BIOENGINEERING SEMINAR

“The Human Transcriptome across Tissues and Individuals”

Wednesday – July 6, 2016 – 11:30 a.m.
EPFL – room SV1717

Prof. Roderic Guigó

Center for Genomic Regulation, Universitat Pompeu Fabra, Barcelona (E)

host: Prof. Bart Deplancke

Abstract

The pilot phase of the Genotype-Tissue Expression (GTEx) project has produced RNASeq from 1,641 samples originated from up to 43 tissues from 175 post-mortem donors, and constitutes a unique resource to investigate the human transcriptome across tissues and individuals. Clustering of samples based on gene expression recapitulates tissue types, separating solid from not solid tissues, while clustering based on splicing separates neural from non-neural tissues, suggesting that post-transcriptional regulation plays a comparatively important role in the definition of neural tissues. About 47% of the variation in gene expression can be explained by variation of across tissues, while only 4% by variation across individuals. We find that the relative contribution of individual variation is similar for IncRNAs and for protein coding genes. However, we find that genes that vary with ethnicity are enriched in lncRNAs, whereas genes that vary with age are mostly protein coding. Among genes that vary with gender, we find novel candidates both to participate and to escape X-inactivation. In addition, by merging information on GWAS we are able to identify specific candidate genes that may explain differences in susceptibility to cardiovascular diseases between males and females and different ethnic groups. We find that genes that decrease with age are involved in neurodegenerative diseases such as Parkinson and Alzheimer and identify novel candidates that could be involved in these diseases. In contrast to gene expression, splicing varies similarly among tissues and individuals, and exhibits a larger proportion of residual unexplained variance. This may reflect that stochastic, non-functional fluctuations of the relative abundances of splice isoforms may be more common than random fluctuations of gene expression. By comparing the variation of the abundance of individual isoforms across all GTEx samples, we find that a large fraction of this variation between tissues (84%) can be simply explained by variation in bulk gene expression, with splicing variation contributing comparatively little. This strongly suggests that regulation at the primary transcription level is the main driver of tissue specificity. Although blood is the most transcriptionally distinct of the surveyed tissues, RNA levels monitored in blood may retain clinically relevant information that can be used to help assess medical or biological conditions.