Respirology (2007) **12**, 22–28 doi: 10.1111/j.1440-1843.2006.00957.x

INVITED REVIEW SERIES: LUNG CANCER

Proteomics analysis in lung cancer: Challenges and opportunities

TAKEFUMI KIKUCHI AND DAVID P. CARBONE

Vanderbilt Ingram Cancer Center, Nashville, Tennessee, USA

Proteomics analysis in lung cancer: Challenges and opportunities
KIKUCHI T, CARBONE DP. *Respirology* 2007; **12**: 22–28

Abstract: Recent technological developments in proteomic analysis are bringing us new insights into the molecular classification of tumours. Although proteomic analysis in cancer profiling is still under development both in terms of the instruments used and the data analytical tools, this method has great potential advantages for the analysis of biospecimens of many types. Direct measurement of abnormally expressed or modified proteins in the tumour tissue and/or patient blood may be an effective approach for discovering new biomarkers. Proteomics has the significant advantage of being able to discern not only changes in expression levels but also in post-translational modifications. Thus, the proteomics approach to protein profiling and biomarker discovery uncovers biomarkers from a different viewpoint than microarray analysis. This review summarizes the range of proteomics technologies employed for cancer profiling, and how they have been used to derive new classification models for human lung cancer.

Key words: fingerprint, lung tumour, proteomics.

INTRODUCTION

Lung cancer is the leading cause of death in the USA and all over the world.[1] While lung cancer comprises only 15% of new cancer diagnoses, it causes over 30% of all cancer deaths. It is estimated that in 2006, 190 000 people will be diagnosed with lung cancer in the USA,[2] and 170 000 patients will die from this disease. Recent developments in the early detection and treatment of cancer have resulted in improved cancer survival rates in some kinds of cancers. For example, in breast cancer, the 5-year survival rate improved from 75% in 1974 to 88% in 1995. However, the 5-year survival rate for lung cancer remains much lower than that of other malignancies, at around 15% in 1995. Thus, lung cancer is both common and extremely difficult to cure once diagnosed.

In recent years, 'omics' analyses have been developed and promise to define 'fingerprints' of crucial molecular patterns in malignant cells. These analyses derive their power from the simultaneous measurement of the expression level of multiple transcripts, protein products and/or protein modifications.

There are increasing numbers of reports using gene expression profile techniques, and perhaps the first report of possible disease outcome prediction by using microarray analysis was described in leukaemia about 10 years ago.[3] Since this report, there have been several reports of similar analyses for non-small cell lung cancer (NSCLC), describing candidate signatures predicting survival or subclassifying adenocarcinomas.[4–6]

Intrinsically, protein expression level must depend on transcription levels, but the correlation between mRNA and protein expression is poor in many cases due to vast differences in protein degradation rates and stability. The discovery of functional RNA, such as endogenous miRNA, did explain a part of this discrepancy, but all of the determinants of the amount of a given protein in a cell are not fully understood, and this is especially true in tumour cells. Although the functional players are proteins and not transcripts, microarray analysis has multiple advantages over proteomics in assay platform stability and standardization, available amplification technologies and lower complexity (no post-translational modifications), so the preponderance of convincing studies to date involve transcriptomics.

Correspondence: David P. Carbone, 2200 Pierce Avenue, 685 PRB, Vanderbilt Ingram Cancer Center, Nashville, TN 37232, USA.
Email: d.carbone@vanderbilt.edu

Received 31 July 2006; accepted 1 August 2006 (Associate Editor: Neil Watkins).

However, given the fact that almost all of Food and Drug Administration-approved cancer biomarkers are protein markers, it is increasingly important to look into global protein expression analysis in order to understand what changes are crucial steps in carcinogenesis and to make reliable predictions of cancer patient outcomes with or without clinical interventions. Improved software applications and instrument technologies are making it possible to directly measure protein complexes in cancer specimens including serum and plasma. In this review, we will describe several different kinds of approaches that appear to be feasible for uncovering protein profiles and for discovering new biomarkers capable of predicting tumour responses or patient outcomes towards the goal of individualized therapy.

