ACTA ONCOLÓGICA BRASILEIRA

Acta Oncol. Bras. Vol. 24 - nº 01, 572-575, jan-fev-mar, 2004

PROTEOMIC AND TRANSCRIPTIONAL PROFILING OF HUMAN BREAST CANCER

MICHAEL J O'HARE[1], SUNIL R LAKHANI[2], AND A. MUNRO NEVILLE[3]

Introduction

The LICR/UCL Breast Cancer Laboratory was set up 10 years ago, under the auspices of the Ludwig Institute, to search for new markers in breast cancer. The objective has been to find molecules that can be used for improved diagnosis and prognosis as well as to find new targets for future therapies. As a truly translational laboratory it has used human clinical samples exclusively in its quest for these targets. It has accrued over 1000 fresh tissue samples, consisting of normal breast tissues (reduction mammoplasties) primary cancers and metastatic disease (lymph nodes and pleural/ascitic effusions). Its strategy has been threefold. Firstly it has developed cell biological methods for the purification and enrichment of specific cell types from the different tissues as well as novel methods for the immortalization of the normal cells. Secondly it has employed a variety of progressively more global analytical techniques to profile purified normal luminal epithelial cells and compare them with their malignant counterparts. These methods have included proteomics with large-scale 2D gel electrophoresis, gel imaging and quantitative analysis, followed by annotation of selected protein isoforms by mass spectrometry (in collaboration with Oxford Glycocsciences). Analysis at the transcriptional level has included differential display, cDNA microarrays (Sanger Consortium) and most recently MPSS (Massively Parallel Signature Sequencing), an extremely powerful method invented by Sydney Brenner and colleagues and implemented commercially by Lynx. In addition, samples from the laboratory have contributed to both the Ludwig SEREX program and the ORESTES database. As well as primary cell samples we have also analysed widely available breast cancer cell lines and a novel in-house longitudinal model of breast cancer based on an immortalized normal luminal cell line (HB4A) which has been susequently transformed by over-expression of c-erbB-2 at different levels (C3.6, C5.2). As a result of these activities and collaborations we have helped generate what is probably one of the largest existing proteomic and transcriptional database set for human breast cancer. Analysis and mining of these datasets is now in progress to identify promising targets for validation. The validation of putative markers and targets is the third part of our strategy and involves both analysis at the pathological level by examining protein distribution across many tumor samples using tissue microarrays (TMA), and the development of appropriate biological systems to examine the effects of knocking out or knocking in genes of interest.

Cell separations

Separation techniques for isolating purified populations of normal and malignant cells have been based largely on the use of cell-type specific surface antigens (1), such as Epithelial Membrane Antigen (EMA) on luminal cells in the breast, CD10 (Common Acute Lymphoblas-

1. *LICR/UCL Breast Cancer Laboratory, Royal Free and University College Medical School, London, UK*
2. *The Breakthrough Toby Robins Breast Cancer Research Centre, Institute of Cancer Research, London, UK*
3. *Ludwig Institute for Cancer Research, London, UK*

Proofs and reprints to: 1 - Ludwig Institute for Cancer Research - 91 - Riding House Street, London - W1W 7BS - UK - e-mail: m.o'hare@ucl.ac.uk

Acta Oncol. Bras. Vol. 24 - nº 01, 572-575, jan-fev-mar, 2004

tic Leukemia Antigen) on the myoepithelial cells, thrombomodulin on microvascular endothelial cells and Fibroblast Activation Protein (FAP) on activated fibroblasts found in breast cancers. Initial separations using FACS were soon replaced by immunomagnetic methods for improving cell yields, as several of the global profiling methods require substantial amounts of protein or mRNA, corresponding to several million cells. For normal cells we use positive selection as the antigens are generally well-expressed on all cells of any given type. To improve purities we often combine positive selection with a subsequent negative selection to remove any residual contaminating cells. This is possible by combining two types of immunomagnetic microbeads, small MACS (Miltenyi) beads (~50nm diameter) with large Dynabeads (~4 microns) as the smaller MACS beads are not 'seen' by the magnet used to separate Dynabeads. Different bead sizes are also combined with primary antibodies of different species or isotypes, to produce an efficient high yield system whereby upwards of 108 cells can be prepared. A different principle is used for tumor cell enrichments because no truly tumor cell-specific antigen exists, and antigens such as EMA are often expressed sporadically or internally by malignant cells. Tumor cell separations are therefore based on disaggregation and filtration to remove normal breast ducts and lobules, as well as in situ disease, followed by removal of activated fibroblasts using the F19 Ludwig antibody against FAP (4). Although highly efficient methods for immortalizing normal cells have been developed (7), we have failed to significantly improve either the take rate or growth rate of primary human breast cancer cell in vitro. We therefore rely on the cell separation methods to provide sufficient freshly isolated cells for analysis. There is one clinical situation in which sufficient freshly isolated tumor cells can be prepared for dynamic experiments involving multiple aliquots for analysis. This is the malignant effusion of either pleural or ascitic origin. We have isolated upwards of 10g of cells in some such cases. Separation of malignant from non-malignant cells in these samples is relatively simple. The latter consist predominantly of reactive mesothelial cells or macrophages, both of which have a propensity to attach in culture within a few hours, unlike the tumor cells which are often in the form of well-differentiated acinus-like spheroids, which are slow to attach and can be harvested in the supernatant.

