breast cancer cell lines**
Farnaz Mohammadi [1], Sean Gross [2], Laura M. Heiser[2], Aaron S. Meyer[1]

38. **The Gene Expression Deconvolution Interactive Tool (GEDIT): Accurate Cell Type Quantification from Gene Expression Data**
Brian Nadel, David Lopez, Dennis J. Montoya, Hannah Waddel, Misha M. Khan, Matteo Pellegrini

39. **Multiplexed Decoding of the Dynamics of Fractional Killing**
Maeve Nagle[1], Anna Pilko[1], Melton Zheng[1], Roy Wollman[1,2]

40. **Subcutaneous adipose transcriptomes reveal a novel master *trans* regulator, TBX15, controlling a co-expression network with a high polygenic risk for abdominal obesity**
David Z. Pan[1,2], Zong Miao[1,2], Kristina M. Garske[1], Sandhya Rajkumar[1,3], Arthur Ko[4], Dorota Kaminska[1,5], Janet S. Sinsheimer[1,6], Karen L. Mohlke[7], Markku Laakso[8], Jussi Pihlajamäki[5,9], Päivi Pajukanta[1,2,10]

41. **Identification of Micro-insertions and deletions (MicroInDels) in RNA sequencing data**
Giovanni Quinones-Valdez[1], Xinshu (Grace) Xiao[1,2,3,4,5]

42. **Ultrasensitive and Non-Invasive Cancer Detection and Tissue-of-Origin Prediction using Cell-free DNA Methylation Sequencing Data**
Mary Same[1], Wenyuan Li[1], Shuli Kang[2], Qingjiao Li[1], Yonggang Zhou[1], Fengzhu Sun[2], Chun-Chi Liu[3], Lea Matsuoka[4], Linda Sher[5], Wing Hung Wong[6,7], Frank Alber, Steven M. Dubinett[9], Xianghong Jasmine Zhou PhD[1,8]

43. **Time-dependent antigen diversity in vaccination for eliciting broadly neutralizing antibodies against highly-mutable pathogens**
Jiming Sheng[1], Shenshen Wang[1]

44. **Signal-dependent transcription factor activity remodels chromatin to mediate melanoma dedifferentiation**
Katherine M. Sheu[1,2] ,Yeon Joo Kim[1,3], Alexander Hoffmann[2], Thomas G. Graeber[3], Antoni Ribas[3]

45. **Unsupervised discovery of structure and protein composition of macromolecular complexes**
Jitin Singla[1,2], Min Xu[3], Frank Alber[1,2]

46. **Quantifying temporal information accumulation by learning hidden Markov model for biochemical signaling dynamics**
Ying Tang[1,2], Adewunmi Adelaja[1,2], Xiaofeng Ye[3], Eric Deeds[1,4], Roy Wollman[1,4,5], Alexander Hoffmann[1,2]

47. **Comparisons Between Genetic Pathways related to copy number alterations and chromosomal instability**
Estelle (Ning) Yao[1*], Nikolas Balanis[2*], Thomas Graeber[2*]
* These authors contributed equally

48. **Population Based 3D Structure Analysis of the Human Genome**
Asli Yildirim[1], Nan Hua[1], Guido Polles[1], Frank Alber[1]

49. **Studying single cell variation in chromosome structures by dimension reduction**
Yuxiang Zhan[1,2], Frank Alber[1,2]

Collaboratory Fellows 2019-2020

https://qcb.ucla.edu/collaboratory/people/

Please visit our website to learn more about the Collaboratory, our classes offered, and class schedule and of course to learn more about our Postdoctoral Fellows.

http://qcb.ucla.edu/collaboratory

Welcome our Incoming Bioinformatics Students!

Welcome our Incoming Medical Informatics Students!

Welcome our Incoming Biomathematics Students!

- **Christine Craib** (Math, University of North Carolina Wilmington)
- **Mariana Harris Heredia** (Applied Math, Instituto Tecnológico Autónomo de México)
- **Rachel Mester** (Applied Math, Columbia University New York - two years in analytics at EPIC)
- **Gaiting 'Gary' Zhou** (Math, Duke University)
- **Xinzhe Zuo** (Math and Physics, UCLA)

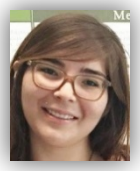
Welcome our Incoming Genetics & Genomics Students!

