Research Experiences (400 char summaries)

Bulyk Lab
Contact: Laurie

I worked in Prof. Bulyk's lab where I analyzed homeodomain transcription factor binding sites by performing protein-binding microarray experiments, resulting in a 2008 publication in the journal *Cell*. I was fascinated by how a computer could interpret the minutely polka-dotted microarrays and transform them into binding motifs, which inspired me to further explore quantitative analysis of biology.

Eddy Lab
Contact: Sean

Admitted to the competitive Janelia Farm (JF) Summer Scholars program, I worked with Dr. Eddy at Howard Hughes Medical Institute JF Research Campus in Ashburn, VA. I used Hidden Markov Models to create a robust null model for HMMER, sequence homology software developed in the Eddy lab. I completed this project in two months in Python and then rewrote it in two months in C, the language of HMMER.

Gifford Lab
Contact: Robin

I worked in Prof. David Gifford's lab in the MIT Computer Science and Artificial Intelligence Laboratory, where I worked on yeast metabolic genome networks. I hypothesized methods of information content in graphs can inform synthetic lethality of a gene, and learned Java to implement the network in JGraphT. My metrics indicated purely genomic methods do not fully explain biological interactions.

Seung Lab
Contact: Srinivas

In Prof. Sebastian Seung's lab at MIT's Department of Brain and Cognitive Sciences, I pursued computational image analysis. I analyzed neuron orientation in the rabbit retina inner plexiform layer (IPL) and found that neuron fragments in this 3D image lie along the axis of information transmission, and do not significantly traverse the direction perpendicular to of signal flow.

Broad Institute
Contact: Pablo

Working with Dr. Jill Mesirov at the Broad Institute, I developed REVEALER, an algorithm that integrates genomic and functional data to infer new associations. A pathway may be broken in many ways, but produce the same phenotypic output.

REVEALER finds candidate explanations of a phenotype by removing samples with known causal events and uncovering new events via a mutual information metric.

Leadership Experiences

Describe or list any leadership experiences, such as class president, committee chairperson, or scouts, with dates; also include a reference for each experience with the reference's name and contact information for verification. If there is no specific contact person, please list someone at the organization who can confirm your experience. Click "Save leadership experiences" when you are finished. There is a 400 character limit.

Co-chair, Intelligent Systems for Molecular Biology Student Council Symposium 2012 - Alex Goncearenco [redacted]

DanceTroupe; Choreographer, Publicity Chair – Catherine Johnson [redacted]

Baker House Social Chair (2008) – Jonathan Nolan [redacted]

Bernard M. Gordon-MIT Engineering Leadership Program – Leo McGonagle [redacted]

Leadershape – Kirk D. Kolenbrander [redacted]

Teamwork Experiences

Describe or list any teamwork experiences (which can include research or academic projects as well as varsity sports, extracurricular groups, or clubs) with dates; also include a reference for each experience with the reference's name and contact information for verification. If there is no specific contact person, please list someone at the organization who can confirm your experience. Click "Save teamwork experiences." There is a 400 character limit.

Co-chair, Intelligent Systems for Molecular Biology Student Council Symposium 2012 - Alex Goncearenco [redacted]

Biological Engineering Design final project (Spring 2010) – John Essigmann [redacted]

MIT DanceTroupe; Dancer (Fall 2007 – Spring 2011) – Catherine Johnson [redacted]

MIT Lightweight Men's Crew; Coxswain (Fall 2006 – Spring 2007) – Seth Davis [redacted]

Memberships & Certifications

Describe or list any educational or professional memberships such as IEEE, SWE, or Tau Beta Pi, and describe or list any certifications, such as Engineer-In-Training. Please include membership start date or certification date. Click "Save memberships & certifications" when you are finished. There is a 400 character limit.

Developing biomedical research at MIT/Skolkovo Institute of Technology in Russia (Summer 2012, expected)

Co-Chair for Intelligent Systems for Molecular Biology Student Council Symposium 2012

Member, International Society of Computational Biology (since November 2010)

Member, Society of Women Engineers (since May 2011)

Certified Russian-English Medical Interpreter (since August 2011)

Community & Volunteer Work

Provide a summary of volunteer work and experiences, interests and/or hobbies; also include the approximate number of hours a week for each experience, and a reference for each experience with the reference's name and contact information for verification. If there is no specific contact person, please list someone at the organization who can confirm your experience. Click "Save community & volunteer work" when you are finished. There is a 400 character limit.

