

Quantitative Biology Colloquium Series

Monday, February 8th

11:00 am - 12:00 noon

HN 310

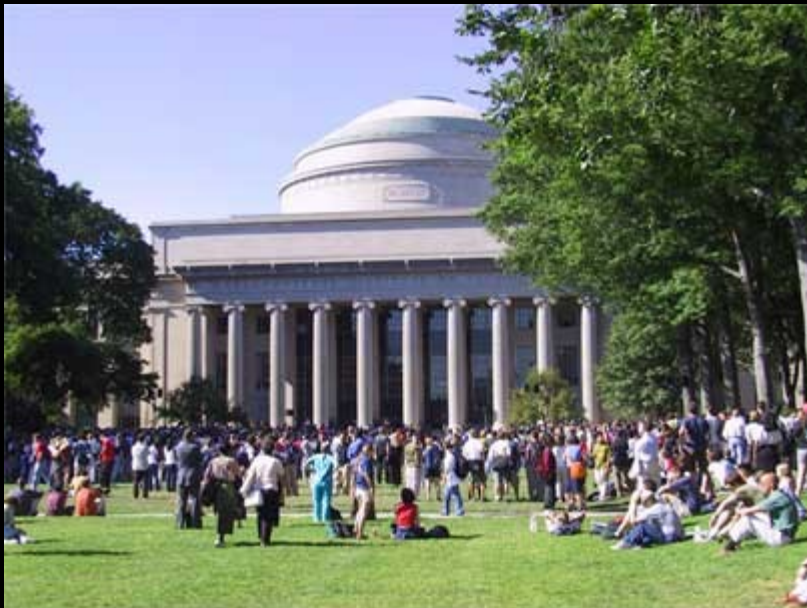
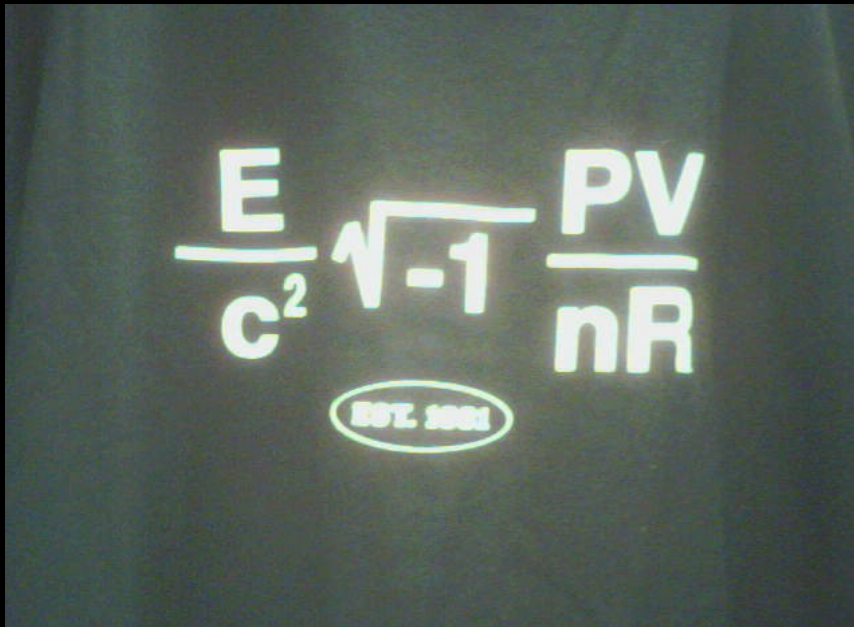
**The View from MIT:
New Bioinformatics
Techniques**

Evan Genest, Vincent Xue and Devin Ghamand

Opportunities at M.I.T

MIT Summer Research Program
HST Summer Institute

MIT Facts #1



MIT Facts #2 A Prominent Biology Department

Top Employment Sectors for MIT Graduates

	Undergrads	Masters
Aerospace/defense	5%	7%
Biological science	8%	3%
Computer technologies	19%	20%
Consulting	13%	18%
Finance	14%	15%
Other engineering	12%	9%

The top employers were McKinsey, Microsoft, MIT, Intel, Amazon, and Bain & Company.