Proteomic analysis

Proteomic analysis by 2D SDS-PAGE and MALDI-MS or MS-MS has now been carried out on 10 normal luminal and myoepithelial cell preparations (8), a total of 33 enriched tumor cell samples (4), and a panel of 30 breast-derived cell lines from the public domain (ATCC) (3). The first study, of normal luminal and myoepithelial cells, was carried out as a proof of principle that breast purified cell populations would yield a large number of differentials which could be quantified and annotated. That in fact proved to be the case with a total of over 43,000 protein isoforms detected from 20 samples, yielding a total of over 1,700 unique proteins, of which 170 were differentially expressed by 2-fold or greater. A total of 51 differential and 134 non-differential proteins were identified by tandem mass spectrometry. Using these data as a primary reference, a second set of separated normal luminal cells was compared with a set of enriched freshly isolated tumor cells. It is worth noting that cell separation of some form is an essential pre-requisite of proteomic analysis of tissue. Intact tissues contain too much plasma and extracellular protein to yield much information of value when analysed in this manner. A total of 8118 protein isoforms were included in the analysis. The statistical analysis of the normal versus tumor data was carried out in such a way (LDA, Linear Discriminant Analysis) as to identify a protein 'signature' consisting of 11 protein isoforms, whose relative abundance not only separated normal from tumor samples but also sub-classified the tumors in a manner that corresponded with their pathological classification, e.g. DCIS, LCIS etc. One of the unique features of the proteomic approach is that it identifies protein isoforms, including post-translationally modified variants. These are often biologically significant and would not be detected by transcriptional analysis. Several of the members of the protein breast cancer 'signature' were specific isoforms, with no differences between other isoforms of the same core protein. Isoforms pose some problems at the validation level as isoform-specific antibodies are often not

Acta Oncol. Bras. Vol. 24 - nº 01, 572-575, jan-fev-mar, 2004

available, and pan-specific antibodies would not reveal subtle isoform differences. Nevertheless, the differential expression of several of these proteins has been verified on tissue microarrays, and studied in tumor progression from in situ to invasive disease. One protein was also shown to provide independent prognostic information on the outcome of disease in individual cases in terms of survival. In an attempt to define further the proteins associated with enhanced malignancy we have analysed the proteomes of a panel of established cell lines that included those which establish poorly as xenografts, and others which establish well and which metastasise in this model. A total of six differentially expressed proteins associated with xenograftability were identified and annotated, including examples know to be associated with poor outcome in human breast cancer, such as vimentin.

Transcriptional profiling

Our first exercise in global transcriptional profiling was the analysis of the HB4A/cErbB-2 longitudinal model of transformation (2,9). The objective here was to identify novel genes either up or down-regulated as a consequence of erbB-2 overexpression, independently of the chromsome 17q12-q21 amplicon associated with amplification of this gene in human breast cancer. The platform we used for this study was the Sanger 10K cDNA array available as part of the LICR-ICRF-Sanger Consortium. It consists of approximately 6,000 unique genes and we hybridised our cell line samples as pairwise comparisons of normal versus transformed. This yielded a total of 61 significantly up or down-regulated genes, including examples associated with cell-matrix interactions, proliferation and transformation, as well as a number of genes associated with MYC signalling. Several of these genes have been verified as differential at the protein level on tissue microarrays of primary tumors. The transcriptional data has been prepared in concordance with MIAME (Minimal Information About Microarray Experiments) and can be viewed via the Sanger Centre via ArrayExpress (E-SNGR-8). The second study (5) was to compare purified normal luminal and myoepithelial cells, using the same platform. However, to enable comparisons to be made across a large number of different samples (we analysed 9 separate samples of each cell type) hybridization was against a common reference RNA sample, comprising a pool of breast cancer cell lines. Using SAM (Statistical Analysis of Microarray) software a total of 132 myoepthelial and 77 luminal differential genes were identified from an intrinsic gene set of approximately 1,600 genes. Using PAM (Prediction Analysis of Microarray) 33 unique genes that enabled all samples to be accurately classified as either luminal or myoepithelial were defined. Interestingly, among the myoepithelial genes there were several that were expressed in primary breast cancer by the malignant cells. These corresponded to the 'basaloid' phenotype that has been identified by other microarray studies of solid tumors. One of the myoepithelial markers had a significant adverse impact on outcome when analysed across tissue microarrays by immunohistochemistry, showing that studies of normal cells can yield valuable information about tumors. Our most recent transcriptional analysis has involved the use of MPSS. (6) This technique analyses quantitatively and qualitatively a very large number (millions) of sequences derived from mRNA samples, as a sort of superSAGE. When the HB4A cell line was analysed, for example, approximately 12,000 unique genes were identified with transcript numbers per gene ranging from less than 5 per million to over 20,000 per million, a linear range of 4 logs. MPPS analysis has now been carried out on four breast samples, the HB4A cell line and its C5.2 erbB-2 overexpressing variant, a pool of 12 normal luminal epithelial cell preparations and a pool of 15 enriched malignant cell preparations from primary breast cancers. These enormous data-sets (over 50,000 data points) are now being reviewed and analysed with the objective of identifying further novel markers of human breast cancer for subsequent pathological and biological validation.

