Systems Biology PhD Program at Harvard University

Program Contacts

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
<th>Phone</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tim Mitchison</td>
<td>Program Co-director</td>
<td>432-3805</td>
<td><a href="mailto:timothy_mitchison@hms.harvard.edu">timothy_mitchison@hms.harvard.edu</a></td>
</tr>
<tr>
<td></td>
<td>Asst: Andrea Chen</td>
<td>432-6839</td>
<td><a href="mailto:Andrea_chen@hms.harvard.edu">Andrea_chen@hms.harvard.edu</a></td>
</tr>
<tr>
<td>Andrew Murray</td>
<td>Program Co-director</td>
<td>496-1350</td>
<td><a href="mailto:amurray@mc.b.harvard.edu">amurray@mc.b.harvard.edu</a></td>
</tr>
<tr>
<td></td>
<td>Asst: Linda Kefalas</td>
<td>496-1350</td>
<td><a href="mailto:lkefalas@mc.b.harvard.edu">lkefalas@mc.b.harvard.edu</a></td>
</tr>
<tr>
<td>Liz Pomerantz</td>
<td>Program Coordinator</td>
<td>432-5876</td>
<td><a href="mailto:Elizabeth_pomerantz@hms.harvard.edu">Elizabeth_pomerantz@hms.harvard.edu</a></td>
</tr>
<tr>
<td>Becky Ward</td>
<td>Executive Director</td>
<td>432-3701</td>
<td><a href="mailto:becky@hms.harvard.edu">becky@hms.harvard.edu</a></td>
</tr>
<tr>
<td>Jason Millberg</td>
<td>ChemBio Coordinator</td>
<td>432-7935</td>
<td><a href="mailto:Jason_millberg@hms.harvard.edu">Jason_millberg@hms.harvard.edu</a></td>
</tr>
</tbody>
</table>

Program Advising

The class advisors will lead a week-long orientation for incoming students at the end of August. The orientation will include a set of lectures and activities that will introduce students to the many resources at and around Harvard and will answer their questions regarding research, academics and the graduate program. Students will also be paired with a senior graduate student mentor during the orientation. Incoming students will meet with the class advisors individually at the beginning of each semester to plan their initial program of graduate study. Class advisors will be available to meet with students at any time during their graduate career.

Class Advisors

<table>
<thead>
<tr>
<th>Name</th>
<th>Office Location</th>
<th>Phone</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doeke Hekstra</td>
<td>Harvard University 52 Oxford St Northwest 311.11</td>
<td>617-496-4740</td>
<td><a href="mailto:doeke_hekstra@harvard.edu">doeke_hekstra@harvard.edu</a></td>
</tr>
<tr>
<td>Sahand Hormoz</td>
<td>360 Longwood Avenue LC-01052 Boston, MA 02215</td>
<td>617-582-9717</td>
<td><a href="mailto:sahand_hormoz@hms.harvard.edu">sahand_hormoz@hms.harvard.edu</a></td>
</tr>
</tbody>
</table>

IDP

Systems Biology students are required to complete an IDP meeting with their Dissertation Advisor (or an alternate Harvard faculty mentor of their choosing) annually. An Individual Development Plan (IDP) provides students with the opportunity to think about their training objectives, their progress towards them, and to set and/or refine goals for the future with their mentor. The National Institutes of Health (NIH) encourages trainees to make Individualized Development Plans to help them prepare for academic and nonacademic careers.

Because the beginning of a new year is an ideal time for self-reflection and planning, we ask that students **complete the IDP in January, and have the yearly planning meeting before February 15th** of every year. Students must notify Liz Pomerantz, the Program Coordinator, (elizabeth_pomerantz@hms.harvard.edu) of the IDP meeting date and the name of the faculty member that they met with. For G1 students, the Spring semester meeting with their faculty advisors that takes place in January will serve as this planning meeting.

Students and their mentors should read the *Molecular Cell* article **“Yearly Planning Meetings: Individualized Development Plans Aren’t Just More Paperwork”** that describes an approach to implementing IDPs and the benefits of IDPs. We also encourage students to visit the Harvard
Office of Career Services and [http://myidp.sciencecareers.org](http://myidp.sciencecareers.org) where they provide additional IDP resources and excellent articles related to mentorship and science careers.

Sharing the IDP form is not a requirement, nor will the IDP form be kept on file by SysBio. Students are free to share as much or as little of the plan as they feel comfortable. Note that the IDP process will be most effective if used to guide candid discussions with an advisor.

**Blank IDP Form**
**Example IDP Form**

**Laboratory Rotations**
Rotations allow students to explore different research areas, identify potential collaborators, and experience the environment in different research groups. The purpose of the rotation is to facilitate the choice of the dissertation laboratory, not to accomplish a research project. Students in the Systems Biology Program are expected to take 2-4 laboratory rotations before selecting a Dissertation Advisor. The program does not set time limits on rotations, but most rotations are expected to be 4-12 weeks long. Rotations with non-training program faculty are permitted but require approval of the program. Students should inform the program coordinator when they begin and complete their rotations.

**First year students must choose their dissertation laboratory no later than June 30th.**

**Teaching**
Students are required to act as teaching fellows (TFs) in at least one course. We recommend that students complete this requirement by the end of their second year. Students should inform the program coordinator office as they make plans to fulfill their teaching requirement. TF positions are arranged by each individual department so there are no universal deadlines or contacts. It is usually best to contact the faculty member or preceptor of the course that you are interested in teaching. Students who wish to perform additional teaching, beyond the required one semester, must receive permission from the program. They will fill out the [additional teaching request form](http://myidp.sciencecareers.org). This is to ensure that your teaching aligns with your career and research plans.

**Course Requirements**
Students are required to take SB212: Communication of Science (fall G1), SB320: Quantitative Measurement and Analysis (spring G1), SB300: Introduction to Systems Biology (all year G1), MedSci300: Conduct of Science (fall of the G2) and three science courses chosen in consultation with their class advisors. A current list of courses students commonly take is provided to students at the beginning of the year.

**Registration**
Students can register from Monday August 20th, 9 a.m. through Tuesday September 4, 11:59 p.m. at [www.my.harvard.edu](http://www.my.harvard.edu). Once a student successfully registers, their registration hold will be lifted and removed from their Student Home.

**Enrollment**
Course enrollment opens on **Tuesday, September 4, 2018** - the first day of classes. It is an online process. Students officially enroll in courses by submitting a study card at [my.harvard.edu](http://my.harvard.edu).
In order for students to be considered a full time student, they must sign up for 16 units of credit each semester (the equivalent of 4 courses).

- Students who are not taking 4 “real” courses and have not joined a lab should sign up for the rotation course, **SysBio399**, catalog # 5863, the appropriate number of units.
- Students who have permanently joined a lab should use **SysBio350**, catalog # 8370, (under their PI’s name) the appropriate number of units.

There is no charge for adding or dropping a course during the first three weeks of the term. After the first three weeks there is a $10 fee each time a petition is filed.

**Study Card Day: Wednesday**, September 12<sup>th</sup>.  
Completed Study Cards are due by 11:59 p.m. Students who fail to submit a Study Card will be charged $40 for each week late. Any Study Card submitted after Study Card Day requires the approval of the instructor for every course listed.

October 22, 2018 is the deadline to add a fall 2018 course.  
November 5, 2018 is the deadline to drop a fall 2018 course.

**Cross Registration** [https://registrar.fas.harvard.edu/cross-registration](https://registrar.fas.harvard.edu/cross-registration)  
Students can cross-register with other Harvard Schools and MIT if there are courses that are of interest to them. Please note that some schools have different registration deadlines.

**Quarter and Nanocourses**  
A quarter course is a half semester course that focuses on a specific topic, usually in the area of expertise of the faculty. The class meets for one 2-hour session per week. Meeting times are usually arranged at the initial session for convenience of faculty and students.

Nanocourses are short graduate-level courses consisting of two class meetings which cover a specific subject in depth. Six nanocourses are equivalent to one half course. Students register for credit on their study cards in the semester that they plan to complete their sixth nanocourse, or when they plan to complete a combination of 3 nanocourses and one quarter course.

**Preliminary Qualifying Exam Guidelines (PQE)**

**Part 1 of the PQE**  
Part 1 of the PQE must be completed no later than June 1<sup>st</sup> of the first year.

A project proposal of 1-2 paragraphs must be submitted no later than February 1<sup>st</sup> to Liz (elizabeth_pomerantz@hms.harvard.edu).

This exam is intended to be a creative exercise in biological theory, computation or informatics. Students will formulate a question related to any problem in biology (on any scale) and develop a simple set of equations and/or a computational or bio-informatic analysis designed to address the question in a quantitative way.

This exam is intended to catalyze a period of exploration and creativity as well as exploration of an area of biology that is new to the student. Students are encouraged to discuss possible questions and models with each other and with faculty in preparing for the exam, but the final
project should be their own work. The project should not be based directly on work from a rotation or a course.

Students will prepare a short written summary and an oral presentation on their project. The written summary should be no more than 4 pages.

- Background
- Question
- Approach
- Results/Conclusion
- Future Directions
- What I learned

Students should also think about whether the results are or are not consistent with the literature and think about some experiments or other approaches to test their project.

The proposal should be submitted to the committee members and Elizabeth Pomerantz 1 week prior to the oral defense. In the oral exam students will present the results of their analysis. If students choose to use PowerPoint, the number of slides is limited to 10.

The presentation will be made to 3 Program faculty members, with one of them being a Co-director. The examination committee will ask questions about the project itself and about background information associated with the project.

The committee may also ask explore the student’s general knowledge beyond that of the project. It is expected that students will be familiar with the content covered in *Essential Cell Biology* by Bruce Alberts et al. These general questions will help the faculty committee assess if the student is adequately prepared to undertake a PhD project. Gaps in general knowledge not associated with the project will not affect the outcome of the exam (pass/fail) but may result in the student being assigned additional reading or coursework.

**Possible outcomes:**

Students pass and move into their dissertation research.

Serious flaws are found in the proposal. Students will be asked to correct flaws or submit a new problem and solution and defend it no later than August 31. Failure to do so will result in the student not being allowed to register for the second year of graduate school. In some rare cases, students may be counseled to consider leaving the program.

Important gaps are found in the student’s education. In this case the committee may assign additional reading or coursework, in consultation with the student’s advisor. This will not affect passing or retaking the exam.
The PQE Part 2
Exam must be held no later than March 31st of the student’s second year.

Assembling the Committee
Subject to program approval, any three Systems Biology Program faculty may be on the PQE Part 2 committee (excluding the Dissertation Advisor). Members of the PQE committee may also be appointed to the student’s DAC, however they are not required to be. A Program Co-director will attend every exam to ensure uniformity.

The student should email the Systems Biology Coordinator with their proposed committee members and a brief abstract summarizing their PQE topic in December. Once the program has approved the committee members, the student should invite the faculty. Students should notify the program coordinator when the faculty members have agreed to be on the committee. The coordinator will then schedule the exam.

