Missing values in gel-based proteomics.


Abstract

Gel-based proteomics is a widely applied technique to measure abundances of proteins in various biological systems. Comparison of two or more biological groups involves matching of 2-D gels. Depending on the software, this can result in spots showing missing values on several gels. Most studies ignore this fact or substitute all missing data by zero. Since a couple of years, scientists have realized that this is not the optimal way of analyzing their data and several studies were published presenting methods of imputing missing proteomics data. Most of these methods have already been applied to microarray data before; the phenomenon of missing data is well known in this field, too. With this review, we intend to further raise awareness of the problem of missing values in gel-based proteomics. We summarize reasons for missing values and explore their distribution in data sets. We also provide a comparison and evaluation of hitherto proposed imputation methods for gel-based proteomics data.

Involved Units and Groups

Molecular and Applied Microbiology Systems Biology and Bioinformatics

HKI-Authors

Prof. Dr. Axel A. Brakhage  Dr. Olaf Kniemeyer  Prof. Dr. Reinhard Guthke  Dr. Daniela Albrecht-Eckardt

Topics
Management of heterogeneous experimental data

Identifier

doi: 10.1002/pmic.200800576 PMID: 20077407