Bioinformatics & Systems Biology Prospective Student Weekend 2014

Student Speakers

Roy Ronen   "**Whole genome sequencing of Andean highlanders uncovers the genetics of Chronic Mountain Sickness**"

The hypoxic conditions at high altitudes present a challenge for survival, causing pressure for adaptation. Interestingly, many high-altitude denizens (particularly in the Andes) are maladapted, with a condition known as chronic mountain sickness (CMS) or Monge's disease. To decode the genetic basis of this disease, we sequenced and compared the whole genomes of 20 Andean highlanders (10 with CMS and 10 without). By applying tests of selection based on the site frequency spectrum (SFS), we discovered 11 genomic regions with a strong signature of selective sweep, unique to the adapted (non-CMS) individuals. Two genes in these regions (an erythropoiesis regulator, SENP1, and an oncogene, ANP32D) show higher transcriptional response to hypoxia in individuals with CMS relative to those without. We further found that down-regulating the orthologs of these genes in flies dramatically enhanced their survival rates under hypoxia, demonstrating that suppression of SENP1 and ANP32D plays an essential role in hypoxia tolerance. Our study provides an unbiased framework to identify and validate the genetic basis of genetic adaptation and identifies potentially targetable mechanisms for CMS treatment.

Dinh Diep "**Genome-wide unbiased identification of tissue specific DMRs and design of HOTSPOTS 450K padlock probes**"

This project is inspired by recent data showing that DNA-protein interactions result in differential DNA methylation patterns in a portion of the genome. Although such regions can include biomarkers of epigenetic programming, we have yet to comprehensively study them due to a lack of an appropriate methylation assay to characterize these regions in large enough sample sizes. In the first part of the project, we utilized the Hidden Markov Model to identify differentially methylated regions, which might be affected by DNA binding factors in whole genome bisulfite sequencing data. Next, we developed bisulfite padlock probes as a tool to deeply interrogate the development of those low methylated regions in cells that are undergoing epigenetic programming.

Boyko Kakaradov "**Secrets from a Single Cell**"

Most of the 10^13 cells in your body share nearly identical genotypes, yet each cell has a unique molecular phenotype which differs with its function. To decipher this relationship, we combine single cell sequencing with powerful statistical methods to answer fundamental questions in molecular and cellular biology that have remained unsolvable until recently:

1) Gene expression in cells of the same tissue is a lot noisier than previously known. The prevalence of bimodal expression is particularly puzzling.

2) The immune system relies on asymmetric division of naive T-cells to separate effector and memory lineages as early as the first division after infection.

3) Development and cell-type differentiation are more stochastic than previously shown, especially in neurons.