One of the advantages of proteomics analysis in the cancer research field is in the materials. Microarray analysis usually requires tumour cells to obtain RNA, but it is not necessary to have direct access to tumour tissue for proteomics analysis. Various surrogate tissues, including blood and other clinical fluids from cancer patients, can be readily used for the analysis. We and others have reported proteomic analyses using lung cancer tissue, serum, plasma and pleural effusions. Considering the limited access to tumour cells within the chest, serum and plasma proteomics analysis have the great advantages, especially for the early detection of lung cancer.

TWO-DIMENSIONAL GELS

A common and standard technique for proteomic profiling analysis is two-dimensional (2-D) gel electrophoresis. This technique employs isoelectric separation according to protein charge as a first dimension coupled with MW resolution by polyacrylamide gel electrophoresis. These gels can then be stained with silver or Coomassie stains and scanned for comparison between samples.

Two-dimensional gel analysis allows us to see hundreds of proteins at the same time in a semiquantitative manner, and to identify the proteins making up each of the spots by picking up discriminating spots from the gels robotically and analysing their sequence by tandem mass spectrometric methods. In order to more accurately align these gels and quantitatively compare two different complex protein mixtures, the two protein lysates (e.g. tumour vs. normal) can be covalently labelled with two different fluorescent dyes, mixed together and run on the same 2-D gel. Thus, identical proteins from the two mixtures will precisely comigrate and can be independently detected by quantitative fluorimetry. In addition to improved alignment and normalization, these fluorescent dyes and sensitive detection methods significantly reduced the amount of protein required for 2-D gel analysis, and enabled robotic automation to pick up hundreds of spots from thin gels for subsequent identification. The widespread availability of genome and protein sequence databases enable us to identify proteins from 2-D gel analysis with mass spectrometry (MS) instruments. Since the first 2-D gel profiling analysis reported with tumour and normal tissues[7] describing differences in electrophoretic patterns, the usefulness of 2-D gel analysis advanced from pattern recognition to the discovery of candidate tumour markers. So far several reports have been published about lung cancer with 2-D gel techniques.[8–13] For example, Chen *et al.* reported the successful identification of PGK1 as a survival predictor from both microarray analysis and 2-D gel analysis using nearly 100 lung cancer patients. They also confirmed their survival prediction power by using tissue microarrays.

While 2-D gel analysis is still the most established 'gold standard' in protein profiling of complex protein mixtures, it does have significant drawbacks. It requires relatively large amounts of starting material and can only detect the most abundant proteins, which might be less important than the lower abundance proteins for cancer biomarkers.

MASS SPECTROMETRY INSTRUMENTS FOR PROTEIN PROFILING

Mass spectrometry instruments are expensive and complex, but all are made from three different basic elements: the method of ionization by which proteins are charged and converted to the gas phase, the mass analyser to separate these ions and the detector to detect them. There are two major ionization methods, matrix-assisted laser desorption ionization (MALDI), in which proteins are desorbed from the solid phase directly to the gas phase using a laser and electrospray ionization (ESI), which ionizes and vaporizes them from liquid protein solutions. The major mass analysers used most commonly in biological sample analysis are time of flight (TOF), quadrupole ion trap mass spectrometry (QIT), linear ion trap (LIT) and Fourier-Transform Ion Cyclotron Resonance (FT-ICR). Tandem mass spectrometry (MS/MS) is a combination of the same or different mass analysers, one feeding ions to the other. For example, Q-TOF is the combination of QIT and TOF. Although the MALDI-TOF and ESI-Q-TOF are most commonly used in the analysis of biological materials, each MS analysis has different capabilities and gives us different insights into proteomic features of a biological sample. Because selecting an appropriate instrument for the scientific question at hand is one of the key points for successful experiments, several groups are now undertaking a systematic evaluation of the different instruments and comparing data from other institutions and other data analytic platforms with human plasma.[14] Also MS instruments have been used for direct measurement of comprehensive profiling of other complex protein mixtures. In this approach, the primary object is not always identifying particular protein sequences, but defining patterns of protein expression from these complex protein mixtures associated with clinical features, such as histology, chemotherapy response or prognosis. Surgically resected tumour specimens, serum and plasma are the primary sources of biological material for these studies.

We will describe the MALDI-TOF instrument first, because other instruments such as ESI-Q-TOF and ESI-LIT are mainly used to obtain amino acid sequences or for shotgun proteomics approaches, to be described below.