Advisor Role
The advisor should view the process of preparing for the exam as a training opportunity and should guide their students in planning the dissertation project and in writing the proposal. However, the advisor must not write any part of the proposal, and the ideas in the proposal must be the student’s. The Program expects that preparation for the PQE Part 2 will be an iterative process in which the student expresses their ideas to the advisor, either verbally or in writing, and receives constructive criticism and feedback, leading to another cycle of creative thought, literature review, and discussion with the advisor.

Prior to the exam, the Advisor will be invited to speak briefly with the chair of the committee.

Written Proposal
Students will prepare and defend an original research proposal derived from the student’s proposed dissertation research. The proposal should define the important questions to be addressed, provide adequate background and describe some details of experiments, computation and/or theoretical work to be undertaken. Examples of successful proposals will be provided to students. The proposal should be designed so that the work can realistically be completed in approximately four years. It is appreciated that research directions may change during the course of the dissertation research.

Students are strongly encouraged to discuss their proposal and practice their oral presentation with faculty, postdocs and students in preparing for the exam.

The final written proposal should be submitted to the student’s Committee and the Systems Biology Coordinator by 5pm one week prior to the oral defense. The overall length of the proposal, not including figures and references, should not exceed ten pages.

The general format for the written proposal is:
• An abstract summarizing the questions, aims and significance of the proposed research
• A section on relevant background
• The specific aims of the research proposal
• A brief description of how the research aims will be accomplished, and what the anticipated outcomes are – both positive and negative.

Diagrams and figures can be included in the body of the proposal for clarity; these items will not be counted against the proposal’s length limit.

**Oral Exam**
Students should not supply food or drinks for their committee. The format of the oral portion will be:

• The exam is organized around the presentation of the proposal in the form of a “chalk talk.” PowerPoint presentations are NOT allowed.
• During the presentation and after it, faculty members will ask questions regarding the proposed research. Faculty may also probe the student’s general knowledge, beyond the specifics of the proposal.
• The oral exam is expected to last no longer than two hours.

**Outcomes**
The faculty committee will convene to evaluate the student’s performance on the exam. The evaluation will take into consideration both the written and the oral parts of the exam. The outcome of the examination will be communicated to the student by the examining faculty within a week of the end of the exam (often on the same day as the exam).

The faculty will consider whether or not the student successfully justified the proposed research, the significance of the research, and will evaluate the degree of independent thinking that went into the proposal, the clarity of the writing, and the student’s breadth of knowledge relevant to the proposed research.

There are several possible outcomes:

• The student will receive a pass and continue with the proposed research.

• The faculty may find problems with the proposed research and/or the presentation (either written or oral). The student may then be asked to prepare and defend a new proposal. The student must retake the exam no later than the end of June.

• If the problems with the proposal and/or presentation are serious, the faculty may conclude that the student is best served by leaving the Program to pursue other interests. In such cases, the student will be asked to leave the program at the end of the semester.

**The Next Step**
After passing the PQE, the student will assemble a Dissertation Advisory Committee (DAC). The initial meeting of the DAC should take place within 6 months of approval of the student’s PQE proposal.
Dissertation Advisory Committee (DAC)
After passing the PQE, a DAC of at least three faculty members and the student’s Dissertation Advisor(s) must be appointed 4 months after their PQE exam and a meeting scheduled for 6 months after their PQE exam. This will allow students to take maximal advantage of their committee’s expertise. At least one committee member should be from the Program in Systems Biology. Other members may be from outside the Program and, if they choose, one member may be from outside Harvard.

The Committee must meet with the student at least once a year through G5 and every six months thereafter, until Ph.D. dissertation writing is underway. However, students are encouraged to consider more frequent meetings (every 6 months is ideal). The Chair of the DAC is responsible for the preparation of the DAC Report, which should be signed by all committee members at the conclusion of each meeting and submitted to the Systems Biology Coordinator. Students will be allowed to register for the upcoming year only if their Dissertation Advisory Committees have met and filed a formal report (see the attached form) within the past twelve months.

Role of the DAC
The role of the DAC is to assist the student in defining the dissertation project, review scientific progress, offer critical evaluation, suggesting extension or modification of objectives, arbitrate differences of opinion between the student and the advisor if they arise, and decide when the work accomplished constitutes a dissertation. Our hope is that the committee will help students in the early stages to get their research off to a good start, and that they will be a resource for students at any point during their graduate career.

At the start of the meeting, the student will be asked to leave the room for a few minutes. This gives the committee an opportunity to speak with the Dissertation Advisor but this should not be considered evaluative. Similarly, the Dissertation Advisor will also be asked to leave the room so that the student can voice any issues that they might feel more comfortable discussing only with the committee

Procedures for Setting up DAC meetings
Students will schedule all meetings and will notify the systems Biology Coordinator once a date and time is set.

Summary of Progress and DAC Report Policy
Students are to submit a brief summary of progress (five or fewer pages not including images and references) to their Dissertation Advisory Committee and Systems Biology Coordinator at least one week before the meeting and be prepared to give a twenty minute presentation. The student is also responsible for bringing a copy of the DAC report to each meeting. This report is to be filled out by the Chair of the committee and returned to the Systems Biology Coordinator immediately following the completion of the meeting.

The presentation can include the student’s overall goals, progress that has been made (show data) and plans going forward. The student should also discuss any other plans for the coming year such as teaching, meeting, courses, etc with the committee so all can get some sense of how time will be spent during the coming year.
Dissertation Preparation and Defense

Preparation for the Dissertation Defense:
The Dissertation Advisory Committee, in consultation with the Dissertation Advisor, determines when it is time for a student to stop laboratory work and begin to write their dissertation. Once a student has been given permission to write their dissertation, they must contact the Program Coordinator to schedule an appointment to discuss requirements, dates and receive their dissertation information packet.

The FAS registrar specifies deadlines by which the dissertation must be submitted and the dissertation examination passed to receive the Ph.D. diploma in November, March, or May of each academic year. A dissertation information packet is available in the Program Office specifying the steps to be taken when the student is ready to apply for the Ph.D. degree and the various forms that need to be submitted. The information packet will be thoroughly reviewed with the student by the program coordinator. The first step is completion of the “application for degree” form. The deadline for submitting this form can be more than three months before the student expects to receive the degree.

Examiners:
The Dissertation Examining Committee evaluates a student’s dissertation defense. The student and the student’s dissertation advisor must select at least three examining committee members: an examination chair and two examiners. The Program Director must approve the members of the examining committee.

- The chairperson of the DAC should preferably chair the examination. The examination committee chair is the only member of the DAC permitted to serve on the examination committee.
- The student may choose to have one faculty member from outside of Harvard, but it is not required.
- The examining committee must have at least one faculty member of the Systems Biology Program.
- All proposed examiners must be the rank of assistant professor or higher.
- The dissertation advisor is not eligible to be an examiner or the chair, but usually attends the exam ex officio.

If an alternate examiner is requested either by the student, dissertation advisor or the program, then the alternate must receive a copy of the dissertation and be available on the date of the defense.

The student is expected to give a seminar of approximately one-hour as part of the examination, on the day of the examination, prior to a defense of the dissertation with the examination committee.
Conference and Professional Development Funding
Students in the third year or beyond, who have been deemed eligible by the Director of Graduate Studies, will receive a one-time professional development fund of up to $2,500 to cover approved expenses.

Events
Systems Biology Student Symposia: December 4, 2018 and April 30, 2019
All Systems Biology students and faculty are invited to attend two half-day symposiums where second year students will give talks on their proposed dissertation project. Each student speaker will be assigned two faculty coaches who will help them to develop their talks and give feedback.

Systems Biology Program Retreat: November 2-3, 2018 (Newport, RI)
All current Systems Biology students and faculty are invited to attend the Systems Biology Program Retreat. The purpose of the retreat is to bring our community together to learn about current research in systems biology. Advanced students will present their research to date orally or through a poster. The schedule for the weekend will also include faculty talks, a poster session, and social gatherings.

Other Seminars of Interest

MCB Events Calendar - https://mcbpublic.unix.fas.harvard.edu/mcb/events/mcb-calendar/ Many talks are listed on the MCB events calendar. Students may be particularly interested in the Thursday Noon Seminars that are held in the Northwest Building

Theory Lunch Talks - https://sysbio.med.harvard.edu/calendar
Theory Lunch is held on Fridays at noon, in Alpert 563 (HMS). A catered lunch is provided and is followed by a "chalk" talk (whiteboard only, no slideware). The talks can be on any biological subject that might be of interest to theoretically minded people, including, but not limited to, on-going work, off-the-wall ideas, a recent paper, an interesting proposal, etc, with vigorous but friendly audience participation.

Pizza Talks - https://sysbio.med.harvard.edu/calendar
SB Tuesday Pizza Talks are a way for Systems Biology Department Post Docs & Graduate Students to present their work and receive feedback in a casual setting. Senior systems biology grad students often volunteer to speak.
Systems Biology Students

Tobyn Branck - G1
tobynbranck@g.harvard.edu
Lab: Debora Marks

Hailey Cambra - G2
hcambra@g.harvard.edu
Lab: Debora Marks

Silvia Canas Duarte - G5
silviacanasduarte@g.harvard.edu
Lab: Johan Paulsson

Adam Carte - G5
ancarte@g.harvard.edu
Lab: Alex Schier

Ivana Cvijovic - G6
icvijovic@fas.harvard.edu
Lab: Michael Desai

Kelsey Dean - G4
kelseydean@g.harvard.edu
Lab: Francis Doyle III
   Allon Klein

David Ding - G5
davidding@g.harvard.edu
Lab: Debora Marks

Daniel Eaton - G2
deaton@g.harvard.edu
Lab: Johan Paulsson

Patrick Flynn - G2
patrick_flynn@g.harvard.edu
Lab: Timothy Mitchison

Samuel Finlayson - G3
samuel_finlayson@hms.harvard.edu Lab: Isaac Kohane

Gabriel Filsinger - G5
filsinger@g.harvard.edu Lab: George Church

Charles Fulco - G6
charlesfulco@fas.harvard.edu Lab: Eric Lander

Anna Green - G6
annaggreen@fas.harvard.edu Lab: Debora Marks

Luis Alberto Gutierrez Lopez - G1
luis_gutierrezlopez@hms.harvard.edu

Gloria Ha - G1
gloriaha@g.harvard.edu

Sabina Haque - G1
sabina_haque@fas.harvard.edu
Systems Biology Students

Rebecca Herbst - G6
rherbst@fas.harvard.edu
Lab: Aviv Regev

Luis Hernandez Nunez - G6
luishernandeznunez@fas.harvard.edu
Lab: Aravinthan Samuel