Kevyn Hart
BS in Bioengineering with Honors, Minor in Bioinformatics, UC Santa Cruz

Serina Huang
BS in Chemical Engineering, UC San Diego

Seung Hyuk (Tony) Lee
BS in Biochemistry, Minor in Chemistry, University of Washington

Misty Knight
BS in Biology: Molecular, Cell & Development, University of Washington



Aileen Nava
BS in Biological Science, Minor in Chemistry, Cal State Fullerton

Welcome new Faculty!

Paul Boutros

Professor, Departments of Human Genetics and Urology

Paul Boutros was recruited from the University of Toronto and the Ontario Institute for Cancer Research where he established as a leader in the precision medicine of cancer via multi-omic data integration and modeling. He leads the Clinical Cancer Genomics efforts of the Institute for Precision Health., and Director of Cancer Science Data at JCCC.

Erik Deeds

Associate Professor, Department of Integrative Biology and Physiology

Eric Deeds was recruited jointly between QCBio, the LS Dean's office and IBP to contribute intellectual leadership to the highly successful math for life science pre-majors education program established recently by Alan Garfinkel. His research interests are in the stochastic and dynamical systems properties of molecular assemblies, regulatory networks, and cell fate decision making. His lab moved Jan 1, 2019 from the University of Kansas to Boyer Hall 5th floor.

Elisa Franco

Associate Professor, Mechanical and Aerospace Engineering

Elisa Franco got PhD degrees from the University of Trieste, Italy (Information engineering, 2008) and Caltech (Control and Dynamical Systems, 2012), and She joined UCLA in Fall 2018 as an Associate Professor in Mech. and Aero Engineering and Bioengineering, after 7 years in Mech. Engineering at UC Riverside. Her research is: 1) Design and synthesis of dynamic nucleic acid systems with applications in biomaterials science and biomedicine. 2) Mathematical modeling to elucidate design principles for temporal signal processing in natural and synthetic gene networks.

Nandita Garud

Assistant Professor, Department of Ecology and Evolutionary Biology

Nandita received an M.S. in Statistics and Ph.D. in Genetics at Stanford University with Dimitri Petrov characterizing rapid adaptation in Drosophila melanogaster population genomic data. As a postdoctoral scholar at the Gladstone Institute at UCSF with Katie Pollard she is studied the evolution of bacteria in the human microbiome. She established her lab at UCLA in Spring 2019.

Pavak Shah, Ph.D.

Assistant Professor, Department of Molecular, Cell and Developmental Biology

Pavak earned his BS and PhD in Biomedical Engineering developing low cost imaging systems for infectious disease diagnostics (2009) and microdevices and automated imaging systems for single cell analysis and sorting (2014). During his postdoc in Zhirong Bao's lab at Memorial Sloan Kettering Cancer Center, he studied neural morphogenesis and developed a real-time image analysis system for optically manipulating single cells in developing embryos and living tissues.

Noah Zaitlin, Ph.D.

Associate Professor, Neurogenetics

Dr. Noah Zaitlen uses computational and medical genomics to identify and characterize the processes that are disrupted in human disease and mitigated through clinical treatments. His lab collaborates closely with both clinical neurologists and molecular biologists, linking functional genomic data to patients' medical records. Dr. Zaitlen earned his PhD in bioinformatics and systems biology from the University of California, San Francisco, completed postdoctoral training at the Harvard School of Public Health, and established his lab at the Lewis Sigler Institute at UCSF before coming to UCLA.

A B.I.G. Thank you

To all mentors for a successful 2019 Program!

B.I.G. SUMMER – Bruins-In-Genomics

Bruins-In-Genomics (B.I.G.) Summer Research Program is an 8-week full-time immersion program for undergraduates interested in learning how to read and analyze genes and genomes. Through this program students have the opportunity to experience graduate-level-cutting-edge research in UCLA laboratories and learn some of the latest research methods to solve real-world problems.

Please visit our website to learn more: https://qcb.ucla.edu/big-summer/