Dr. H. Robert Horvitz

Dr. Phillip A. Sharp

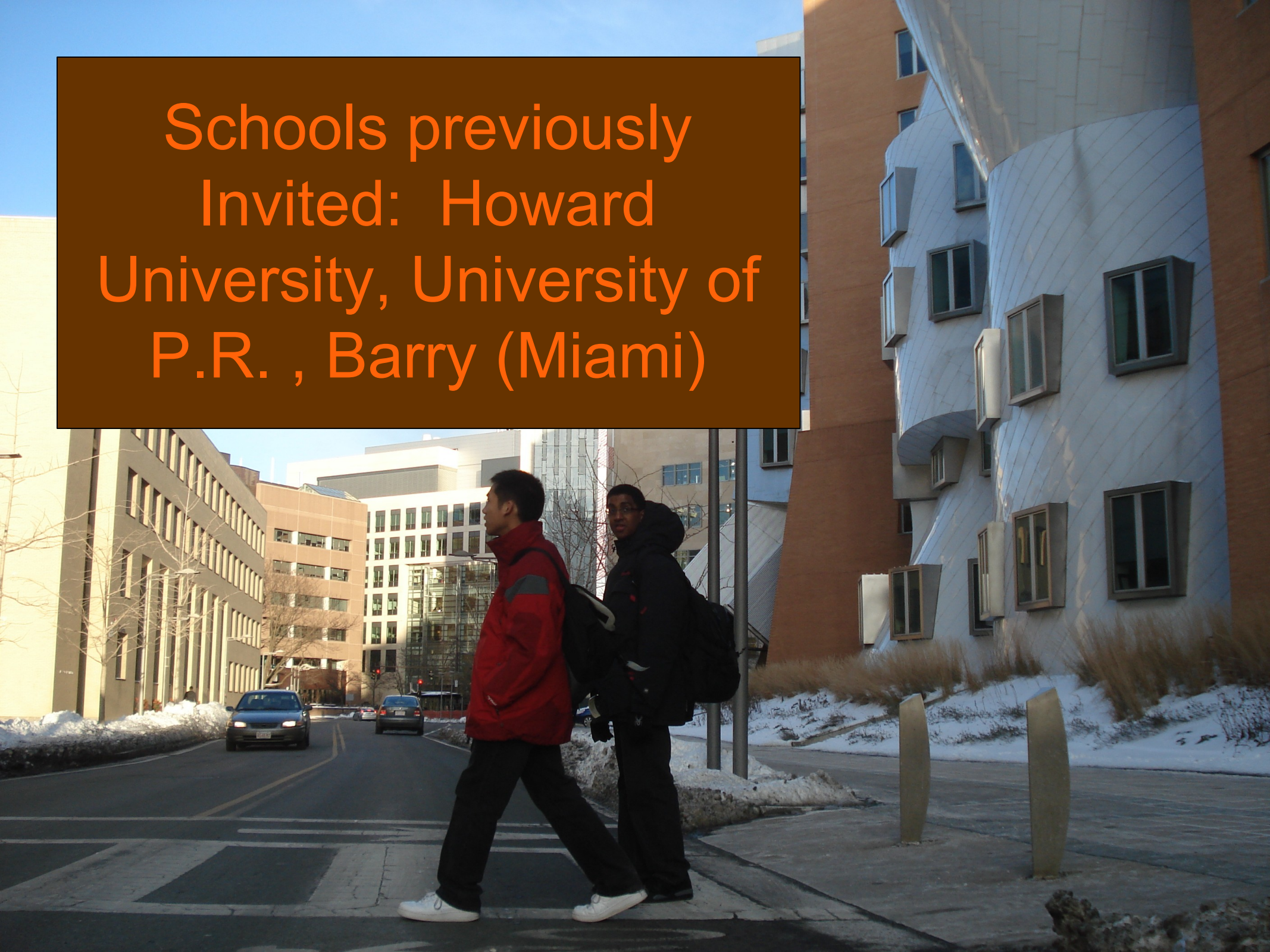
Dr. Susumu Tonegawa

Dr. H. Gobind Khorana

MIT #3 - The Smoot



Schools previously
Invited: Howard
University, University of
P.R. , Barry (Miami)



A Different Topic Each Day

KINETICS

9:30 AM- 11:30 AM Professor Bob Sauer: "Binding and Kinetics"
12:30 - 3:00 PM Dr. Caroline Koehler: "Determining RNA-protein binding constants using filter binding assay and surface plasmon resonance (BIAcore)"
3:30- 4:30 PM Dr. Mandana Sassanfar- Journal club

GENETICS

9:00 AM - 9:30 AM Review and discussion
9:30 - 12:00 Professor Chris Kaiser: "Fundamentals of Genetic Analysis."
1:00 - 3:00 PM Mary Kay Thompson and Jamie Newman: "Using RNAi screens to explore regulatory mechanisms in mammalian cells."

SYSTEMS BIOLOGY

9:30 AM -11:15 AM Professor Ernest Frankel: "Introduction to Systems Biology."
12:15 - 2:15 PM Dr. Tatjana Degenhardt " Chromatin Immuno Precipitation-Sequence and microarray analysis"
2:30 - 4:30 PM Dr. Laura Riva: "Chlp Seq analysis using Python."

BIOSTATISTICS


9:30 - 11:00 AM Professor Michael Laub: "statistics used in Biology"
11:15 - 12:15 Faculty Lunch with Professor Chris Burge
12:30 - 3:30 PM Noah Spies: "Data simulation and analysis using R"

PROTEIN FOLDING + LOW TECH SYSTEMS BIOLOGY

9:30 -11:30 AM Professor Amy Keating: "Protein Structure and Molecular Modeling"
12:30 - 2:30 PM Dr. Emiko Fire/Christopher Negron/ Orr Ashenberg "Protein 3D Visualization, Folding, and Design with PyMol and FoldIt "
3:00 - 4:45 PM Hyun Youk "Imaging data analysis using MATLAB and imagej"

Mornings



A man with dark hair is sitting at a desk, smiling. He is wearing a dark jacket. In front of him is a white laptop with the Google logo on the lid. Behind him is a large projection screen displaying a 3D molecular model of a protein structure. The room appears to be a laboratory or a computer lab.

MON - “Determining RNA-protein binding constants using filter binding assay and surface plasmon resonance (BIAcore)”

TUES - “Using RNAi screens to explore regulatory mechanisms in mammalian cells.”

WED - " Chromatin Immuno Precipitation-Sequence analysis using Python.”

THURS - "Data simulation and analysis using R”

FRI - “ PyMol and FoldIt ”
"MATLAB and imageJ”

Evenings we returned to the hotel and did homework



Biochemistry

The study chemical processes in living organisms.

Includes pathway analysis

Protein interactions

Rate Reactions

Lecturers:

Dr. Robert Sauer

Uses biophysical, genetic, structural, and design strategies to study the relationship between the sequence, stability, and three-dimensional structures of proteins.