Allison Hicks - G3
allison_hicks@g.harvard.edu
Lab: Yonatan Grad

Tanush Jagdish - G1
tanush@fas.harvard.edu

David Jacobowitz - G3
djacobowitz@g.harvard.edu
Lab: Sahand Hormoz

Mike Jin - G2
mikejin@g.harvard.edu
Lab: Peng Yin

Gemma Johnson - G3
gjohnson1@g.harvard.edu
Lab: Jessica Lehoczky

Dylan Kotliar - G5
dylan_kotliar@hms.harvard.edu
Lab: Pardis Sabeti

Danylo Lavrentovich - G1
dlavrent@g.harvard.edu

Alan Le Goallec - G4
alanlegoallec@g.harvard.edu
Lab: Chirag Patel

Ang Li - G3
angli01@g.harvard.edu
Lab: Michael Springer

Allen Lin - G5
allenlin@g.harvard.edu
Lab: Alejandro Balazs

Maximillian Marin - G2
maximilliangmarin@gmail.com
Lab: Michael Springer

Hector Medina-Abarca - G6
hector_medina-abarca@hms.harvard.edu
Lab: Walter Fontana

Eran Mick - G7
eranmick@fas.harvard.edu
Lab: Vamsi Mootha

Kee Myoung Nam - G4
kmnam@fas.harvard.edu
Lab: Jeremy Gunawardena
Systems Biology Students

Nathan Nakatsuka - G4
08nanaka@gmail.com
Lab: David Reich

Erica Normandin - G2
ericanormandin@g.harvard.edu

Anjali Nelliat - G1
anelliat@fas.harvard.edu

Jackson Nyman - G2
nyman@g.harvard.edu
Lab: Eli Van Allen

Elizabeth Perez - G2
emperez@g.harvard.edu
Lab: Mario Suva
    Aviv Regev

Natalia Quinones - G2
nquinones@g.harvard.edu
Lab: Michael Baym

Luke Rast - G4
lukerast@g.harvard.edu
Lab: Jan Drugowitsch

Mary Richardson - G1
mrichardson@fas.harvard.edu

Elmer Rho - G2
erho@g.harvard.edu

Tal Scully - G1
Tal_Scully@hms.harvard.edu

Michael Schnebly - G3
schnebly_md@g.harvard.edu
Lab: Daniel Needleman

Jacob Shenker - G3
jqshenker@g.harvard.edu
Lab: Johan Paulsson

Jung-Eun Shin - G3
shinj@g.harvard.edu
Lab: Debora Marks

Sarah Shareef - G4
sshareef@g.harvard.edu
Lab: Bradley Bernstein

Yekaterina Shulgina - G4
shulgina@g.harvard.edu
Lab: Sean Eddy

Daphne Sun - G3
dssun@g.harvard.edu
Lab: Yonatan Grad
Systems Biology Students

Armin Schoech - G5
aschoech@hsph.harvard.edu
Lab: Alkes Price

Narendra Tallapragada - G6
ntallapragada@fas.harvard.edu
Lab: Allon Klein

Arielle Tripp - G1
arielletripp@g.harvard.edu

James Valcourt - G5
valcourt@g.harvard.edu
Lab: Sharad Ramanathan

Adrian Veres - G5
adrianveres@g.harvard.edu
Lab: Doug Melton

Debra Van Egeren - G3
dvanegeren@g.harvard.edu
Lab: Franziska Michor

Elizabeth Van Itallie - G7
vanitallie@fas.harvard.edu
Lab: Marc Kirschner

Roger Vargas - G3
roger_vargas@g.harvard.edu
Lab: Maha Farhat

Samuel Wolock - G5
swolock@g.harvard.edu
Lab: Allon Klein

Suyang Wan - G2
suyang_wan@g.harvard.edu

Caleb Weinreb - G4
caleb_weinreb@hms.harvard.edu
Lab: Allon Klein

Qiu Chang Wu - G1
q_wu@g.harvard.edu

Muhammed Yildiz - G5
myildiz@g.harvard.edu
Lab: Johan Paulsson

Eric Zinn - G4
ericzinn@g.harvard.edu
Lab: Luk Vandenberghe
We develop and apply statistical methodology and algorithms that enable new lines of inquiry in systems biology and integrative genomics. Our focus is on developing mechanistic models that produce predictions testable at the bench. Our approach leads to quantitative insights into molecular and chemical aspects of cellular biology that are not directly measurable with experimental technologies. We are broadly interested in regulatory mechanisms that drive cellular proliferation and growth, the cell cycle, the metabolic cycle, C and N metabolism, nutrient and stress response, protein-mRNA regulation, and cancer systems.

Publications:


Although numerous eukaryotic genomes have been sequenced, much still remains to be understood about how the genes in those genomes are regulated. The interactions between transcription factors (TFs) and their DNA binding sites are an integral part of the regulatory networks within cells. We employ numerous combined experimental and computational strategies in studying these regulatory interactions, to discover cis regulatory codes in the genome.

To permit a genome-wide scan of all possible TF binding sites in a given genome, we developed universal protein binding microarrays (PBMs), a DNA microarray-based technology that allows rapid, high-throughput characterization of the DNA binding site sequence specificities of proteins. We are using PBMs to study hundreds of TFs in Drosophila and mammals, so that we may predict sets of co-regulatory TFs acting together at candidate transcriptional cis regulatory modules, such as transcriptional enhancers. We are examining the effects of protein-protein interactions in DNA binding. We are also investigating divergence within TF families and how different members of a TF family gain distinct regulatory roles. We are investigating the effects of mutations and naturally occurring genetic variation on human TFs. We are combining sequence, evolutionary, structural, and PBM data to understand the molecular determinants of protein-DNA binding specificity and its evolution. We have developed computational strategies that employ comparative genomics methods for prediction of tissue-specific transcriptional enhancers in fly, and separately in mammals. Many of these predicted enhancers recently have been validated in vivo. In our analyses, we infer each TF’s relative importance for a given cell-type-specific gene expression pattern. We are interested in understanding the quantitative nature of information encoded within transcriptional enhancers, as it pertains to spatiotemporal specificity of enhancer activity. Technology development projects include novel approaches for high-throughput, experimental identification of tissue/cell-type-specific enhancers. The results of these experiments and analyses aim for a better understanding of the functions, locations and organization of DNA regulatory elements, and their evolution.

**Publications:**


George Church  
Professor of Genetics

Harvard Medical School  
New Research Building, Room 232 & 238  
77 Ave. Louis Pasteur  
Boston, MA 02115  
Tel: 617-432-7562  
Email: gmc@harvard.edu  
http://arep.med.harvard.edu/gmc/  
Lab Size: Over 20

Our research focuses on new technologies for "omic" measures, synthesis and modeling/CAD/analysis tools. We apply these to biomedical & ecological systems -- in particular, personal genomics and microbial genome engineering for new genetic codes, novel amino acids and multi-virus resistance. We developed next-generation sequencing (NGS) and fluorescent in situ sequencing (FISSEQ) methods to analyze the output of combinatorial selections as well as comprehensive gene-environment-trait data for affordable personalized medicine. We developed CRISPR technologies to reprogram the (epi)genomes of human pluripotent stem cells to connect cis-regulatory motif variants in populations to allele-specific and cell-type-specific RNA measures and further to causal effects on cell and larger-scale morphologies. We helped initiate the BRAIN project and are contributing "innovative neurotechnologies" including fluorescent connectome, transcriptome, ticker tape and developmental lineage barcode collected via a single integrated "Rosetta brain".

Publications:


Diverse control mechanisms converge to ensure that gene transcripts are expressed and processed accurately. Dissection of these interactions has proven challenging, because most experimental approaches record downstream products fed by multiple pathways – for example, mature mRNA as the combined product of transcription and splicing.

The Churchman lab enables direct mechanistic insights into fundamental biological processes by developing and applying quantitative approaches that create high-resolution views of genome function. Our group has developed methods for genome-scale, high-precision measurement of Pol II transcription in yeast and mammalian cells (Churchman and Weissman, Nature 2011; Mayer et al., Cell 2015). These approaches have enabled fundamental insights into many aspects of eukaryotic transcriptional control, such as transcriptional pausing, and they bridge the divide between the wealth of in vitro biophysical studies and in vivo genomics.

Aside from eukaryotic transcription regulation, we are interested in the mechanisms that couple gene expression processes, with two current areas of focus: 1) the coupling of transcription elongation with co-transcriptional processes, particularly splicing; and 2) the coordination of mitochondrial and nuclear gene expression in the assembly of oxidative phosphorylation complexes. By determining the molecular mechanisms that control transcription and couple gene expression processes, we aim to open new vistas on potential therapeutic strategies for correcting the defects in splicing dysfunction and energy production that are increasingly recognized as drivers of disease states.

Publications:

Philippe Cluzel
Professor of Molecular and Cellular Biology and of Applied Physics

Harvard University
NorthWest Building, Room 435.10
52 Oxford St.
Cambridge, MA 02138
Tel: 617-384-9447
Email: cluzel@mcb.harvard.edu
http://www.mcb.harvard.edu/Cluzel/index.php
Lab Size: Between 5 and 10

The Cluzel laboratory draws together a range of scientific disciplines and skills in order to tackle biological problems at the systems level. We use real-time systems analysis to investigate how single cells respond to information in their environment. Our systems of interest include multi-drug resistance in E. coli and S. aureus, transcriptional dynamics of flagellar genes in bacteria, and degeneracy in the genetic code. Our research offers interdisciplinary training opportunities for students with either a biological or physical sciences background. Techniques in the lab come mainly from microbiology, molecular biology, statistical physics, and computational biology; these include fluorescence correlation spectroscopy, microfluidics, bacterial genetics, time-resolved fluorescence, video microscopy, and protein engineering.

Publications:
Adam Ezra Cohen  
*Professor of Chemistry and Chemical Biology*

Harvard University  
Mallikcrdott Building, Room 115  
12 Oxford St.  
Cambridge, MA 02138  
Tel: 496-9466  
Email: cohen@chemistry.harvard.edu

http://cohenweb.rc.fas.harvard.edu/  
Lab Size: Between 10 and 15

My lab develops new physical tools to study molecules and cells, and we apply these tools to make new measurements. We combine nanofabrication, optics, microfluidics, electronics, and biochemistry to generate data; and we apply statistics and physical modeling to understand the data. Current projects include: development of nanofabricated devices for trapping and studying single biomolecules in free solution; development of fluorescent voltage-indicating proteins; studies on optical spectroscopy in highly contorted electromagnetic fields; experiments on magnetically sensitive photochemical reactions; and studies of the mechanochemistry of biological hydrogels.

Publications:


The Denic lab takes advantage of novel methodologies for studying large-scale genetic interactions in budding yeast as well as mass spectrometry-based characterization of natively-isolated protein complexes in order to identify the essential components required for several membrane-associated cellular processes. We then carry out targeted and systematic biochemical reconstitution strategies using the identified components in order to go from parts lists to functional and mechanistic insights.

**Publications:**


Our long-term goal is to understand how regulatory DNA dictates transcriptional network behavior and, ultimately, organismal phenotype and evolution. Our approach is mechanistically motivated: we believe that understanding the molecular mechanisms that drive transcription will lead to models of gene regulation that can predict the functional consequences of regulatory sequence changes and guide production of new types of regulatory circuits.

