Overall quantitative pattern extraction analysis of two-dimensional electrophoresis images from prostatic cancer mouse model

ID - 140

Mazzara Saveria[1], Pattini Linda[1], Vicquery Francesca[2,3], Bellone Matteo[3], Alessio Massimo[2], Cerutti Sergio[1]
[1]Department of Biomedical Engineering, Polytechnic University, Milan, Italy
[2]Proteome Biochemistry - San Raffaele Scientific Institute, Milan, Italy
[3]Cellular Immunology v San Raffaele Scientific Institute, Milan, Italy

Motivation

Two-dimensional gel electrophoresis (2DE) is still a key component of current proteomics, since it represents an indispensable tool for the analysis of protein expression profiles in complex biological systems such as whole cells and tissues.

The technique provides a 2D map which couples orthogonally a charge-based (isoelectric point, pI) to a size-based separation (relative molecular mass, Mr).

The maps obtained from proteins migration are acquired as grey level images which can be processed to allow the analysis and the comparison of the experimental outcome for different samples. However the complexity and the intrinsic variability of the maps make this task difficult and time consuming. The analysis of gel images is a very labour intensive step and involves a considerable expertise to properly extract information. The process, in general, takes advantage of dedicated tools for quantitative analysis and comparison of the gels, meant as collections of identified spots, that are matched through the different samples by means of image registration techniques.

Beside these tools, or better, in a complementary way, it could be useful to develop strategies for the automatic classification of the whole gels, based on the assessment of the complete ensembles of spots shown in the maps, providing a schematic representation of the overall outcome of the experiments without considering the individual spots in detail. We developed such a strategy and it was applied to the gel images set acquired in a study on prostatic carcinoma, accomplished through the use of an animal model of the pathology: the TRAMP mouse (TRansgenic Adenocarcinoma Mouse Prostate). The 2DE maps were generated from serum samples drawn from transgenic and wild-type mice, in three different developmental stages.

Methods

Some works may be retrieved in the recent literature where efforts towards the classification of 2DE maps are described; briefly quantitative image descriptors, extracted on the basis of pixel intensities quantified in the different zones in which the image is partitioned, are analyzed with techniques of dimensionality reduction to discriminate different clinical conditions. Starting from this kind of approach we intervened on the definition of the gel descriptors: the idea was that the features refer to areas of the image segmented as spots excluding from the quantitative description artifacts and background signal; then the position of the detected spots are considered, no longer in pixel, but in terms of pI and molecular weight (MW), after ad hoc calibration. This simple step makes the samples comparable, without the accomplishment of a canonic image matching by means of registration techniques; indeed the new space (pI, MW) is invariant in respect of the alterations of the protein migration, allowing the inclusion in the analysis of gel images that otherwise had to be excluded, because of the lack of the necessary homogeneity in respect of the other samples. The quantitative descriptors so defined were processed through the principal component analysis.

Results

The informativeness of the chosen descriptors allows to see the gels of the data set as items in a privileged three-dimensional space, segregating according to their biological conditions. The samples, both for the transgenic mice and for the control mice, obtained in different times of development, appear as linearly separable.

The scatterplots for the transgenic vs. wild type comparison, performed in the three developmental ages, show a degree of separability according to the disease course, in agreement with the biological knowledge.

The method proposed is highly repeatable, does not need any a priori information. It may provide an effective visualization tool and important clues for classification; thus it could represent a reliable complement to the routine of a proteomics lab to perform rapid and systematic screening tests capturing the essential impression of the 2DE gel image.

References

Marengo E, Robotti E, Antonucci F, Cecconi D, Campostrini N, Righetti PG: Numerical approaches for quantitative analysis of two-dimensional maps: A review of commercial software and home-made systems. Proteomics 2005, 5:654-666.

Email: bruno.molina@fastwebnet.it