Dr. Caroline Koehrer-

Lectured on RNA-protein binding constants using filter binding assay and surface plasmon resonance

Calculating K_D

$$(1) \quad K_D = \frac{[A][B]}{[AB]}$$

$$[A]_{\text{total}} = [A] + [AB]$$

$$[B]_{\text{total}} = [B] + [AB]$$

$$(2) \quad K_D = \frac{([A]_{\text{total}} - [AB])[B]}{[AB]}$$

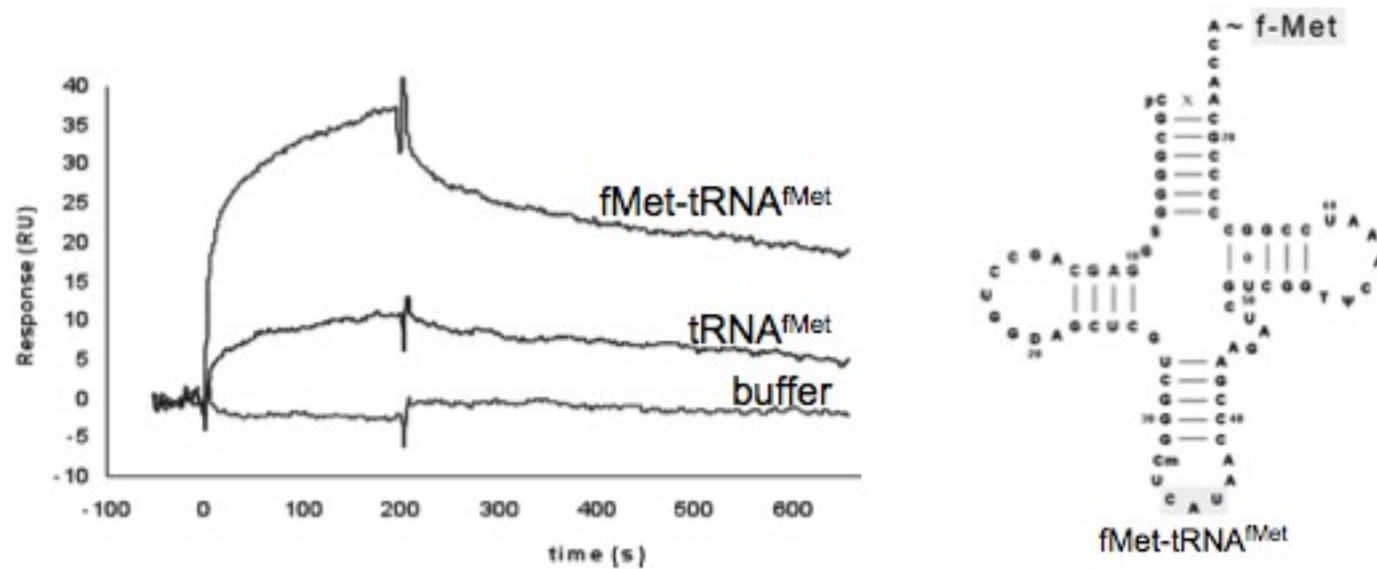
$$(3) \quad K_D = \frac{[A]_{\text{total}}[B] - [AB][B]}{[AB]}$$

$$(4) \quad K_D = \frac{[A]_{\text{total}}[B]}{[AB]} - \frac{[AB][B]}{[AB]}$$

$$(5) \quad K_D = \frac{[A]_{\text{total}}[B]}{[AB]} - [B]$$

$$(6) \quad \frac{[AB]}{[A]_{\text{total}}} = \frac{[B]}{[B] + K_D}$$

Case study: initiator tRNA – initiation factor IF2 complex



Surface ligand: IF2

Analyte: fMet-tRNA^{fMet}; tRNA^{fMet}; buffer

[Mayer, Koehrer et al., 2003: Anticodon Sequence Mutants of Escherichia coli Initiator tRNA: Effects of Overproduction of Aminoacyl-tRNA Synthetases, Methionyl-tRNA Formyltransferase, and Initiation Factor 2 on Activity in Initiation.]

Genetics

The study of heredity

Terminology: Dominance, Recessive, Epistasis

Lecturers

Chris Kaiser









Uses combined genetic, biochemical and cell biological methods to gain an understanding of the molecular mechanisms that underlie protein folding and intracellular trafficking in yeast *S. cerevisiae*

Mary K. Thompson: Graduate student, Ph.D. program in Biology, Gilbert Lab

Jamie Newman: Graduate student, Ph.D. program in Biology Young Lab, Whitehead Institute

Properties Set as a parent Mate **New experiment** Add to strains Remove from strains Punnett Square Feedback

Strains



- m: 9  Mutant 1
- m: 9  Mutant 2
- m: 8  Mutant 3
- m: 9  Sup 1
- m: 6  Wildtype M
-  Wildtype F
-  Homo-m1-F
-  Homo-m1-M


Active experiment


exp-9


Individual Summary

Mating site


Wildtype F

homo-m3-M


Mate


Rename




Discard


All experiments


P(M1-WT) F1(M1-WT) P(M3-WT) F1(M3-WT) Cross_homo-m1_homo-m3 P(M2-WT) **P(M4-WT)** Getting started









































