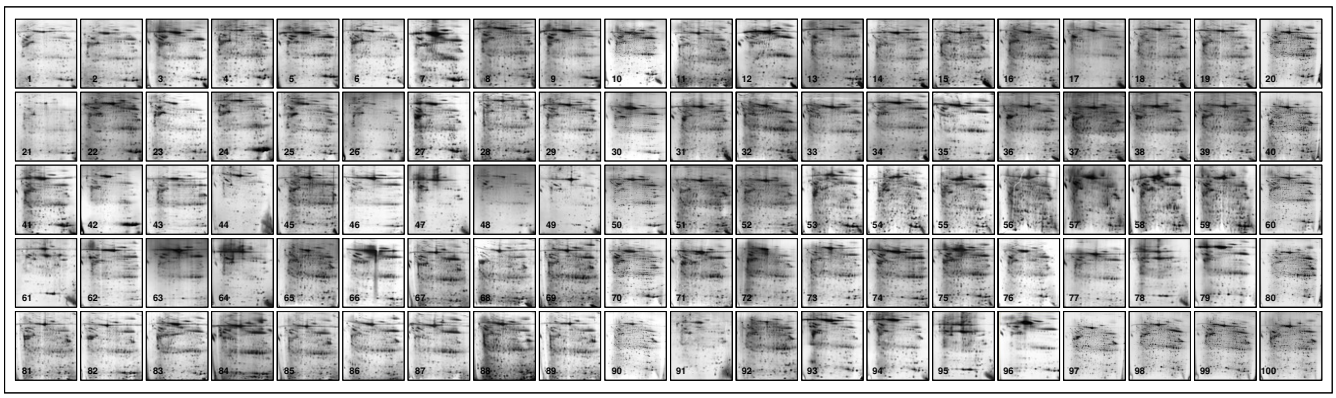
GENE ONTOLOGY-BASED ANNOTATION AND COMPARATIVE ANALYSIS OF PROTEINS EXTRACTED FROM PROTEOMICS OF 100 BREAST CANCER PATIENTS

Pucci-Minafra[1,2] I., Cancemi[1,2] P., Albanese[1,2] NN., Di Cara[1,2] G., Costantini[1,2] F., Marabeti[1,2] MR., Marrazzo[3] A., Taormina[3] P., Minafra[2] S.
[1]Dipartimento di Oncologia Sperimentale e Applicazioni Cliniche (DOSAC), Università di Palermo,
[2]Centro di Oncobiologia Sperimentale (COBS),
[3]Casa di Cura di Alta Specialità La Maddalena, Palermo, Italy.
E-mail: idapucci@unipa.it

Background: Current clinical parameters for breast cancer diagnosis and therapy are: tumour size, axillary lymph node status, histological grading and presence or absence of metastases. Prognostic/predictive properties, such as oestrogen and progesterone receptor status, and human epidermal growth factor receptor (HER-2/neu) status are currently used for therapeutic decision. Conversely, it is now emerging that the number of genetic mutations and epigenetic deregulations in cancer is far more higher than previously thought. Therefore, proteomic screening for differential protein expression in subsets of tumor samples is an essential tool for generating data bases and biomarker discovery. The aim of present study was to extend to a significant number of proteins and patients our previous data, contributing to the knowledge of biological pathways involved in breast cancer and to the molecular-based classification of breast cancer patients.
Methods: Aliquots of breast cancer tissues and their adjacent non-tumoral tissues were obtained during surgical intervention from patients who did not receive any neoadiuvant therapy. Diagnosis of ductal infiltrating breast cancer (DIC) was confirmed by histopathology. Sample preparation for proteomics was performed as described (Pucci-Minafra et al. Proteomics Clinical Applications, 1, 118-129, 2007; Pucci-Minafra et al. Journal of Proteome Research 7, 1412-18, 2008). Quantitative determination of protein spots was normalized for actin.
Results: We have collected for the present study 100 proteomic maps from G2/G3-DIC patients, and 13 from non-tumoral mammary tissues. All detected proteins were identified, or confirmed, by mass spectrometry. Collectively we have annotated 209 protein spots, corresponding to 122 genes. Genes were analyzed by the instruments of DAVID Bioinformatics Resources (Dennis et al. 2003; Huang et al. 2009). The Gene Accession Conversion tool recognized 112 unambiguous Gene IDs, over the 122 ones present in our list. The gene list was correlated to 136 functional terms, but only 36 correlations were highly significant (Benjamini values from $1x10^{-8}$ to $5.6 x10^{-3}$). Nine terms over the 36 corresponded to the function of apoptotic processes and in particular 24 genes were related at the highest significant value ($1x10^{-8}$) with the regulation of programmed cell death; 50% of the genes belonging to this category, codify for proteins with anti-apoptotic functions. The cluster of anti-apoptotic proteins, corresponding to 12 genes and 22 protein isoforms was compared among the 100 proteomics maps and the 13 reference non-tumoral tissue maps. The comparative proteomic profiling showed: 1) a highly significant overexpression of several members of the anti-apoptotic protein cluster in the cancer tissues *vs* non tumoral counterparts; 2) a relative variability of the expression levels of the normalized proteins within patients. Among proteins of this category reaching the high levels of expression, we observed the nucleophosmin (NPM), a crucial regulator of p53; the translationally-controlled tumor protein (TCTP1), a protein involved in calcium binding and microtubule stabilization; cofilin (COF1), an actin-binding protein responsible also for the signal translocation from cytoplasm to nucleus; annexin A1 (ANXA1) a calcium/phospholipid-binding protein which promotes membrane fusion and ruffling, and the glutathione s-transferase (GSTP1), which plays important roles in detoxification but having also a role in susceptibility to cancer.
Conclusions: The application of the powerful Bioinformatics Resources for gene/protein classification provided by DAVID knowledgebase, while confirming our previous protein classification, introduced new terms for further remodulation of protein clusters on the basis of the multiple functions for individual proteins. In particular we found a high number of proteins with specific biochemical functions, converging towards common pathways. Overall, a predominant pathway was the programmed cell-death, which resulted to be the most robust among patients, both as number of proteins involved and as level of significativity. We believe that the present collection of human breast cancer proteomics represents a valid contribution for clinical applications to breast cancer.



Overview of the miniaturized proteomic maps of the 100 patients utilized for present investigation. Numbers on the maps correspond to individual patients.