

You are cordially invited to a talk in the Edmond J. Safra Center for Bioinformatics Distinguished Speaker Series, held jointly with The Department of Biochemistry and Molecular Biology, The George S. Wise Faculty of Life Sciences. The speaker is Prof. Mark Gerstein, Computational Biology & Bioinformatics Program, Yale University.

Title: "Comparative Genome Analysis"

Time: Wednesday, June 10 2015, at 14:15 sharp (refreshments from 14:00)

Place: Sherman building, room 105, Tel Aviv University

Host: Prof. Nir Ben-Tal, bental@tauex.tau.ac.il, The Department of Biochemistry and Molecular Biology, Life Sciences Faculty, Tel Aviv University

Abstract: The ENCODE and modENCODE consortia have generated a resource containing large amounts of transcriptomic data, extensive mapping of chromatin states, as well as the binding locations of over 300 transcription-regulatory factors for human, worm and fly. The consortium performed extensive data integration on this data set. Here I will give an overview of the data and some of the key analyses. In particular:

(1) Conservation & Divergence of Transcription

(1a) A novel cross-species clustering algorithm to integrate the co-expression networks of the three species, resulting in conserved modules shared between the organisms. These modules are enriched in developmental genes and exhibited hourglass behavior.

(1b) The extent of the non-coding, non-canonical transcription is consistent between worm, fly and human.

(1c) In contrast, analyses of pseudogene (fossil genes) show that they diverged greatly between the organisms, much more so than genes.

Nevertheless, they had a consistent amount of residual transcription.

(2) Conservation of Regulation

(2a) A global optimization algorithm to examine the hierarchical organization of the regulatory network. Despite extensive rewiring of binding targets, high-level organization principles such as a three-layer hierarchy are conserved across the three species.

(2b) The gene expression levels in the organisms, both coding and non-coding, can be predicted consistently based on their upstream histone marks. In fact, a "universal model" with a single set of cross-organism parameters can predict expression level for both protein coding genes and ncRNAs.

encodenets.gersteinlab.org

encodeproject.org/comparative

pseudogene.org/psicube