MALDI-TOF PROFILING

The MALDI-TOF instrument ionizes proteins with a laser, accelerates the resulting ions through a high voltage tunnel and captures the ionized proteins at the detector. Since higher mass ions will fly slower than lower mass ions with the same charge, the time it takes for each ion to reach the detector will form a spectrum, which displays ion intensity in the *y*-axis, and mass–charge ratio (m/z) in *x*-axis. A MALDI-TOF instrument is widely used in a variety of experiments including direct profiling of protein complexes and amino acid sequence analysis, because of its mass accuracy, which is approximately 1 Da in 10 000. The MALDI ionization method does not require protein solubilization, but with the assistance of a matrix solution can analyse proteins directly from tissue sections. While simple and accurate, this technique suffers from variability in desorption efficiency due to salt or other tissue factors, limitation to lower MW (typically less than 50 kDa), and ion suppression artefacts from abundant species in the mixture.

In spite of these limitations, it has many advantages, and one of the major ones is the small sample size that can be analysed, down to just a few cells. Techniques have been developed to analyse cells on metallized glass slides allowing direct visualization of the cells of interest (Fig. 1).[15] This ability to acquire independent spectra from arrays of very small spots in a tissue section enables us to assemble virtual 'images' of protein distributions within the 2-D tissue samples with high resolution from frozen sections.[16] This image is akin to immunohistochemistry, but does not require an antibody or even protein identification, has a 2 order of magnitude dynamic range and can readily detect post-translational modifications.[17,18]

We have been attempting to assess the value of direct tumour tissue MALDI-TOF techniques to generate clinically useful protein expression profiles. In a proof of principle study using 66 NSCLC tumours,[19] protein profiles were capable of distinguishing normal or tumour tissue, histology, primary tumour or metastatic tumour and patient survival.

We have continued these studies using 250 NSCLC surgically resected tumours and 80 normal lung tissues. This study shows that tumour-normal discrimination was 93% accurate, prediction of occult nodal involvement was around 80% (Fig. 2).[20] These results are showing feasibility of prediction based on MALDI-TOF protein profiling analysis.

Surface-enhanced laser desorption/ionization (SELDI)-TOF is a variation of the MALDI technique. This method is a combination of protein array/preseparation and MALDI-TOF MS analysis. The solubilized sample is allowed to adsorb directly onto a variety of proprietary surfaces with different chemical characteristics (charge or hydrophobic, hydrophilic and immobilized metal affinity) or biological characteristics (antibodies, antigen binding to protein fragments, receptors). Then the matrix solution is applied to the surface and the proteins bound to the substrate

