Terry Speed

University California, Berkeley, USA & The Walter and Eliza Hall Institute, Melbourne, AUS

"Normalization of RNA-Seq Data:

Are the ERCC Spike-In Controls Reliable?"

Friday, March 7th, 2014 at 11:00

Génopode Auditorium B

Host: Mauro Delorenzi

Abstract

Normalization of RNA-Seq Data: Are the ERCC Spike-In Controls Reliable?

Joint work with Sandrine Dudoit, Davide Risso and John Ngai, all from UC Berkeley.

The External RNA Control Consortium (ERCC) developed a set of 92 synthetic polyadenylated RNA standards that mimic natural eukaryotic mRNA (Jiang et al., 2011). The standards are designed to have a wide range of lengths (250-2,000 nucleotides) and GC-contents (5-51%). The ERCC standards can be spiked into RNA at various concentrations prior to the library preparation step and serve as negative and positive controls in RNA-Seq. Ambion commercializes spike-in control mixes, ERCC ExFold RNA Spike-in Control Mix 1 and 2, each containing the same set of 92 standards, but at different concentrations.

We investigate the use of the ERCC spike-in controls for two main purposes: (a) Quality assessment/quality control (QA/QC) of RNA-Seq data and benchmarking of normalization and differential expression (DE) methods, and (b) Direct inclusion in between-sample normalization procedures.

We have two RNA-seq data sets which make use of the ERCC controls: a local one concerning treated and untreated zebrafish tissue, and some of the SEQC samples.

A variety of normalization methods will be compared, both using and not using the ERCC controls.