The strength of our system lies in the breadth and depth of experimental, computational and theoretical approaches that are possible in Drosophila embryos. The nuclei in blastoderm embryos are in a monolayer at the surface, making it easy to image transcription quantitatively in both fixed and live tissue. All the players in the patterning network are known making it ideal for systems-level studies. We can also perturb the network using genetics, transgenesis and genome engineering. Finally, the availability of many genome sequences from related species and natural populations makes insects an outstanding clade for evolutionary studies. Combined, these tools allow us to go “soup to nuts” — from genotype to phenotype — at a level of detail that is impossible in other animal systems. The lessons we learn will be widely applicable to other animal systems including humans, as has historically been the case for research on Drosophila transcription.

**Publications:**


Wunderlich, Z., Bragdon, M.D. & DePace, A.H. (2014). Comparing mRNA levels using in situ hybridization of a target gene and co-
Michael Desai  
Professor of Organismic and Evolutionary Biology and of Physics

Harvard University  
NorthWest Building, 435.20  
52 Oxford Street  
Cambridge, MA 02138  
Tel: 617-496-3613  
Email: mdesai@oeb.harvard.edu  
http://sysbio.harvard.edu/csb/research/desai.html  
Lab Size: Between 5 and 10

Natural selection and other evolutionary forces leave characteristic signatures in the genetic variation within populations. My group uses a combination of theoretical and experimental approaches to study how this genetic variation is created and maintained, and to develop methods to infer the evolutionary history of populations from the variation observed in sequence data. Our focus is primarily on natural selection in asexual populations such as microbes and viruses. We are developing new approaches to population genetic theory to better understand the structure of genetic variation in these populations. We complement this with high-throughput experimental evolution in budding yeast, evolving thousands of lines simultaneously to explore the distributions of phenotypic changes and their correlations with the evolution of genetic variation within and between populations.

Publications:


Our laboratory has identified essential signaling components of the mammalian vomeronasal organ including novel multigene families encoding candidate pheromone receptors. Using molecular and genetic tools we are analyzing the coding of pheromone signals in the mammalian brain and the specificity of the pheromone response leading to gender discrimination and aggression.

**Publications:**


The experimental effort is focused on aging in C. elegans. The phenomenon of aging raises questions about the limits of biological processes. What type of damage, or "garbage", and how much of it, is generated as a by-product of which molecular processes? Is this damage inescapable? What can be repaired and at what cost? Aging is also a life history trait that has been shaped by evolution. This raises the question about the plasticity of aging. Aging research seems to be lagging other areas of molecular biology in adopting a more quantitative and theoretically founded approach.

The challenge of systems biology is not only experimental in kind. It also is the challenge of reasoning about facts that are rapidly evolving while remaining highly fragmented across research communities. I see a fundamental role for models as reasoning instruments in biology. Models, not databases as we know them today, will become the main vehicles for the computer-assisted storage, communication, and retrieval of biological knowledge. Computer scientists and I have joined forces with several other researchers to design a computational environment that represents biological knowledge, as it pertains to signaling, in an editable and executable fashion. This instrument will lend itself to the collaborative construction and critique of models.

**Publications:**


Jeremy Gunawardena  
Associate Professor of Systems Biology

Harvard Medical School, Department of Systems Biology  
Goldenson Building, Room 509  
200 Longwood Ave,  
Boston, MA 02115  
Tel: 617-432-4839  
Email: jeremy_gunawardena@hms.harvard.edu  
http://vcp.med.harvard.edu/index.html  
Lab Size: Between 5 and 10

How do mammalian cells process the information they receive from chemical signals and decide what to do? We take two broad approaches to studying this. From the “inside out”, we study particular molecular mechanisms which are implicated in cellular information processing, such as protein post-translational modification (PTM). Most cellular proteins are post-translationally modified in multiple ways on multiple sites, giving an enormous range of combinatorial molecular states and, with it, the potential capability for sophisticated information “encoding” [4]. We have developed biophysical methods (mass spectrometry and nuclear magnetic resonance spectroscopy) for accurately measuring PTM states, at least for proteins with small (< 10) numbers of sites, [6]. We use this in combination with “systems biochemistry” to try and understand how networks of enzymes regulate the distribution of PTM states and thereby control cellular behaviour.

The combinatorial complexity in PTM is a particular challenge to our conceptual understanding [4]. We have pioneered two mathematical strategies for overcoming this. We have developed a “linear framework” for undertaking time-scale separation, which can analytically simplify extremely complex biochemical networks [5,7,8]. We are now exploiting this framework to study gene regulation in eukaryotic genomes, where the framework allows us to accommodate dissipative, non-equilibrium mechanisms like histone modifications and nucleosome reorganisation. We have also developed the method of “invariants”, in which we exploit the fact that, because of mass-action, a network of (bio)chemical reactions gives rise to a polynomial dynamical system. We can therefore use methods from algebraic geometry to analyse such networks at steady state [2,3,10]. For instance, we can determine (at least in principle) the algebraic invariant that holds among specified components in a network, while eliminating the influence of all the remaining components [3,10]. This distillation can make it much easier to understand what the network is doing. Among other things, these new mathematical developments give us the capability to rigorously infer the properties of certain systems, including PTM systems, irrespective of their underlying molecular complexity [2,8]. We are seeing in this way the emergence of new kinds of mathematics for overcoming biological complexity.

Our second broad approach is from the “outside in”. This arises from a different perspective, which suggests that the complexity found in biochemical networks cannot be fully explained by just studying the networks. Instead, this internal complexity reflects the external complexity of the environments in which these networks were evolved. We have developed new types of programmable microfluidic device to subject cells to complex “interrogation”. We use fluorescence microscopy to assay downstream behaviours at single-cell level and use mathematical models of the molecular networks to predict what we should see. We have developed flexible computational infrastructures (“little b” and, more recently, “Proteus”) to support this kind of modelling, [9]. Instead of studying networks by pulling them to pieces, as in our first approach, we try to ask cells more complicated questions, in the hope that their

Publications:


John Higgins  
*Associate Professor of Systems Biology*

Massachusetts General Hospital  
Center for Systems Biology  
185 Cambridge St  
Boston, MA 02114  
Tel: 617-643-6129  
Email: JMHIGGINS@PARTNERS.ORG  
http://higgins.hms.harvard.edu  
Lab Size: Between 5 and 10

My research focuses on human pathophysiologic processes. I am interested in developing mathematical descriptions of complex human disease phenotypes and how they change over time. Pathophysiology may be described at the molecular, cellular, tissue, and organismal levels and may show clinically significant variation over time scales ranging from seconds to years. My research combines clinical and pathophysiologic insight with dynamical systems theory in order to (1) advance fundamental understanding of the dynamics of human pathophysiology, and (2) improve patient diagnosis, monitoring, and treatments.

Publications:


My research focuses on computational biology at the intersection of microbial community function and human health. The human body carries some four pounds of microbes, primarily in the gut, and understanding their biomolecular functions, their impact on human hosts, and the metabolic and functional roles of microbial communities generally is one of the key areas of study enabled by high-throughput sequencing. First, computational methods are needed to advance functional metagenomics. How can we understand what a microbial community is doing, what small molecule metabolites or signaling mechanisms it's employing, and how its function relates to its organismal composition? Second, our understanding of the human microbiome and its relationship with public health remains limited. Pathogens have been examined by centuries of microbiology and epidemiology, but we know relatively little about the transmission or heritability of the normal commensal microbiota, its carriage of pathogenic functionality, or its interaction with host immunity, environment, and genetics. Finally, more broadly, novel machine learning methodology is needed to leverage structured biological knowledge in high-dimensional genomic data analysis. In other words, we need to turn all of these new data into biology! My group works on a variety of computational methods for data mining in microbial communities, model organisms, pathogens, and the human genome.

Publications:


Organizing Space and Time. In the development of an organism, as in the theater, timing is everything. Imagine if, one night, the actors in a play were to miss every single cue, delivering each line perfectly, but always too early or too late. The evening would be a disaster. The same is true in embryonic development. Starting at the moment when sperm and egg meet, cells in the embryo send signals to each other to coordinate the growth of organs, limbs, and tissues. Not only do the signals have to be correct, they also must be perfectly timed. Otherwise, disasters like cancer can result.

The Kirschner lab studies, among many other things, the way a developing frog embryo orchestrates numerous signals to yield the final, complex organism. Just as multiple cues would destroy an actor's ability to deliver his lines at the right time, it would seem like the existence of multiple signals ought to result in cellular cacophony. But, somehow, the cells in the embryo can sort out the meaning of the different signals that are bombarding them. In particular, the lab is investigating the signals that tell cells when to divide.

Publications:


Our lab studies how individual cells translate internal and external signals into decisions such as growth, death, movement or differentiation. We quantitatively measure the changes in level, activity, or localization of proteins in single cells at high temporal resolution and correlate these behaviors with specific cellular fates. By visualizing how dynamical behaviors vary between different cells, we aim to tease out the reasons for varying behavior both in cell populations and in different cell types. Understanding these issues will be enormously important for understanding how drugs act on different cell types and organs, and to begin to gain insights into the reasons why different cells and people respond differently to specific drugs.

We mainly focus on two networks; the p53 network and the DNA damage response. In the p53 network we are investigating the dynamics of the tumor suppressor protein p53 in response to irradiation and chemotherapeutics drugs in individual cells and ask how p53’s dynamic behavior is controlled, why different cells show different dynamical behaviors and what consequences these behaviors have on cell survival. In the DNA damage response we ask how the kinetics of DNA repair and the choice of repair mechanism are affected by the cell cycle and what are the consequences of activating one repair mechanism versus the other under various cellular backgrounds.

In the long term we are optimistic that these studies will help us predict how signaling networks in human cells will behave in response to new stimuli; how they can be modified or rebuilt to give a desired cellular output; and how to selectively increase the tendency for cancer cells to go in the direction of apoptosis by modulating the dynamics of the networks controlling this decision.

**Publications:**

Giorgio Gaglia, Yinghu Guana, Jagesh V. Shaha, and Galit Lahav. Activation and control of p53 tetramerization in individual living cells PNAS, 2013; PDF.


Jeremy Purvis and Galit Lahav. Encoding and Decoding Cellular Information through Signaling Dynamics Cell, 2013; 152:945-956 PDF.


The Megason Lab has pioneered the use of in toto imaging for revealing the origin of biological form. With the completion of the genome projects, it is now clear that animals have very similar sets of genes. Yet, as is clear to any toddler at a zoo, the diversity of biological form, from the stripes of a zebra to the trunk of an elephant, that can be generated with this same toolbox is astounding. How can the linear sequence of information encoded by the A’s, G’s, C,’s and T’s in the genome be converted into elaborate tissues of the precise size, shape, and organization required for their function?