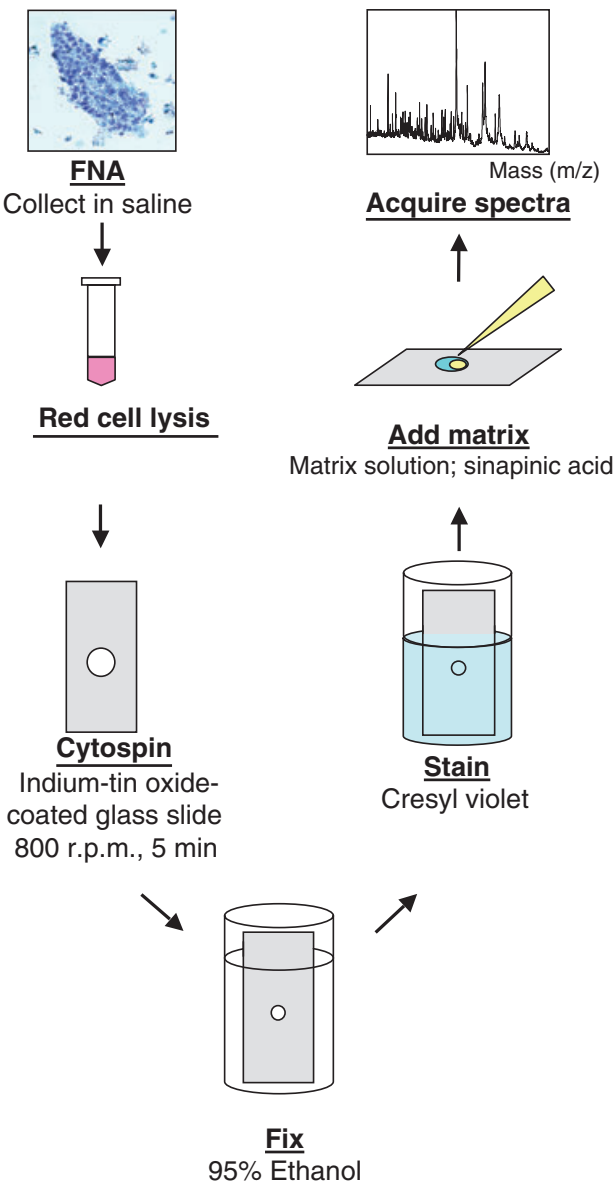

Figure 1 Principle of proteomics analysis of fine needle aspirated (FNA) sample. The cells collected from patient by fine needle aspiration (FNA) will be placed on conductive (indium-tin oxide-coated) glass slides by cytospin. After staining with cresyl violet and washed by ethanol, tumour cells can be visualized under microscopy. By spotting matrix solution (sinapnic acid) on the tumour cells directly, spectrum will be obtained by matrix-assisted laser desorption ionization time of flight instrument.

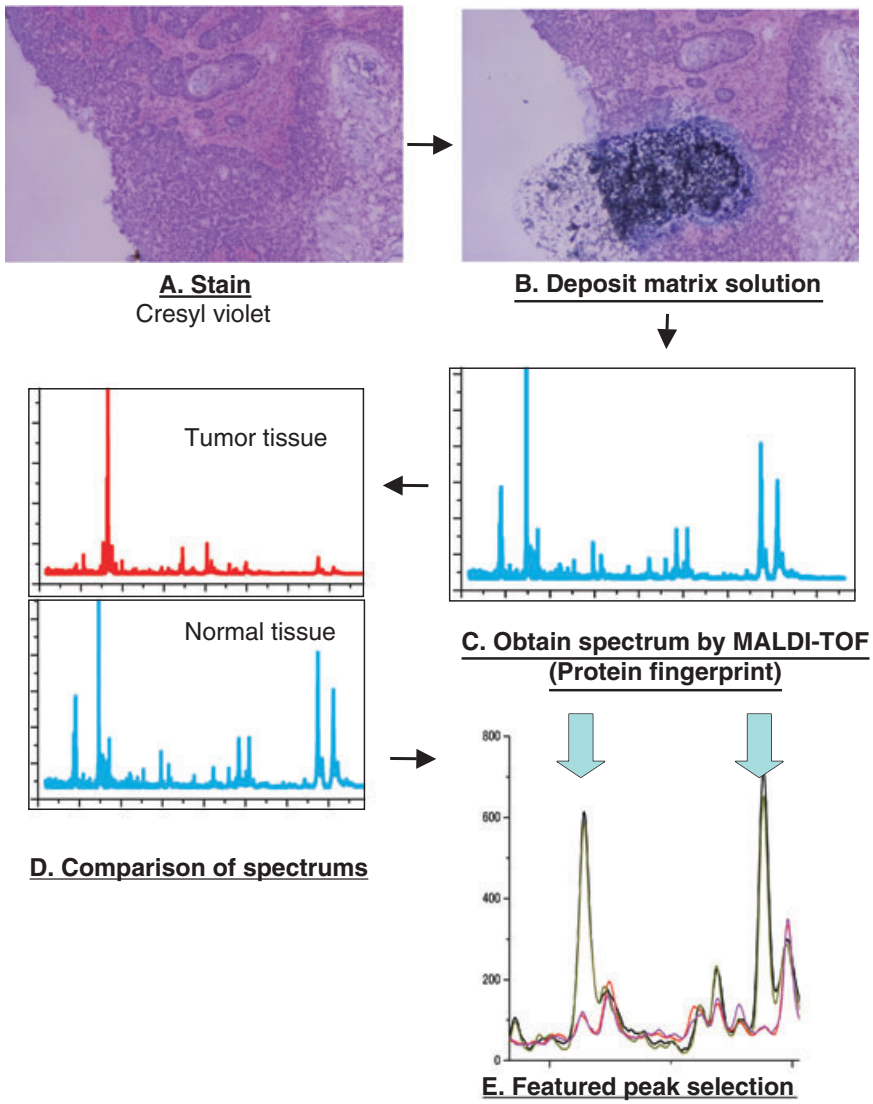

Figure 2 Tissue visualization for tissue proteomics profiling analysis. Tissue sections were placed on conductive glass slides and stained with cresyl violet. (A) Matrix solution will be placed with the aid of microscope. (B) Obtained spectra by matrix-assisted laser desorption ionization time of flight (MALDI-TOF). (C) Processing the spectrum files and comparison of spectrums. (D) Magnified view of spectrums from tumour tissue and normal tissue. Arrow heads indicate representative peaks discriminating tumour tissue from normal lung tissue. (E) These featured peaks are usually selected by statistics analysis. Using multiple significant peaks will enhance the accuracy of proteomics diagnosis on blinded sample.