We Teach Science Foundation, 8th grade Algebra Tutor (2h/wk 10/11–present) – Camille Stone [redacted]

Science Club for Girls, 2nd grade Mentor (4h/wk 1/11-5/11) – Rosalind Gould [redacted]

Mission Hill Middle School, Math tutor (2h/wk 2/2009–5/2011) – Ann Ruggerio [redacted]

Kappa Alpha Theta (2h/wk 9/2007–6/2010) – Victoria Thomas [redacted]

Summary of Goals

In your own words, provide a summary of your educational program objectives and your long-range professional goals.

As part of this statement, we are interested in your ideas about:

- the kinds of research in which you would like to be engaged during your graduate study or in the longer term; or

- specific research questions that interest you and how you became interested in them.

Please discuss these research interests in sufficient detail for an expert who is technically competent in your field to judge your understanding of the questions to be addressed. This includes relevant hypotheses and approaches one might take to answering the questions, and other research principles required to investigate the research area you identify.

We are interested in not only the science, but also your longer-term goals and how the science fits into your life as an individual. We do not want this to look like a grant submission.

Your response will be limited to 3,000 characters, including spaces. There is no extra space for citations. If you are writing this text elsewhere and copy-pasting it into this box, be aware that some word processing programs will transfer spaces and returns differently.

Current *in vitro* models of tumors are wrong. Working at the Broad Institute, I saw results from cell lines extrapolated to a whole class of cancers, but less than 10% of tumor cells even survive *in vitro* and monoclonal lines do not represent tumor heterogeneity. Certainly, there are successful *in vitro* cases, such as PLX4032 designed to target BRAF's V600E mutation, but most drugs designed for a cell line are unsuccessful in clinical trials. Before a drug even reaches the clinic, we don't know why applying a drug to monoclonal cell lines results in resistance, and whether survival occurred on a single-cell basis or was driven by intercellular interactions. Relying on statistics and success "most of the time" is an antiquated view – we now have the tools to understand the mechanisms behind an ineffective drug.

I want to shatter current acceptance of monoclonal cell lines as usable models of tumors. I will develop an *in silico* model of *in vitro* heterogeneity of breast cancer (BRCA) cell line SK-BR-3 in response to Trastuzumab (Herceptin, a BRCA drug used in HER2 positive cases such as SK-BR-3) by analyzing single-cell RNA-Sequencing (RNA-Seq) data, high-resolution snapshots of a cell's transcriptome and robust determination alternative splice sites. I hypothesize the heterogeneity of a cell population can be modeled by extracting network information from many single-cell transcriptomes. First, I will develop RNA-Seq assemblers tailored to the very small DNA sample amount of a single cell, less than a nanogram. Second, I will analyze differential transcripts between pairs of neighboring cells, accounting for differences in cell cycle stage using documented cell cycle genes. I expect to see a spectrum of active transcription in secretory pathways of neighboring respondent and healthy survivors, indicating communication between these cells and identifying specific transcripts potentially useful to survival. Using these differentially expressed genes and known interaction networks, I will develop a network of cell-cell interaction between adjacent SK-BR-3 cells in response to Herceptin.

While I use cancer as an example in my proposed research, I will develop methods for use across disease boundaries and throughout biomedicine. This project can be applied to biological warfare by finding mechanisms of resistance and driving development of vaccines and prophylactics against biological agents. To maximize my impact on revolutionizing bioinformatics, I want to work at the interface of biomedicine and computation as a professor at a research institution or in a national lab, where my quantitative skills will filter noisy biological data into novel discoveries. Seizing opportunities to work with people of diverse experiences and expertise, our strengths together can solve problems in completely unexpected ways. By obtaining the NDSEG fellowship, I would have the freedom to forge high-risk collaborations and work on groundbreaking science.