Biological form is generated during development as the seemingly simple egg elaborates into the morphologically complex embryo. I believe that the mechanistic principles that turn an egg into an embryo span many levels from molecules to cells to whole tissues. Developmental biologists have been quite successful in recent decades in determining molecular level mechanisms, such as the genes governing cell fate decisions, but have made comparatively little progress in other long standing questions in development such as patterning, morphogenesis, and size control. There are two reasons for this failure. One is that these questions cannot be answered using the reductionist logic of one gene, one function. Understanding these phenomena will require a more systems biological approach. The other reason progress has been slow is because the key mechanisms regulating these phenomena do not only occur at the molecular level but also occur at higher levels such as cell dynamics and tissue mechanics.

We have developed a technological approach called in toto imaging which is capable of capturing biological information across all these scales—molecular, cellular, tissue, and whole embryo—in a dynamic fashion. For in toto imaging, we use transgenic zebrafish embryos that are labeled with multiple colors of fluorescent proteins; high-resolution, timelapse confocal/2-photon microscopy is performed to generate image sets containing 100,000 images; and custom developed software called GoFigure is used to track all the cells and quantify molecular, cellular, and tissue level data. We are using in toto imaging in combination with mathematical and computational modeling to elucidate the multiscale principles controlling patterning, morphogenesis, and size control in the zebrafish inner ear and spinal cord.

**Publications:**


The Mitchison group works on fundamental questions of how cells are spatially organized as well as some applied problems in pharmacology and drug development. Our main focus in cell organization is how microtubules self-organize to promote mitosis and cell division. We use frog eggs, cell-free extracts and reconstituted protein systems to investigate how microtubules self-organize to promote mitosis and cytokines. Our main approaches are quantitative microscopy and biochemistry. In the pharmacology area we are working on how microtubule- and mitosis-targeting drugs kill cancer cells, in particular how the important anti-cancer drug taxol works to promote tumor regression in man. We are also developing small molecules that target interferon pathways in innate immune cells with the goal of novel therapeutic approaches to treating and preventing cancer and infectious disease.

Publications:


Our laboratory focuses on mitochondria. These tiny organelles found in virtually all human cells, serving as the center stage for energy metabolism, ion homeostasis, and apoptosis. Their composition, copy number, and efficiency are dynamic properties, varying across cell types and remodeling during growth and differentiation. Mitochondrial dysfunction underlies rare, inborn errors of metabolism, as well as some of the most common human diseases, such as diabetes, cancer, and neurodegeneration. Given their importance in basic biology and clinical medicine, mitochondria represent an excellent "model" for basic and clinical systems biology.

Our group is broadly interested in characterizing the structure and dynamic properties of mitochondria, understanding how genetic variation influence mitochondrial physiology, and exploiting the network properties of the organelle to design therapies for human disease. To achieve these goals, we combine classic biochemistry and physiology with the new tools of genomics, proteomics, and chemical biology. In recent years, we have used mass spectrometry, microscopy, and computation to define the mitochondrial proteome (an inventory we call MitoCarta). We have subsequently coupled MitoCarta with human genetics to discover over one dozen Mendelian disease genes. We have used computational and comparative genomics in combination with CRISPR screens to predict and validate the function of proteins comprising the mitochondrial calcium uniporter.

Current areas of focus include: (1) nuclear:mitochondrial cross-talk in the control of mitochondrial biogenesis, (2) membrane biochemistry and biophysics of ion and metabolite transport, (3) next-gen sequencing and functional studies of human mitochondrial disease, (4) metabolomics approaches to mitochondrial function, and (5) chemical biology approaches to modulating mitochondrial copy number and function.

**Publications:**


Andrew Murray
Professor of Molecular Genetics

Harvard University
NorthWest Building, Room 496.20
52 Oxford Street
Cambridge, MA 02138
Tel: 617-496-1350
Email: amurray@mcb.harvard.edu
http://www.mcb.harvard.edu/murray/index.html
Lab Size: Between 10 and 15

We try to understand the “rules of the game” that explain how cells function and evolve. We study budding yeast, using experimental evolution, genetic analysis, synthetic biology, and cell biology. We try to make quantitative measurements that discriminate amongst different classes of models. Members of the lab come from both biology and physics backgrounds.

How does biological novelty evolve? Because we lack time travel, this process is difficult to study in nature, and we therefore apply selective pressure in the laboratory. We have evolved multicellularity, altered mating preferences, circadian oscillators, genetic instability, and new connections between signaling pathways and have developed methods to find the mutations that cause these new phenotypes. We are interested both in general questions about what determines evolutionary trajectories and the specific mechanisms that organisms invent to produce novel traits.

How do cells accomplish specific tasks and how did these solutions evolve? We follow the Feynman principle of “What I cannot create, I cannot understand” by engineering and analyzing the behavior of new yeast strains. As examples, we have used synthetic biology to support the notions that the efficient use of secreted public goods drove the evolution of multicellularity, that multicellularity arose before cellular differentiation, and that novel symbioses could arise without requiring previous evolutionary co-adaptation.

How do cells respond and adapt to their environment to maximize the chance that they survive and reproduce? Achieving these aims requires the coordination of thousands of reactions under a wide range of inter- and extracellular conditions. We are exploring how yeast cells respond to sudden starvation and have discovered that they can rapidly halt their cell cycles, at any stage, and then, later, slowly resume cell division. We are asking how they arrest, whether the arrest destabilizes the genome, and how cells adapt to start dividing again.

Finally, we collaborate with David Nelson (Physics) to combine theory and experiment to investigate population dynamics and evolution in space and time.

Publications:


The Needleman laboratory combines quantitative experiments and theory to study the architecture and dynamics of self-organizing subcellular structures. We presently focus on trying to understand the cell division and spindle assembly and function. Research in the group can be divided into three broad categories: spindle biophysics, spindle physics, and spindle evolutionary biology. This work is highly interdisciplinary and combines approaches from engineering, physics, and biology.

Spindle Biophysics: Spindles are composed of microtubules and associated proteins. It is unclear how the activity of microtubules gives rise to the spindle, largely because we do not know how microtubules act in spindles. We are using quantitative measurements, theory, and computer simulation to determine how microtubules behave in spindles in human cells. This project entails extensive technique development. The ultimate goal of this work is to create mesoscopic models that can predict the response of the spindle to disease state and novel therapeutic drugs.

Spindle Physics: Spindles can be composed of hundreds of thousands of microtubules, whose collective interactions are ultimately responsible for biological behaviors. We are testing field theoretic models of spindles in Xenopus egg extracts to gain insight into these collective effects. We are exploring these spindle’s structure, dynamics, and response to perturbations. Our approach is inspired by methods successfully used in soft condensed matter physics.

Spindle Evolutionary Biology: Spindles in different organisms and different tissues within organisms are different, but very little is known about the mechanistic reasons for these differences or the evolutionary forces ultimately responsible for them. We are address these issues using biophysical, quantitative genetic, and comparative approaches on spindles in early nematode embryos.

Publications:
Reza Farhadifar, Charles F. Baer, Aurore-Cecile Valfort, Erik C. Andersen, Thomas Muller-Reichert, Marie Delattre, Daniel J. Needleman, Scaling, Selection, and Evolutionary Dynamics of the Mitotic Spindle. Current Biology, 2015 PDF


The Program for Evolutionary Dynamics is an interdisciplinary group, bringing together scientists from different disciplines to study theoretical and experimental aspects of evolution. Under the direction of Martin Nowak, Professor of Mathematics and of Biology, researchers at the Program for Evolutionary Dynamics use mathematical approaches and computer simulations to study the evolution of cooperation (evolutionary game theory), the evolution of social behavior, the somatic evolution of cancer, the dynamics of virus evolution, the evolution of language, and the origins of life. Experimental methods from behavioral economics are also used to explore human interactions, bridging theory and experimental observations.

Evolution is the one theory that permeates all of biology. The evolutionary biologist Theodosius Dobzhanski once remarked: "Nothing in biology makes sense if not seen in the light of evolution." We add: "Nothing in evolution makes sense if not seen in the light of evolutionary dynamics."

Evolutionary dynamics is the study of the fundamental principles of evolutionary change. Evolution requires populations of reproducing individuals. Mutation generates novelty. Natural selection and random drift determine the fate of new mutants. Cooperation leads to organization.

The natural laws of evolution can be formulated in terms of precise mathematical equations. These equations describe reproduction, mutation, selection and cooperation. Reproduction can be genetic or cultural. Therefore, evolutionary dynamics describe in principle both genetic and cultural evolution - although there are important differences. These methods provide a powerful set of tools for investigating issues across biology and the social sciences.

Publications:


Life in single cells is dictated by chance: reactions that involve small numbers of molecules generate spontaneous fluctuations that then enslave all dependent processes. Our goal is to identify and understand the guiding principles behind such phenomena, by developing mathematical methods to interpret fluctuations, experimental methods to count molecules in single cells, and combining the two to study the simplest natural and engineered networks, primarily in bacteria and yeast cells. The mathematical methods focus more on general theorems for fluctuations than individual models, drawing on probability theory and statistical physics, as well as control and information theory. Our efforts particularly focus on analytical methods to interpret fluctuations in terms of physical observables rather than kinetic parameters. The experimental methods focus on various tricks from genetics and biochemistry to count the integer number of macromolecules in cells, combined with techniques from microfluidics and single molecule detection under the microscope. The applications range from stress response, replication control, partitioning of molecules at cell division, horizontal gene transfer, bacterial epigenetics, toxin-antitoxin systems, stochastic proteolysis, and the evolution of cooperation.

Publications:


Our laboratory studies how cells and organisms make decisions. To arrive at a decision, organisms must measure multiple environmental signals and interpret them appropriately. The questions we are interested in are how cells and organisms interpret their environment, how this interpretation depends on prior experiences, as well as the spatial, temporal modulation and the statistics of environmental cues. We want to achieve a quantitative understanding of the underlying signaling and transcriptional circuits that lead to discrete decisions. Since our goal is to uncover general design principles of the circuits that underlie decision making, our lab works on several model systems. Our recent work has focused on the yeast Sacchromyces cerevisiae, and we are now working on circuits that make developmental decisions in mammalian cells and behavioral decisions in the worm C. elegans. We are also developing several new optical and micro-fluidic techniques to interrogate the dynamics of signaling and transcriptional networks in single cells.

Publications:


Molecular circuits are the information processing devices of cells and organisms, transforming extra- and intra-cellular signals into coherent cellular responses. Past studies to chart key circuits in mechanistic and functional detail have typically required decades of serial work to identify and connect a few components or interactions at a time. In recent years, there has been hope that genomic approaches would make it possible to reconstruct circuitry on a systems level. However, genomic studies have largely been observational and rarely involve large-scale testing of the models and subsequent refinement.

We address this challenge with a systematic computational and experimental approach, based on iterating three steps. First, we measure the circuit’s output (e.g. mRNA levels) or internal state (e.g. protein-DNA or protein modification states) along a relevant time course using genomic tools. Next, we create a computational model that explains the observed data, using algorithms we develop. We perturb every key component proposed by our model (e.g. using RNAi). Then, we repeat the process by measuring the circuit’s output or internal state following the perturbation, refining the model, and testing it again, until data and model converge.