Individual Summary

Mating site


Sup 1

Wildtype M

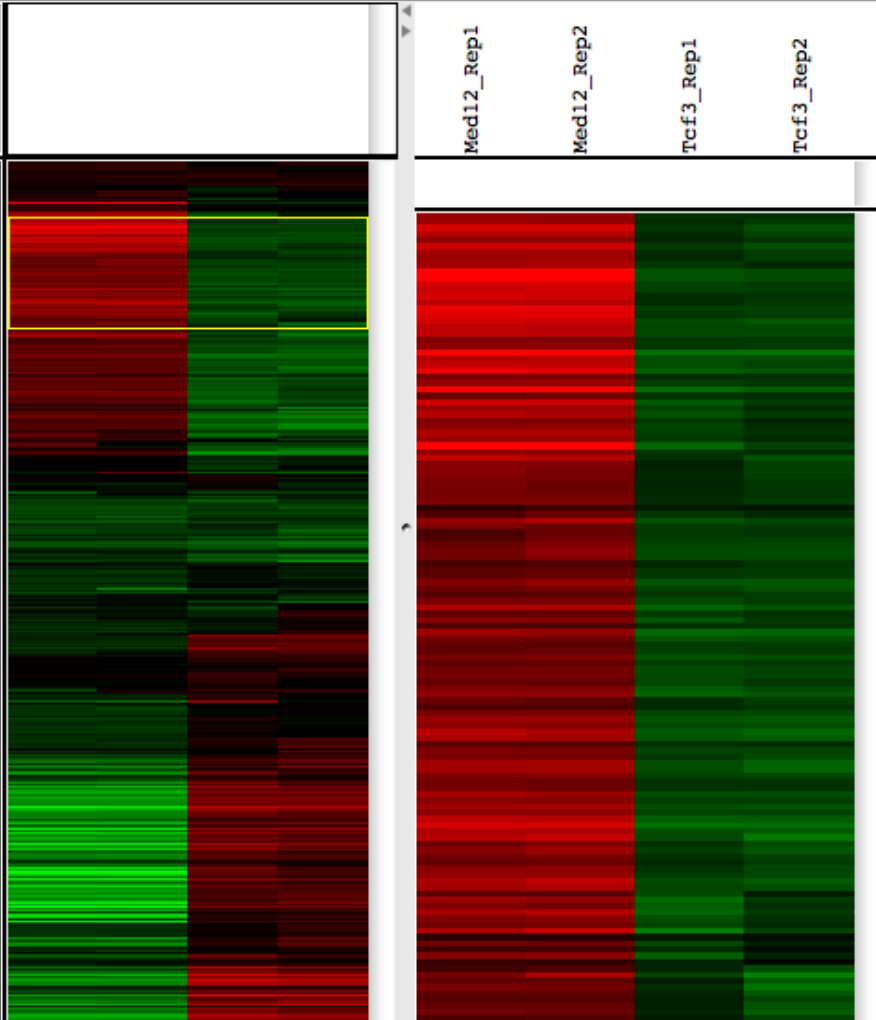
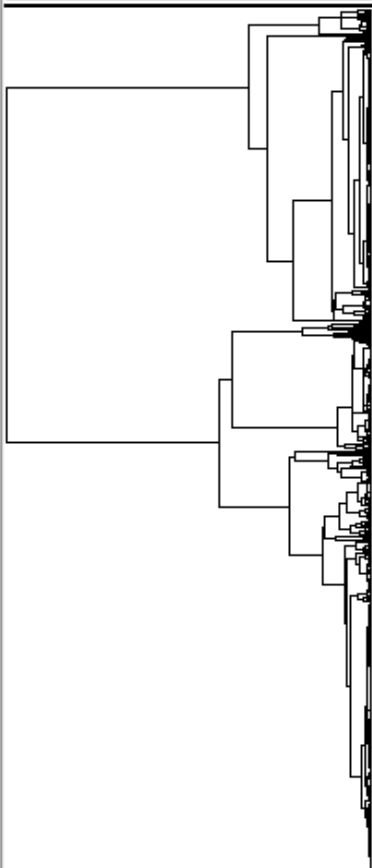

Rename


Discard

									
P(M4-WT)-1P(M4-WT)-2P(M4-WT)-3P(M4-WT)-4P(M4-WT)-5P(M4-WT)-6P(M4-WT)-7P(M4-WT)-8P(M4-WT)-9									
									
1(M4-WT)-10(M4-WT)-11(M4-WT)-12(M4-WT)-13(M4-WT)-14(M4-WT)-15(M4-WT)-16(M4-WT)-17(M4-WT)-18									
									
1(M4-WT)-19(M4-WT)-20(M4-WT)-21(M4-WT)-22(M4-WT)-23(M4-WT)-24(M4-WT)-25(M4-WT)-26(M4-WT)-27									
									

Dendrogram Dendrogram

View Status
Select Node to
view annotation.



Usage Hints
Click to select node
- use arrow keys to navigate tree

462
08E11Rik
07008Rik
17
18016Rik
0
05I04Rik
044
09N05Rik
05R14Rik
a1
40A07Rik
11229Rik
33851Rik
670000050829
4808Rik
54A13Rik
103
196Rik
114
8499
06E14Rik
2106Rik
37106Rik
537
41N04Rik
102918Rik
1he1
1747
2
2676-1
1007
2
1e1

Systems Biology

Purpose:

Discover molecular changes in an unbiased way.
Determine the functional consequences of these changes

Methods:

High throughput measurement
microarray
RNAi screening
Imaging - mass spectroscopy
chIP seq

Computational analysis
Differential equations models
Fishing for information
Visualizing data in an effective way

Ernest Frankel

Systems Biology

Lecturers:

Ernest Frankel

Develops computational and experimental approaches to search for new therapeutic strategies for diseases.

currently focus on the study of cancer,
neurodegenerative diseases and diabetes.

