Pathway-based analysis of multi-level omics data for tumour subtypes classification

Loredana Martignetti*[1], Laurence Calzone[1], Arnau Montagud[1], Andrei Zinovyev[1], and Emmanuel Barillot[1]

[1]Cancer et génôme: Bioinformatique, biostatistiques et épidémiologie d'un système complexe – Inserm : U900, Institut Curie, MINES ParisTech - École nationale supérieure des mines de Paris – 26 rue d'Ulm - 75248 Paris cedex 05, France

Abstract

In many analysis of high-throughput data in systems biology, there is a need to quantify the activity of a set of genes in individual samples. In cancer the same pathway can be affected by defects in different individual genes in different patients and application of pathway-based approaches in the analysis of genomic data can help to capture biological information that is otherwise undetectable by focusing on individual genes.
We present here ROMA (Representation and quantification Of Module Activities) software designed for fast and robust computation of the activity of gene sets (modules) with coordinated expression. Focusing on small set of genes at time in omics data analysis avoids the curse of dimensionality and generates robust results much less sensitive to noisy measurements and to heterogeneity of samples than ranking procedures of single genes.
ROMA can be applied in many contexts, from estimating differential activities of transcriptional factors to finding activated/inactivated pathways in single-cell transcriptomics data. We present here the principles of ROMA providing a practical example of its use. We applied it to integrate multi-level omics data to compare distinct cancer subtypes in terms of activated/inactivated signalling pathways and transcriptional programs.

*Speaker