All circuits are dynamic, and rewire in response to perturbation, at time scales from minutes to eons, as cells respond to new environmental conditions, differentiate, or evolve. We focus on a selected model system at each time scale. For short-term responses, we study the regulatory circuit of primary mouse dendritic cells (DCs) responding to pathogen components. For long-term responses, we study the differentiation of immune cells, especially T helper cells. For evolutionary changes, we study the rewiring of nutrient responses in 15 yeast species.

We develop and apply an extensive experimental and computational toolbox for each step. Experimentally, this includes novel methods to profile RNA, proteins and their interactions, nanotechnology-based delivery to primary cells, and mesoscale ‘signatures’ to monitor the effect of hundreds of perturbations. Computationally, we have pioneered sophisticated algorithms to reconstruct dynamic circuit models from time course and perturbation data, to design time course, perturbation and signature experiments, and to facilitate analysis of large-scale genomics datasets, especially RNA-Seq.

**Publications:**


The goals of The Sabeti Lab are to use computational methods and genomics to understand mechanisms of evolutionary adaptation in humans and pathogens. We are pursuing these goals through 3 research foci:

(1) Developing analytical methods to detect and investigate evolution in the genomes of humans and other species
(2) Examining host and viral genetic factors driving disease susceptibility to the devastating and deadly disease widespread in West Africa, Lassa hemorrhagic fever virus.
(3) Investigating the genomes of microbes, including Lassa virus, Ebola virus, Plasmodium falciparum malaria, Vibrio cholerae, and Mycobacterium tuberculosis to help in the development of intervention strategies.

Publications:


We seek to both enhance our understanding of natural biological design, and to develop tools and concepts for designing cells, tissues and organisms. In the long term, we hope to develop principles for building synthetic cells that act as sensors, memory devices, bio-computers, producers of high value commodities and energy from the sun, and to build novel subsystems such as proteins with designed properties for therapeutic use. Current projects use mammalian cells, simple eukaryotes and prokaryotes. Understanding how to program cells in a rational way will have value, for example, in stem cell design, drug therapy and the environment. These experiments use a combination of theoretical and experimental approaches that are well suited to students with backgrounds in biology, engineering, or any allied field.

**Publications:**


We explore design principles for self-assembling molecular machines, primarily using structural DNA nanotechnology to build our model systems. We seek to apply our knowledge towards construction of artificial systems that help solve problems of biological and medical interest. Currently we focus on building tools for molecular biophysics (NMR structure determination of membrane proteins, single-molecule assays) and for therapeutics (drug delivery, tissue engineering). For many applications, our present capabilities are too primitive. Thus we also investigate enabling technologies for increasing the complexity of programmable self-assembled systems.

Publications:


The Sorger Laboratory studies mammalian cancer biology with a focus on cell signaling networks involved in disease and the therapeutic drugs that target them. We combine cell biological, biochemical and mathematical approaches with the overall aim of deriving novel mechanistic and systems-wide insight into cellular physiology and human disease. Modeling methods range from purely statistical and correlative to physicochemical and mechanistic but in all cases we rigorously link models to experimental data on protein states, signaling biochemistry, and cellular phenotypes.

Our laboratory is particularly interested in death and survival signaling in response to death ligands such as TRAIL and TNF and growth factors such as EGF, IGF and HGF. We study how these ligands affect normal cells and how the signals they provoke are misregulated in cancer. A significant fraction of the lab’s effort is devoted to developing and applying new mathematical methods for model assembly, calibration and validation (including logic-based modeling, rules-based physicochemical modeling, Bayesian calibration and optimal experimental design). We also develop and exploit mouse “models” of cancer that recapitulate key features of human disease (particularly breast cancer and hepatocellular carcinoma). Finally, we develop and apply a wide variety of live-cell microscopy methods as well as multiplex approaches to biochemical measurement. Individual lab members are encouraged to undertake projects that combine computation and experimentation.

Publications:


While we have a great mechanistic understanding of how the information encoded in DNA is converted through mRNA to protein, we have a limited understanding of the quantitative relationship between genotype and phenotype. We do not know whether properties such as protein abundance are under strong selective constraints and therefore it is difficult to interpret the myriad of changes that are evident between related individuals and species. Using a combination of theoretical, genomic, and proteomic approaches we are exploring in both a high throughput and direct fashion how evolution has changed (comparative evolution), could have changed (synthetic evolution), and does changes (experimental evolution) quantitative features of networks in multiple yeast species under different selective regimes.

Publications:


Ralph Weissleder
Professor of Systems Biology and of Radiology

Center for Systems Biology, Massachusetts General Hospital
Simches Research Building, Room 5.210
185 Cambridge St.
Boston, MA 02114
Tel: 617-726-8226
Email: rweissleder@mgh.harvard.edu
http://csb.mgh.harvard.edu/
Lab Size: Over 20

Current research in the Weissleder lab is focused on developing better detection methods for specific cells (e.g. circulating cancer cells, immune cells), cell-derived microvesicles or, bacteria implicated in human disease. The lab utilizes a variety of techniques including whole body and intravital microscopic imaging, novel chemical approaches that perturb specific systems or biological pathways, and innovative sensing strategies that include nanotechnology and microfluidic methods. The lab uses these techniques to: a) obtain quantitative and systems-wide global measurements, b) perform dynamic serial measurements, and c) integrate multiple and various data sets into models. Increasingly, the lab has been focused on reconciling the gap that exists between imaging and traditional cell biology research by developing techniques that are powerful enough to image cellular processes in an in vivo setting. Thus far, the lab’s work on nanomaterials and on novel miniaturized nuclear magnetic resonance (μNMR) chips have led to advanced clinical trials. Single cell analysis systems have also been developed for rapid and accurate diagnosis and monitoring of drug therapy. Whilst the research is primarily basic in nature, much of it has a strong translational focus, with in vivo imaging playing a major role.

Publications:


Peng Yin  
*Professor of Systems Biology*

Wyss Institute for Biologically Inspired Engineering  
Center for Life Sciences, 5th Floor  
3 Blackfan Circle  
Boston, MA 02115  
Tel: 617-432-7731  
Email: py@hms.harvard.edu  
http://molecular-systems.net/  
Lab Size: Between 10 and 15

Our research interest lies at the interface of information science, molecular engineering, and biology. We are generally interested in developing programmable molecular systems and technology inspired by biology. Specifically, we focus on engineering information directed self-assembly of nucleic acid (DNA/RNA) structures and devices, and on exploiting such systems to do useful molecular work, such as probing and programming biological processes for imaging and therapeutic applications.

**Publications:**


The Zhuang research lab develops and applies advanced optical imaging techniques to study the behavior of individual biological molecules and complexes in vitro and in live cells. Our understanding of living organisms has greatly benefited from various imaging and visualization tools. In particular, understanding the inner workings of a cell requires imaging techniques with molecular-scale resolution and dynamic imaging capability such that molecular interactions and processes inside the cell can be directly visualized. We are developing imaging methods with single-molecule sensitivity and nanometer-scale resolution to meet these challenges and applying these tools to study problems in cell biology, neurobiology and microbiology. Our current research effort focuses on three main directions: (1) developing super-resolution fluorescence microscopy techniques to allow imaging of cells and tissues with molecular-scale resolution. (2) Apply super-resolution imaging methods to study neural circuitry and sub-neuronal structures, DNA organization and gene expression regulation, as well as how viruses infect cells. (3) using single-molecule approaches to investigate how proteins and nucleic acids interact, with emphasis on chromatin remodeling.

Publications:
David Nelson  
Professor of Biophysics and of Physics and Applied Physics

Harvard University  
19 Oxford St., Lyman 234  
Cambridge, MA 02138  
Tel: 617-495-4331  
Email: nelson@physics.harvard.edu  
http://www.physics.harvard.edu/people/facpages/nelson.html  
Lab Size: Between 5 and 10

Much of Nelson's recent research has focused on problems that bridge the gap between the physical and biological sciences. Early work explored population dynamics in the presence of varying growth rates and convection, leading to universal predictions for the spreading and transverse profile of populations in space- and time-dependent environments. Together with his student, David Lubensky, Nelson developed a theory of force-induced denaturation of double-stranded DNA. Sequence heterogeneity dominates the dynamics of the un-zipping fork (with possible implications for DNA replication in prokaryotes) over a large of forces above an unzipping transition. Energy barriers near the transition scale as the square root of the genome size. Recent observations of jumps and plateaus in the unzipping of lambda phage DNA at constant force (the result of a collaboration with his colleague, Mara Prentiss) are consistent with these predictions. Sequence heterogeneity was also predicted to have a re-markable effect on the dynamics of motor proteins such as helicases, exonulceases and RNA polymerases, leading to sublinear drift in time of these complex enzymes and a possible expla-nation for the nearly horizontal velocity-force curves observed near the stall force in experi-ments on RNA polymerase. Nelson and colleagues have also studied the shapes of viruses, showing that the icosahedral packing of protein capsomeres of spherical viruses becomes un-stable to faceting for sufficiently large virus size. A parametization of the architecture of virus shells in terms of single dimensionless "von Karman number" shows why small viruses are round and large ones are faceted, and allows important information about the elastic constants to be extracted from electron micrographs. Professor Nelson's current interests include population genetics, gene surfing, cooperation and competition

Publications:


We study the structure, organization, and dynamics within prokaryotic cells. By developing a detailed spatial, mechanistic understanding of essential processes within bacteria we hope to uncover new targets for antibiotics. Our approach is based on the historic methodology established in field of eukaryotic cell biology: the biochemical characterization of protein function in vitro and the microscopic examination of dynamics in vivo. We use biochemistry, genetics, high precision particle tracking, super-resolution imaging, and chemical genetics to build spatial, mechanistic understandings of molecular machines.

Bacteria, even though they are quite small, contain a great deal of spatial order, order they to encode developmental functions. As these processes are encoded with a small set of genes, bacteria provide model systems to directly visualize, dissect, and understand how biological machines work.

We are currently studying how bacteria establish the most fundamental step of development, asymmetry of shape. A rod is the simplest developmental shape that can occur from the 1D coordinate system of a from a sphere: By defining a second axis of length, the cell gains a new coordinate system by which cellular components can be organized.

In rod shaped bacteria, these two coordinates of width and length correspond to two modes of PG synthesis: during elongation cell wall material is inserted evenly throughout the length of the rod (setting the width), and during division the synthesis is restricted to occur at the center of the cell (reading out length). Each of these systems is coordinated by a different cytoskeletal polymer. Our work has shown that these enzyme / filaments complexes move circumferentially around the cell width, with complexes moving in both directions. Surprisingly, this motion appears to be powered by the process of cell wall synthesis itself. We are now working to understand how the local actions of these independently moving enzymes are able to impart long range order and give reproducible cell shape. We are also watching the enzymatic motions of division, and are attempting to build a “polymer-up” mechanistic model of how both of these systems function, and how the cell controls these systems in response to growth rate.

