

## Bern Seminar Series for Precision Medicine

### Prof. Mark Gerstein

Albert L Williams Professor of Biomedical Informatics, Molecular Biophysics & Biochemistry, Computer Science, and Statistics & Data Science  
Co-director of the Program in Computational Biology & Bioinformatics and the Center for Biomedical Data Science, Yale University

#### Title: Disease Genomics

**Friday June 14 at 12 noon, room H810 (top floor), MEM Murtenstrasse 35  
(Sandwiches will be provided)**

Host: Prof. Dr. Mark Rubin, Director Bern Center for Precision Medicine (BCPM)



After graduating from Harvard summa cum laude with an A.B. in physics in 1989, Prof. Mark Gerstein earned a doctorate in theoretical chemistry and biophysics from Cambridge University in 1993. He did postdoctoral research in bioinformatics at Stanford University from 1993 to 1996. He came to Yale in 1997 as an assistant professor in the Department of Molecular Biophysics and Biochemistry, and since 1999, in the Computer Science Department. He was named an associate professor in 2001, and the following year became co-director of the Yale Computational Biology and Bioinformatics Program. Gerstein has published appreciably in the scientific literature, with >400 publications in total, including a number of them in prominent venues, such as Science, Nature, and Scientific American. His research is focused on bioinformatics, and he is particularly interested in data science & data mining, macromolecular geometry & simulation, and human genome annotation & cancer genomics.

#### Abstract:

My talk will focus on how to leverage thousands of functional genomics datasets to deeply annotate the disease genome and perform data mining to discover disease-associated regulators and variations.

First, I will introduce our computational efforts to perform large-scale integration to accurately define distal and proximal regulatory elements (MatchedFilter) and then show how our extended gene annotation allows us to place oncogenic transformations in the context of a broad cell space; here, many normal-to-tumor transitions move towards a stem-like state, while oncogene knockdowns show an opposing trend.

Second, I will look at our comprehensive regulatory networks of both transcription factors and RNA-binding proteins (TFs and RBPs). I will showcase their value by highlighting how SUB1, a previously uncharacterized RBP, drives aberrant tumor expression and amplifies the effect of the well-known oncogenic TF MYC.

Third, I will introduce a workflow to prioritize key elements and variants. I will showcase the application of this prioritization to somatic burdening, cancer differential expression and GWAS (LARVA, MOAT & uORF tools). Targeted validations of the prioritized regulators, elements and variants demonstrate the value of our annotation resource.

Finally, I will put all these methods together through application to kidney and prostate cancers.