Bar Ilan University , Colloquium

Prof. Eran Segal

Dept. of Mathematics and Computer Science, Weizmann Institute

Physics Department (Building 202), room 301

Deciphering codes in DNA sequences

Abstract:

Much of the phenotypic differences among people is attributable to genetic variation in regulatory regions that affect the activity levels of the various genes. However, without a 'regulatory code' that informs us how DNA sequences determine gene activity levels, we cannot predict which sequence changes will affect gene activity levels, by how much, and by what mechanism. To address this challenge, we developed a high-throughput method for constructing libraries of thousands of fully designed regulatory sequences and measuring their gene activity levels in parallel, within a single experiment, and with an accuracy similar to that obtained when each sequence is constructed and measured individually. Using this ~1000-fold increase in the scale with which we can study the effect of sequence on gene activity, we designed and measured the activity levels of libraries in which we systematically perturbed different sequence elements. Our results provide several new insights into principles of gene activity regulation, bringing us closer towards a mechanistic and quantitative understanding of how gene activity levels are encoded in DNA sequence.