**Publications:**


Animal genomes endow cells and circuits with the ability to form life-long memories. How does a genome orchestrate this dynamic interaction of neurons with experience? The genome responds to experience by unleashing bursts of new gene expression that rewire circuits to store long-term memories. These bursts of gene expression rely on an exceedingly complicated network of neuronal activity-regulated transcription factors. It is neither known why such an extensive network is necessary nor how it works. We are taking a two-pronged approach to understand how this transcriptional network rewires neuronal circuits. First, we are applying recently developed genomics and systems biology approaches to understand how the activity-regulated transcriptional network responds to increases in neuronal firing rates. At the same time, we are establishing brain slice and in vivo experimental systems in which we can manipulate the transcriptional network in an intact circuit, using electrophysiological, optogenetic, and behavioral tools to assess how these manipulations affect neuronal circuit rewiring.

**Publications:**


Ramy Arnaout  
Assistant Professor of Pathology

My laboratory has two complementary areas of expertise: (1) systems immunology and (2) systems medicine. They have in common the use of big-data analytical methods, high-performance computation, information theory, and network/graph theory harnessed to understanding complex systems.

Systems immunology. B and T cells are important in vaccines, infections, autoimmunity, aging, and cancer. There are millions of these cells in a typical blood sample. High-throughput sequencing and big-data computational biology make it possible to investigate how this unique set of cells carries out its functions—and raise the possibility of using them as early diagnostics and biological therapeutics for a range of conditions. In our lab, we are sequencing and analyzing antibodies and T cell receptors across species and medical conditions to understand this complex system.

Systems medicine. Medicine has a genome. This genome is made up of the health data of the millions of people who interact with doctors, nurses, and others at hospitals and clinics around the world. This includes the results of diagnostic blood tests, the single highest volume medical activity, as well as procedures, diagnoses, and more. Within this genome is knowledge for better patient care and the basis for a nationwide learning health system. We are using big-data analytics and traditional techniques to reveal useful patterns, trends, and features of conventional medical care and of genomic medicine and use real-world hospital data to give structure to problems in clinical medicine that are commonly talked about but not well understood.

Publications:


Dr. Klein is fascinated by the question of how stem cells choose between alternative fates in developing and adult tissues. Today, one can strip differentiated cells of their fate by inducing pluripotency, yet it remains exceedingly difficult to target differentiating stem cells to a specific fate. His work focuses on germ layer specification in the South African claw-toed frog, *Xenopus Laevis*, which is a powerful system for studying cell fate choice and differentiation in a "native" biological context (rather than in cell culture). He is currently studying collective fluctuations of groups of genes, measured at the single-cell level, to test several possible hypotheses for how different fates are selected. Dr. Klein's research combines (1) traditional experimental methods used to study developmental perturbations, (2) novel assays for gene expression in hundreds of individual cells, (3) statistical tools for analysing gene expression fluctuations, and (4) theoretical tools for characterizing stochastic cell fate decisions.

**Publications:**


Our lab uses methods and approaches of physics and chemistry to gain fundamental first-principle insights into evolutionary dynamics. It requires vantage points on multiple scales, ranging from the molecular, through the systems, to the cellular/organismal scale. To that end we employ a variety of approaches from computational and theoretical multi-scale modeling to experimental approaches to derive genotype phenotype-relations. Traditional experimental approaches are “top-down” in which phenotype is selected first, thus precluding the pinpoint location of the molecular source of perturbation due to possible accompanying genetic variation elsewhere. In departure from tradition our studies are “bottom up” whereby we introduce, through genomic editing, chromosomal mutations into an essential gene(s), whose effects on protein Biophysical and Biochemical properties have been carefully evaluated in vitro and in silico. The genome editing technique employed by our lab allows us to make precise and Biophysically/chemically-rational genetic perturbation covering a broad range of molecular effects of mutations. A precise control of perturbations at the molecular level enables us to quantitatively track the linkage between genetic variation and phenotype on all scales, including the systems-level and population-level and pinpoint the underlying physical-chemical causes of fitness effects of mutations. This approach also allows us to get crucial insights into evolutionary dynamics of antibiotic resistance; mechanisms of horizontal gene transfer and uncover physical chemical causes of various bacterial phenotypes. In a related development we study viral evolution using computational methods and experimental approaches based on microfluidics. The emerging theory aims to predict incoming strains that escape current immune responses providing a path to proactive vaccine and anti-viral drug development. Other active projects in the lab include rational ligand (drug) design, simulations of protein folding and aggregation at the atomic level and theoretical studies of population genetics in the context of Molecular Biophysics.

Publications:
Debora Marks
Assistant Professor of Systems Biology

Harvard Medical School
200 Longwood Ave
Armenise 630
Boston, MA 02115
Tel: 617-432-5016
Email: debbie@hms.harvard.edu
https://marks.hms.harvard.edu/
Lab Size: Between 5 and 10

One million human genomes, will it make a difference?

The large and growing volume of genome information, from all forms of life, presents unprecedented opportunities for computational biologists. The challenge for our scientific generation is to turn an avalanche of sequence information into meaningful discovery of biological principles, predictive methods, or strategies for molecular manipulation for therapeutic and biofuel discovery. The Marks lab is a new interdisciplinary lab dedicated to developing rigorous computational approaches to critical challenges in biomedical research, particularly on the interpretation of genetic variation and its impact on basic science and clinical medicine. To address this we develop algorithmic approaches to biological data aimed at teasing out causality from correlative observations, an approach that has been surprisingly successful to date on notoriously hard problems. In particular, we developed methods adapted from statistical physics and graphical modeling to disentangle true contacts from observed evolutionary correlations of residues in protein sequences. Remarkably, these evolutionary couplings, identified from sequence alone, supplied enough information to fold a protein sequence into 3D. The software and methods we developed is available to the biological community on a public server that is quick and easy for non-experts to use. In this evolutionary approach to accurately we have predicted the 3D structure of hundreds of proteins and large pharmaceutically relevant membrane proteins. Many of these were previously of unknown structure and had no homology to known sequences; two of the large membrane proteins have now been experimentally validated. We have now applied this approach genome wide to determine the 3D structure of all protein interactions that have sufficient sequences and can demonstrate the evolutionary signature of alternative conformations.

The vision for the Marks lab is to build computational methods that address three critical challenges (i) protein conformational plasticity in health and disease, (ii) genome-wide evaluation of mutations on disease likelihood, antibiotic resistance and personal drug response, and (iii) synthetic protein design.

Publications:
Sequence co-evolution gives 3D contacts and structures of protein complexes

Protein structure prediction from sequence variation. D. S. Marks, T. A. Hopf, C. Sander,. Nature biotechnology 30, 1072-1080 (2012);

Three-dimensional structures of membrane proteins from genomic sequencing

Protein 3D structure computed from evolutionary sequence variation.

RNAs globally perturbs gene regulation by endogenous microRNAs

We are interested in deciphering the evolutionary history of life by comparative analysis of genome sequences. We develop computational algorithms and software tools for genome sequence analysis, including the HMMER and Infernal software packages for identifying distant homologs of biological sequence families. We rely on Bayesian probabilistic inference approaches, including hidden Markov models (stochastic regular grammars) for primary structure analysis of proteins and DNA, and stochastic context-free grammars for RNA secondary structure and sequence analysis.

We have a special interest in RNA. The "ancient RNA world" hypothesis says that an ecosphere of RNA-based life preceded protein/DNA based life. Some RNA genes that we see today are thought to be ancient relics of the RNA world. By studying modern RNAs, we may learn something about the origins of life. But in addition, we and others have also been finding that the RNA World model is pessimistic in a sense: far from being a few scattered relics, RNAs are in widespread use in modern organisms in a variety of roles. We have argued for a "modern RNA world" hypothesis, that many of the RNAs we see today are modern inventions, highly adapted to regulatory roles in complex organisms.

Most recently, our laboratory has begun looking for ways to study one of the great mysteries in biology: how a relatively small genome manages to specify biological complexity, especially something as complex as what we see in the neural circuits of a brain. This is an area where molecular evolutionary biology still largely lacks quantitative language for asking precise questions. As one way forward in this difficult frontier, we collaborate with neuroscientists working on the molecular regulatory specification of neural cell types in fly, worm, and mouse.

Publications:


Harvard School of Public Health
665 Huntington Ave
Building 1, Room 715
Boston, MA 02115
Tel: 617-432-2275
Email: ygrad@hsph.harvard.edu
http://www.hsph.harvard.edu/yonatan-grad/
Lab Size: Between 5 and 10

The Grad laboratory studies how pathogens evolve and spread through populations with the motivation of improving clinical and public health strategies for decreasing the burden of disease. To date, our efforts have focused on several pathogens, including Neisseria gonorrhoeae, respiratory syncytial virus, and E. coli. We aim to use a variety of methods, including genomics, epidemiological tools, and microbiology to define the dynamics of spread and characterize the genotypic and phenotypic diversity of the pathogens.

Publications:


The Impact of Human Evolutionary History on Biology and Disease

(Focus #1) Using the historical perspective to improve human health: A major focus of our lab has been to tailor methods for finding disease genes to populations of recently mixed ancestry, like African Americans and Latinos (1, 2, 3). For example, we found seven independent genetic risk factors for prostate cancer that are sufficient to explain the elevated rate of this disease in African Americans (4). We also developed EIGENSTRAT, a widely used method to correct for population stratification: systematic differences in allele frequency between cases and controls due to differences in their history that can cause false-positive associations to disease (5). A current project is motivated by the fact that thousands of groups in India have likely experienced founder events as strong as those that have occurred in Ashkenazi Jews or Finns (6). This suggests that there are likely to be many rare recessive diseases in India. Mapping and understanding these is likely to provide a major opportunity for public health improvement as well as a natural laboratory in which to study genetic disease.

(Focus #2) A new history and geography of human genes informed by ancient DNA: We have developed methods for using data from modern and ancient DNA to learn about population structure and mixture events (6, 7, 8). We have used these methods to reconstruct the deep population history of Indians (6), Europeans (7), and Native Americans (9). Using ancient DNA in collaboration with Svante Pääbo’s group, we have also found evidence for gene flow from Neandertals into the ancestors of non-Africans (10) and from ancient “Denisovans” into the ancestors of Melanesians (11; 12; 13). We completed a new ancient DNA lab in February 2013, and are using it to study DNA samples dating from the last ice age to further enrich our understanding of history.

(Focus #3) How did population mixture affect human biology? The finding that many human populations descend from population mixture events is likely to be biologically important. For example, at the time that Neandertals met and mated with modern humans, they were pre-adapted to Eurasian environments, while modern humans were likely not. Thus, modern humans could have used genetic mutations inherited from Neandertals to more quickly adapt to the new and challenging environments they were encountering outside of Africa. We are working to understand the biological impact of Neandertal mixture and other mixture events in human history.

