TWO STAGE LINEAR MIXED MODEL FOR PROTEIN EXPRESSION ANALYSIS IN 2D-DIGE

Fernández, Elmer Andrés[12]. Girotti, Romina Maria[23]. Llera, Andréa Sabina[23]. Cantet, Rodolfo[24]. Balzarini, Mônica[25].

1- CONICET, 2-Catholic University of Córdoba, School of Engineering, 3-Leloir Fundation, Laboratory of Molecular and Cellular Therapy, 4- University of Buenos Aires, School of Agronomy-Biometry, 5-National University of Córdoba, School of Agronomy-Biometry.

In two dimensional difference gel electrophoresis (2D-DIGE) the usual technique to find differentially expressed proteins, between treatments, is the Student's t test. Its use does not account for non-biological experimental effects such as dye, gel and subject effects. These sources of variation could invalid t-test statistical assumptions. To overcome this limitation commercial softwares usually apply different ad-hoc normalization strategies prior the analysis of differentially expressed proteins. The selected normalization procedure impacts the protein identification. Here we propose a statistical modelling framework that fit linear mixed models in two stages. The first one (the normalization model) allows to model experimental effects and heterocedasticity in the error. The second stage models residual from the first one to analyze protein and treatment-by-protein interactions. Specific applications of the proposed strategy on 2D-DIGE experiments yielded better results than the normalization procedures and classical t-test available in commercial software. The statistical modeling framework is flexible enough to be applied in different kinds of experimental designs with 2D-DIGE. It allows the researchers to analyze the different source of variation, studying and evaluating their impact in the biological inference. The proposed modelling strategy was illustrated with the analysis of human proteome to study molecular pathways affected by the tumorigenic "Secreted Protein Acid and Rich in Cysteine (SPARC)" in a melanoma cell model. Known differentially expressed proteins associated to SPARC were successfully identified by means of the proposed analysis strategy.