Dr. Tatjana Degenhardt

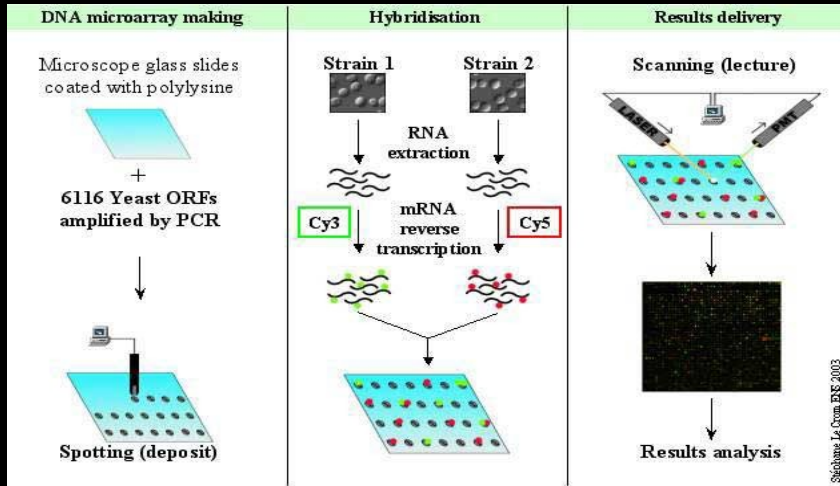
Postdoctoral Associate, Frankel Lab

Dr. Laura Riva

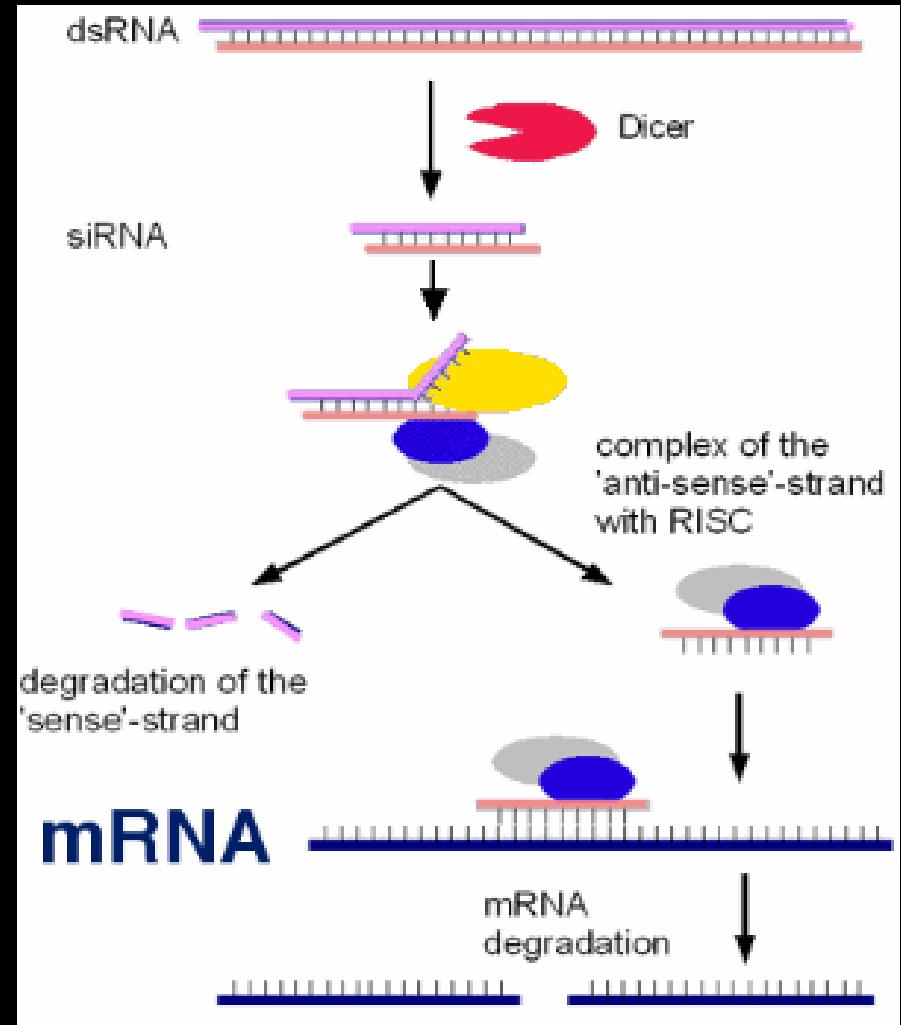
Postdoctoral Associate, Frankel Lab

Methods For High Throuput Analysis

Microarray



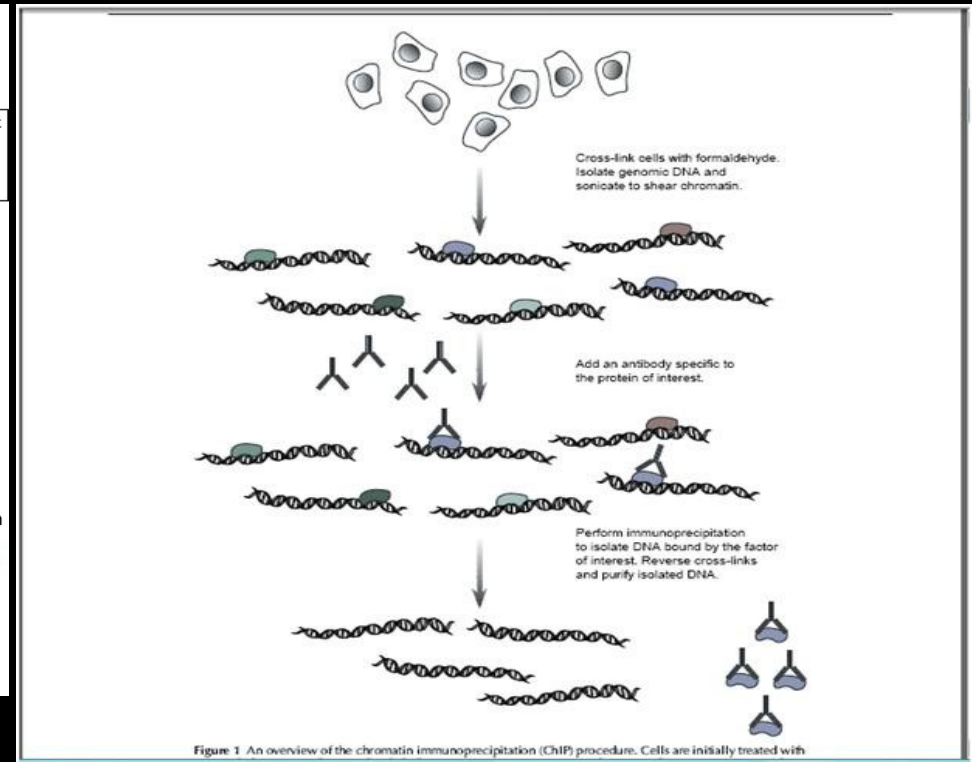
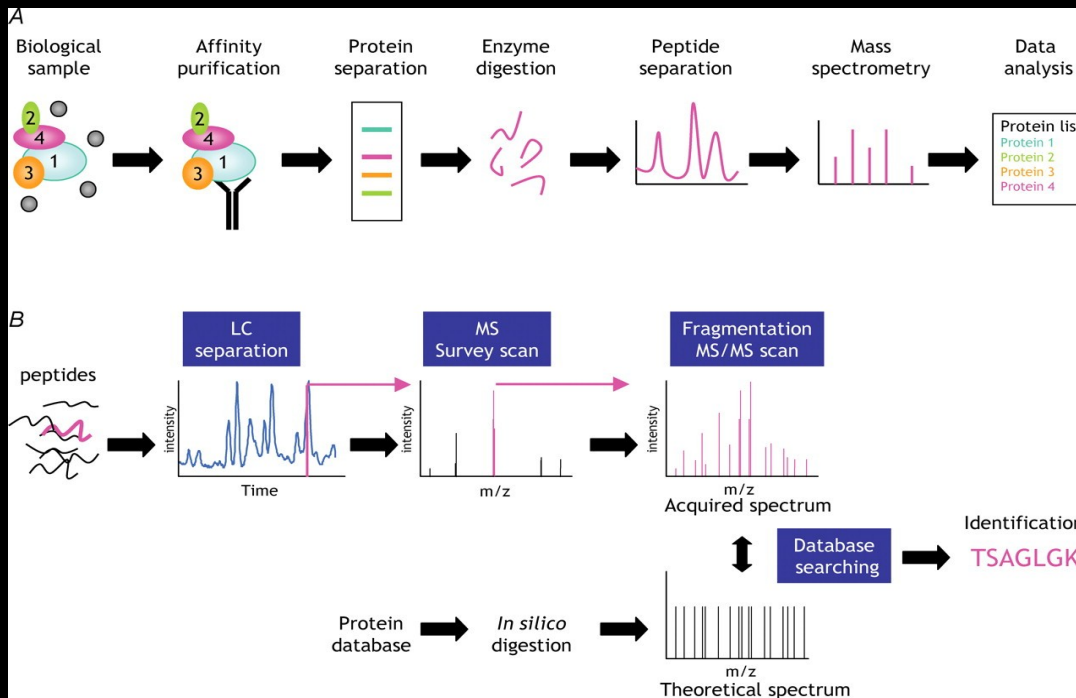
RNAi Screening



Methods For High Throuput Analysis

Mass Spectroscopy

Chromatin immunoprecipitation(chIP) sequencing



Computational Analysis

Differential Equation Models

Fishing for Information

Visualizing data in a useful way

- UCSC Genome Browser

- Panther Database

- DAVID

- oPOSSUM

Python Scripting

```
#!/usr/bin/env python

import string, sys

# If no arguments were given, print a helpful message
if len(sys.argv)==1:
    print 'arguments not provided'
    sys.exit(0)

# Loop over the arguments
for i in sys.argv[1:]:
    file = open(i)
    line=file.read()
    temp=line.upper()
    # Transcribe
    temp=temp.replace('A','U')
    temp=temp.replace('G','Z')
    temp=temp.replace('C','G')
    temp=temp.replace('Z','C')
    temp=temp.replace('T','A')
    print temp
```

Biostatistics:

Using Statistics to solve questions in Biology

Lectures:

Michael T. Laub-

Works to understand how cells process information, make decisions, and control their own behavior

Noah Spies

PhD Student (Bartel Lab)

Protein Modeling and Computational Biology

How does the primary structure of a protein take on its shape?

Lecturers:

Dr. Emiko Fire, Postdoctoral Fellow, Department of Biology

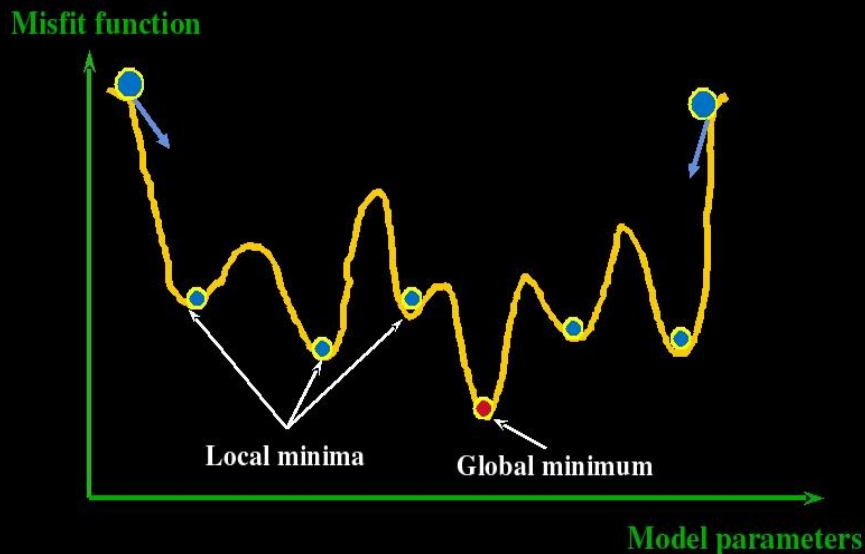
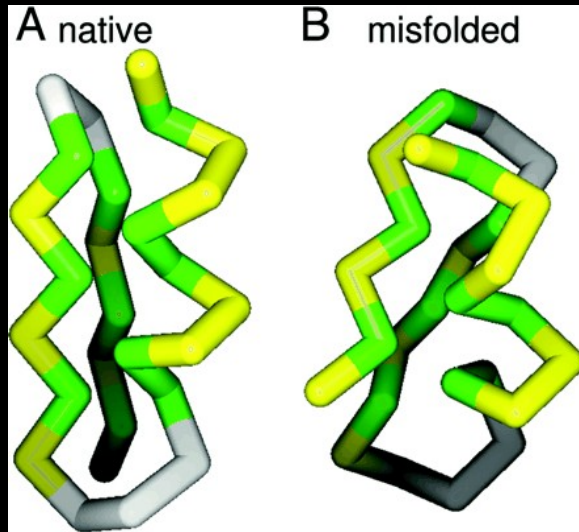
Christopher Negron- PhD student, Computer and System Biology Graduate Program