Publications:
The Doyle group is a leading force on the computational side of the field of systems biology. Our ongoing work on circadian rhythms continues to probe at the sources of regulation that give rise to highly precise periods in the mammalian ‘biological clock’. We have also begun to make preliminary links between clock performance and cognitive function. One of the most profound contributions from our group is the development of a cell autonomous mathematical model of the circadian oscillator, which resolves the experimentally observed discrepancies between the tissue (and animal) scale and the cellular scale. Our ongoing work involves elucidating the mechanisms and networks driving coherence in populations of cellular circadian oscillators. Research efforts in systems biology have expanded into the application domains of ecology, with the broad technical theme of understanding synchronized population-scale phenomena, such as coral spawning, using coupled and driven oscillator models. Our medical systems biology studies have expanded to include Diabetes, Alzheimers Disease, Heat Stroke, and PTSD.

Over the past two decades, our group has been working toward the development of automated control of insulin delivery to regulate glucose in people with type 1 diabetes. This has resulted in a functional artificial pancreas that has been evaluated successfully in numerous clinical studies. Our work has been greatly enriched by our collaborations with the William Sansum Diabetes Research Center in Santa Barbara, CA, along with other clinical research facilities around the world. In the past year, we have worked to achieve regulatory approval to conduct the first fully outpatient evaluation of our control algorithm. Our software system, the portable Artificial Pancreas System (pAPS), has been used by 10 clinical-sites around the globe for clinical trials. Our technical contributions to artificial pancreas research include: a method for hypoglycemia alarming, a zone model predictive control strategy, a personal model predictive control algorithm, a safety mechanism to limit insulin overdosing (insulin on board), monitoring and telemedicine, and schemes for improved day-to-day management of insulin dosing (iterative learning control). We have also conducted pilot studies of innovative new approaches such as the use of intraperitoneal insulin delivery for the artificial pancreas and the use of inhaled insulin to supplement closed-loop glucose control. Our collaborative relationships with leading diabetes technology companies have allowed us to move quickly from bench to bedside to evaluate our ideas in real-life scenarios.

Publications:


The complexity and heterogeneity of biological systems are formidable obstacles that must be overcome for achieving a more quantitative and predictive understanding of physiology and phenotypes on the cellular or organism scale. Such a level of understanding has remained largely elusive in biology, despite the extraordinary level of detail to which molecular interactions have been characterized over the past decades, as it often remains unclear how to harness detailed molecular knowledge to achieve this goal.

In our lab, we try to tackle these challenges by identifying phenotypic patterns that can guide us in decoding the underlying molecular mechanisms and principles, which govern the behavior of complex biological systems. Our approach relies on the close coordination and mutual feedback between experimental and theoretical efforts and we combine careful characterization of physiology, genetic perturbations, omics technology and theoretical models.

Fundamental biological questions that we are interested in include the role of metabolic strategies during growth and adaptation, tradeoffs between competing evolutionary objectives of microorganisms and how cells achieve homeostasis of cell size, cell number and cellular composition, as well as the breakdown of these mechanisms in disease. We use the well-characterized model organisms Escherichia coli and Drosophila melanogaster to address such questions.

Publications:


**Doeke Hekstra**  
*Assistant Professor of Molecular and Cellular Biology*

Northwest Building 311.11  
52 Oxford Street  
Cambridge, MA 02138  
Tel: 617-496-4740  
Email: doeke_hekstra@harvard.edu  
https://www.mcb.harvard.edu/directory/doeke-hekstra/  
Lab Size:

---

**Publications:**


During development, for tissue maintenance, and in diseases, as cells proliferate, they transition between functionally and molecularly distinct states. Dysregulation of transitions between cell states can lead to pathologies such as cancer. In many biological systems, we would like to know which transitions can occur, in what sequence, and at what rates. Current single-cell methods have expanded our ability to identify distinct cell states but not our ability to directly measure their dynamics. For example, state of the art single-cell techniques can measure the expression levels of thousands of genes, but they destroy the cells in the process, and provide only static snapshots.

In our lab, we combine experiments and theory to understand the complex dynamics of cell state transitions in development and in disease. To do so, we expand and combine emerging experimental techniques. For example, we use time-lapse microscopy to track the lineage history of individual cells as they divide, followed by single-molecule imaging to readout the expression levels of multiple genes in the same cells. We also use synthetic biology to engineer novel genetic circuits that can retain the history of major transitions in each cell. One of our favorite model systems is mouse embryonic stem (ES) cells in culture, which can rehash slices of development on a dish, providing us with a rich playground.

At the same time, we also develop new theoretical frameworks to decipher the data generated in our experiments: for example, methods for inferring dynamics from single-cell lineage trees with endpoint gene expression measurements. Often inspired by concepts in statistical physics, we have used ideas from geometry, critical phenomena, and at times number theory. Ultimately, we seek to find general theoretical principles to understand the stochastic processes and collective dynamics that emerge in populations of proliferating cells.

**Publications:**


Accuracy is a remarkable feature of a large number of biological processes. To maintain homeostasis, cellular protein degradation must be carried out with no less efficiency or precision than transcription and translation. My lab is trying to understand how the cell achieves accurate control of protein degradation and how its failure may lead to neurodegeneration and aging. Using modern enzymological techniques, namely single-molecule fluorescence/force spectroscopy and cryo-electron microscopy, we are investigating how the information in ubiquitin configurations is decoded by the proteasome, a universal protein machine in eukaryotic species, to command an accurate rate of protein degradation.

Publications:
Current projects include the following:

(1) Which genetic and epigenetic alterations cause cancer?

Cancer cells often harbor hundreds to thousands of genetic changes. Many of those changes represent neutral variation that does not influence cancer development; such mutations are called passenger or hitchhiking mutations. A few alterations, however, are essential for driving tumorigenesis. Those changes are known as driver mutations and increase the reproductive fitness of the cancer cell. The identification of such mutations is of crucial importance for drug discovery because they represent promising targets for therapeutic intervention. We are developing a novel evolutionary mathematical framework designed to identify alterations that act as drivers during tumorigenesis. This method may aid in the prioritization of candidate mutations for functional validation and contribute to the process of drug discovery.

(2) In which order do oncogenic alterations arise?

Recent technological advances have empowered researchers to examine the cancer genome at unprecedented throughput and resolution. Computational algorithms designed to filter random genetic events have begun to uncover mutational patterns that are typical for a particular cancer type and highly consistent between sample sets. Further functional validation of these recurrent genetic events in non-transformed primary cells and mouse models of human cancer is hampered by the lack of knowledge of the sequence in which these alterations occur during human tumorigenesis. This temporal order can guide the generation of the correct genomic context in animal models of human cancer and can prioritize the validation of potential drug targets since those changes that occur early during malignant transformation may result in rewiring of the signaling circuitry or confer a state of addiction to the new signal. We have designed a novel computational approach, called Retracing the Evolutionary Steps In Cancer (RESIC), to determine the sequence of genetic events using cross-sectional genomic data from a large number of tumors at their fully transformed stage. RESIC represents the first evolutionary mathematical approach to identify the temporal sequence of mutations driving tumorigenesis and may be useful to guide the validation of candidate genes emerging from cancer genome surveys.

(3) Where do oncogenic alterations arise?

Two independent theories for the initiation of human cancers are often presented: a) the “stem cell” theory and b) a theory that

Publications:


The overall goal of our work is to solve biological problems using quantitative methods from the physical sciences. We use computational methods from statistical physics (e.g. maximum entropy, inference of probability distributions), data sciences and statistics (e.g. machine learning), as well as computer science and mathematics (e.g. network analysis, optimization methods). We apply these methods to build predictive network models of molecular and cellular interactions, to support cancer precision medicine, and to aim for discoveries in structural and evolutionary biology.

Predictive network models: In a wet and dry lab, and in collaborations, we are developing systems biology methods that combine systematic perturbation experiments with rich observational readout, for the de novo derivation of predictive and quantitative models. We aim to model cellular singling events, e.g., between proteins, nucleic acids and metabolites, and cell-cell interactions in cell culture or cancer tissues. To solve the computational complexity of systematic model discovery we adapt methods from statistical physics and data sciences. The translational impact of this technology is in the nomination of combination therapeutics in cancer, such as melanoma, prostate cancer and pancreatic cancer. The cell biological impact depends on the availability of rich perturbation data with rich readout, with dual Crispr screens followed by mass spectrometry an attractive next technological advance and using other datasets from the HMS community may be an exciting opportunity.

Cancer precision medicine: Bioinformatics took on a crucial analytical and infrastructure role in the national and international cancer genome atlas projects. Collaborating closely with clinical researchers, we are applying bioinformatics analysis to cancer genomics, for individual cancer types and across the board. For example, pancancer oncogenic signature analysis of about 10K tumor samples provides an estimate as to which groups of cancer patients can be nominated for particular trials based on shared occurrence of genomic alterations. Collaboration with clinicians for the design of ‘basket’ and ‘match’ trials would be attractive.

Structural and evolutionary biology: In collaboration with the Marks group in systems biology at HMS, we are interested in strengthening the application of maximum entropy methods, with improved inference algorithms, to problems of evolutionary biology, structural biology and cell biology. The derivation of interactions that explain the observed correlations in data is key and there is significant potential for inferring interactions in diverse biological systems beyond proteins and RNA. We would like to generate evolutionarily constrained sequences in the laboratory and further develop a quantitative theory relating biopolymer sequences to phenotypic consequences. General goals and style of science: In general terms, I am in favor of choosing problems that have a chance of impacting the real world, such as human health, global future, synthetic biology, and engineering. Working with

**Publications:**


Antibiotic resistance is a quintessential systems problem. While in certain cases it can result from a single mutation in a single gene, the magnitude and effect of resistance can be dynamic and context-dependent – it can be difficult or impossible to predict from simple tests and a classical, reductionist understanding. My lab seeks to understand the basic science of the evolution and spread of antibiotic resistance, while remaining grounded and guided by the applied problem of fighting the growing public health threat.

With a combination of experimental evolution, algorithm development, and big-data analysis, we endeavor to enable a future in which we “play one move ahead” of antibiotic resistance by manipulating evolution to induce antibiotic sensitivity.

**Publications:**


-Y. J. Jiao*, M. Baym*, A. Veres, R. Kishony, Population diversity can jeopardize the efficacy of antibiotic cycling, in review

Preprint: http://www.biorxiv.org/content/early/2016/10/20/082107

These two papers, when taken together with the ideas presented in a review article I wrote with Roy Kishony last year, lay the foundation for the research program above. Through both adding a spatial dimension to experimental evolution (Baym, et al), and increasing the scale of experiments (Jiao, et al), we show how context-dependent and rare mutations can affect the course of resistance evolution and confound efforts to control it. Indeed, the heterogeneous nature of cross-resistance to antibiotics between different resistance mutations to the same antibiotic implies that evolutionarily-guided treatment strategies will have to be responsive to the particularities of the resistance phenotype and its context. Indeed, this shows that a reductionist, single-mutation single-effect model of resistance is insufficient, and that a systems understanding is required to effectively address the problem.


**Shamil Sunyaev**

*Professor Medicine and Biomedical Informatics*

Harvard New Research Building, Rm 0466C
77 Avenue Louis Pasteur
Boston, MA 02115

Office: 617-525-4708
Lab: 617-525-5953
Fax: 617-525-4705
E-mail: ssunyaev@rics.bwh.harvard.edu

**Publications:**