

Abstract

In this paper, we present the brand-new concept of in silico metabolomics. The idea is that metabolomics data can be imputed through computational methods from already available RNA expression data. We show that LASSO regression can effectively predict metabolomics using transcriptomics data under sparsity condition given that the relationship is linear. We validate the importance of this prediction by demonstrating the importance of metabolomics data to immune profiling, and thus potentially to immunotherapy. We first analyze kidney cancer data and generalize our findings to pan-cancer data, including breast cancer, colon cancer, liver cancer, ovarian cancer, prostate cancer, and lymphoma to identify previously unknown potential relationships between metabolites and genes.