Orr Ashenberg- PhD student, Computer and System Biology Graduate Program

Hyun Youk - PhD student, Physics Program



Fold-It ; a human-based supplement to avoid local minima in computer based protein folding (A collaboration at the Univ of Washington between the Baker Biochemistry Lab and Dept of Computer Science)



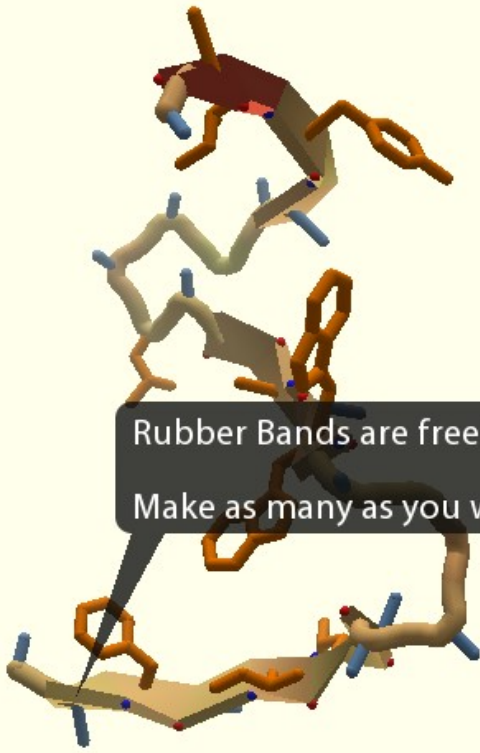


Foldit

Pull Mode

Score: of 8420

Sheets and Ladders



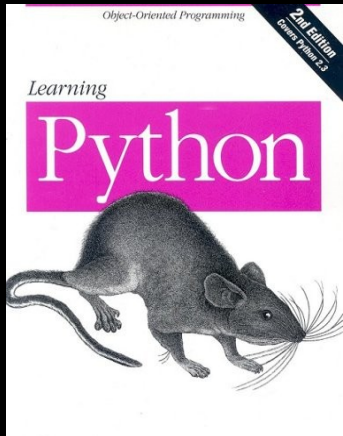
Rubber Bands are free!
Make as many as you want.

Shake Sidechains Wiggle All Freeze Protein Remove Bands Disable Bands Reset Puzzle

▶ Actions ▶ Undo ▶ Menu

▶ Chat - Puzzle Levels ⓘ ✕ auto show
▶ Chat - Global ⓘ ✕ auto show

Wednesday + Thursday: Computing (for systems biology problems)



```
>>>file=open('myfirstfile','w') #create file ('w'=write)
>>>file.write('This is my first python file\n')# write string
into file
>>>file.close() #close file
>>>file=open('myfirstfile','r') #open an existing file
('r'=read)
>>>lines=file.readlines() #read entire file into list of line
strings
>>>lines
['This is my first python file\n']
>>>file.close() #close file
```

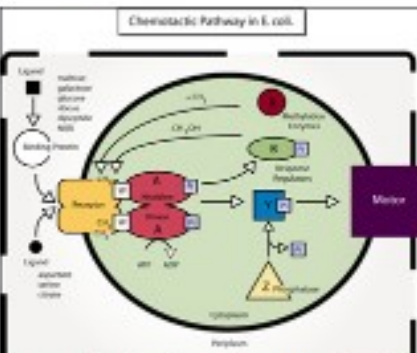
```
1 area <- function(f, a, b, eps = 1.0e-06, lim = 10) {
2   fun1 <- function(f, a, b, fa, fb, a0, eps, lim, fun) {
3     ## function `fun1' is only visible inside `area'
4     d <- (a + b)/2
5     h <- (b - a)/4
6     fd <- f(d)
7     a1 <- h * (fa + fd)
8     a2 <- h * (fd + fb)
9     if(abs(a0 - a1 - a2) < eps || lim == 0)
10      return(a1 + a2)
11    else {
12      return(fun(f, a, d, fa, fd, a1, eps, lim - 1, fun) +
13             fun(f, d, b, fd, fb, a2, eps, lim - 1, fun))
14    }
15  }
16  fa <- f(a)
17  fb <- f(b)
18  a0 <- ((fa + fb) * (b - a))/2
```

MIT Open Course Ware

[<http://ocw.mit.edu/>]

8.591J / 7.81J / 9.531J Systems Biology

As taught in: Fall 2004



Level: Graduate

Instructors: Prof. Alexander van Oudenaarden

Course Features

Course Highlights

Course Description

Technical Requirements

Diagram of the chemotaxis pathway in *E. coli*. (Figure by MIT OCW. Alter figure 4 in Falck, J. J., R. B. Bass, S. L. Butler, S. A. Chen, and M. A. Danielson. "The Two-component Signaling Pathway of Bacterial Chemotaxis: A Molecular View of Signal Transduction by Receptors, Kinases, and Adaptation Enzymes." In *Annu Rev Cell Dev Biol* 13 (1997): 457-512.)