Acknowledgements

It is obvious that the studies outlined above could not have been carried out without the collaboration and cooperation of a large number of individuals, including the many clinicians who gave us access to fresh tissues, and many members of different Ludwig Branches and organisations such as OGS and the Sanger Centre, as well as our own past and present team members. It would be invidious to name some but not others, but we would particularly like to acknowledge the general support of the Ludwig Institute for the past 10 year

Acta Oncol. Bras. Vol. 24 - nº 01, 572-575, jan-fev-mar, 2004

References

1. CLARKE C, TITLEY, J, DAVIES S & O'HARE MJ. An immunomagnetic separation method using superparamagnetic (MACS) beads for the large-scale purification of human mammary luminal and myoepithelial cells. Epithelial Cell Biol. 3, 38-46 (1994)
2. HARRIS RA, EICHOLTZ TJ, HILES ID, PAGE MJ & O'HARE, M.J. New model of ErbB-2 overexpression in human mammary luminal epithelial cells. Int. J. Cancer. 80, 477-484 (1999)
3. HARRIS RA, YANG A, STEIN RC, LUCY K, BRUSTEN L, HERATH A, PAREKH R, WATERFIELD MD, O'HARE MJ, NEVILLE AM, PAGE MJ & ZVELEBIL MJ. Cluster analysis of an extensive human breast cancer cell line protein expression map database. Proteomics 2, 212-223 (2002)
4. HERATH A, O'HARE MJ, BRUSTEN L, DAVIES S, MERRETTT S, , PAGE MJ,, PARRY S, COSSU A, TANDA F, BUDRONI M, TOWNSEND RR, NEVILLE AM, PAREKH RB & LAKHANI SR. Identification of novel prognostic markers by proteomic analysis of sporadic breast cancer. (submitted)
5. JONES C, MACKAY A, GRIGORIADIS A, COSSU A, REIS-FILHO JS, DEXTER T, DAVIES S, BULMER K, FORD E, PARRY S, TANDA F, NEVILLE AM, O'HARE MJ & LAKHANI S. Expression profiling of purified normal luminal and myoepithelial breast cells: identification of novel prognostic markers in breast cancer (submitted)
6. JONGENEEL CV, ISELI C, STEVENSON BJ, GRIGGINS GJ, LAL A, DE SOUZA SJ, MACKAY A, HARRIS RA, O'HARE MJ, NEVILLE AM, SIMPSON AJG & STRAUSBERG RJ . Comprehensive sampling of gene expression in human cell lines with massively parallel signature sequencing (MPSS). Proc. Nat. Acad. Sci. USA 100, 4702-4705 (2003)
7. O'HARE MJ., BOND, J., CLARKE, C., TAKEUCHI, Y., ATHERTON, AJ., BERRY, C., MOODY, J., SILVER ARJ., DAVIES DC, ALSOP AE., NEVILLE, AM & JAT PS. Conditional immortalization of freshly isolated normal human adult mammary fibroblasts and endothelial cells. Proc. Nat. Acad. Sci. USA 98, 646-651 (2001)
8. PAGE M, AMESS B, TOWNSEND R, PAREKH R, HERATH A, BRUSTEN L, ZVELEBIL MJ, STEIN R, WATERFIELD MD, DAVIES SC. & O'HARE MJ. Proteomic definition of normal human luminal and myoepithelial breast cells purified from reduction mammoplasties. Proc. Nat. Acad.Sci. USA 96, 12589-12594 (1999)
9. MACKAY A, JONES, C, DEXTER T, HARRIS R, LAKHANI, S, JAT PJ & O'HARE MJ. cDNA microarray analysis of genes associated with erbB-2 (HER2/neu) overexpression in human mammary luminal epithelial cells. Oncogene 22, 2680-2688, 2003.