Course Features

- Selected lecture notes

Course Highlights

This course features a set of course notes in the [readings](#) section.

Course Description

This course introduces the mathematical modeling techniques needed to address key questions in modern biology. An overview of modeling techniques in molecular biology and genetics, cell biology and developmental biology is covered. Key experiments that validate mathematical models are also discussed, as well as molecular, cellular, and developmental systems biology, bacterial chemotaxis, genetic oscillators, control theory and genetic networks, and gradient sensing systems. Additional specific topics include: constructing and modeling of genetic networks, lambda phage as a genetic switch, synthetic genetic switches, circadian rhythms, reaction diffusion equations, local activation and global inhibition models, center finding networks, general pattern formation models, modeling cell-cell communication, quorum sensing, and finally, models for *Drosophila* development.

Technical Requirements

[MATLAB® software](#) is required to run the .m files found on this course site.

MIT OPEN COURSEWARE

Home Courses **Donors** About OCW Help Contact Us Enter search input

Home > Courses > Biology > Computational Functional Genomics > Readings

Readings

MIT OCW helps support MIT OpenCourseWare by shopping at Amazon.com! MIT OpenCourseWare offers direct links to Amazon.com to purchase the books cited in this course. Click on the Amazon logo to the left of any citation and purchase the book from Amazon.com, and MIT OpenCourseWare will receive up to 10% of all purchases you make. Your support will enable MIT to continue sharing open access to MIT courses.

LEC#	TOPIC	READINGS
Part 1: Using DNA Sequences to Explore Mechanisms		
1	Course Introduction	
2	Positive Alignment	Dorbin, Richard, Sean P. Kelly, Andrew Singh, and Gaurav Mishra. <i>Statistical Sequence Analysis: Probabilistic Models of Language and Protein Abundance</i>. Cambridge, UK: Cambridge University Press, 1999. ISBN 0521629210.
3	Finding Regulatory Sequences in DNA: Motif Discovery	Hughes, J. D., P. W. Boly, B. Tavazoie, and D. M. Church. "Computational identification of cis-regulatory elements associated with groups of functionally related genes in <i>Saccharomyces cerevisiae</i> ." <i>Journal of Molecular Biology</i> 296, no. 1 (March 18, 2000): 1055-102.
4	Finding Regulatory Sequences in DNA: Motif Discovery (cont.)	
Part 2: Observing the Mechanisms of Transcriptional Regulation		
5	Motif Technology	Fay, Guo, et al. "Revealing Global Regulatory Features of Mammalian Alternative Splicing Using a Quantitative Motifway Platform." <i>Molecular Cell</i> 18, no. 5 (December 23, 2004): 629-641. Zhang, Bin, et al. "The Functional Language of Mouse Gene Expression." <i>Journal of Biology</i> 3, no. 3 (2004): 21. Lee, Tony Bin, et al. "Transcriptional Regulatory Networks in <i>Saccharomyces cerevisiae</i> ." <i>Science</i> 298, no. 5584 (October 20, 2002): 799-804.
6	Expression Arrays, Normalization, and Gene Models	Yang, Yan-Pei, et al. "Normalization for cDNA Microarray Data: A Robust Composite Method for Addressing Single and Multiple Side Systematic Variation." <i>Nucleic Acid Research</i> 29, no. 4 (2001).
		Seaton, M. B., et al. "The Differential Variability of Expression Pairs: Improving Statistical Inference about Gene Expression Changes from Microarray Data." <i>Journal of Computational Biology</i> 8, no. 1 (2007): 27-52.
7	Expression Profiles, Clustering, and Latent Processes	Boon, M. B., et al. "Cluster Analysis and Display of Gene-Variable Expression Patterns." <i>PNAS</i> 95, no. 20 (December 9, 1998): 14862-6.
		Liao, J. C., et al. "Robust Composite Analysis: Reconstitution of Regulatory Signals in Biological Systems." <i>PNAS</i> 100, no. 20 (December 23, 2003): 11620-7.
8	Computational Functional Genomics	Dworki, and Fay. "Probabilistic Sparse Matrix Factorization." Technical Report, University of Toronto, 2004. PDF
9	Gene Cells, and Transcriptional Regulation	De, Y., J. Zhou, D. Morrison, and D. B. Melnick. "Adult pancreatic beta cells are formed by self-duplication rather than stem-cell differentiation." <i>Nature</i> 429, no. 6987 (2004): 47-51. Sivakoff, M. B., et al. "Distinct Types of Diffuse Large B-cell Lymphoma Identified by Gene Expression Profiling." <i>Nature</i> 424, no. 6953 (February 3, 2002): 523-27. Perou, C. M., et al. "Distinct Gene Expression Patterns in Breast Cancer." <i>Nature</i> 396, no. 6702 (October 19, 2000): 76-79. Iwama, N. B., et al. "A Stem Cell Molecular Signature." <i>Science</i> 298, 10023 (October 18, 2002): 871-4.
Part One: An Example of Distinctive Expression Data		

Opportunities at M.I.T

MIT Summer Research Program
HST Summer Institute

Teaching style at M.I.T.



Acknowledgements:
Mandana Sassanfar
Hunter QuBi Grant (NIH)
Veronica Lichman