analysed by MALDI-TOF. Thus, the SELDI technique is a kind of prefractionation before the laser desorption for MS analysis. The acquired spectra are then analysed in the same manner as for MALDI-TOF analysis, with peak intensity considered to represent the amount of protein. Several studies have shown the feasibility of SELDI-MS to directly analyse patient serum for protein expression pattern that has capability to distinguish cancer patients.[21,22] Mian *et al.*[23] reported that potential prediction of biological resistance of breast cancer cell line to specific antitumour reagents with SELDI protein profiling. In lung cancer proteomics, several reports have been published with SELDI proteomics profiling analysis with patient serum[24] and tumour specimen.[25]

Both in MALDI-TOF and in SELDI protein profiling analysis, demonstration of reproducibility is crucial for further clinical use and identification of proteins making up the interesting signals is indirect and difficult.

SHOTGUN PROTEOMICS

Shotgun proteomics refers to the analysis of complex protein mixtures of all kinds of materials, serum, tissue and fluids such as pleural effusion and urine. In shotgun sequencing in human genomic DNA, fragments of genome are sequenced and assembled computationally.[26] Similar to shotgun sequencing of the human genome, shotgun proteomics is the approach where peptide sequence information from MS instruments is computationally assembled to yield information on proteins present in the clinical sample. It is called 'bottom-up' proteomics because, different from 2-D gel analysis and MALDI profiling analysis, this analysis is not based on proteins, but on peptides produced by protease digestion. Protein mixtures are usually proteolytic digested by using sequence specific proteases such as trypsin, then the peptide fragments applied to an MS instrument after HPLC fractionation. Computationally, each peptide frag-

ment is then assembled and identified in protein databases to determine the proteins in the original sample. The main difficulty in this technique is that with extremely complex protein mixtures, the mixture is too complex for confident and reproducible identification by MS/MS sequencing. It should also be noted that low abundance proteins could be missed because of the presence of abundant proteins. In an attempt to overcome these problems, several separation methods are commonly used for MS analysis, size exclusion, anion exchange, strong cation exchange and reverse phase chromatography. These front-end separation techniques have been developed and successfully combined with MS instruments. Multidimensional protein identification technology (MudPIT)[27–29] is one of representative this front-end technique, which separates proteins by strong cation exchange chromatography and reverse-phase chromatography before applying protein samples to the MS instrument. In the recent report of mice organ protein profiling analysis, authors reported successful identification of 1500–2500 proteins with MudPIT separation and organellar enriched protein fractionation.[30] In cancer proteomics, Mauri *et al*.[31] reported successful identification of 42 secreted proteins in the culture media from pancreatic cancer cells.

A similar approach has been used in lung cancer pleural effusion proteomic profiling analysis.[32] In this study, the authors reported successful identification of 124 proteins from 43 pooled adenocarcinoma patient pleural effusions with a high confidence level, which means at least two or more unique peptides were found for each protein identified. Using lung cancer patient plasma, Fujii *et al*. reported establishment of automated multidimensional chromatography coupled with MS, and identified 120 proteins uniquely expressed in lung adenocarcinoma patient plasma.

Recent reports have started showing feasibility of this shotgun proteomics approach to discover biomarkers, and development of better separation techniques should allow rapid progress to be made in this area.

PROTEIN MICROARRAY

Protein microarrays are the most efficient way to analyse multiple samples or proteins at the same time. In addition, the dramatic technological developments that have enabled rapid progress in large scale gene expression array analysis will also be applicable to protein microarray analysis. There are two major types of protein microarrays, forward (FPAs) and reverse phase arrays (RPAs). In forward protein arrays, hundreds of specific antibodies are arrayed on glass slides, and directly or indirectly tagged protein lysates applied. This forward type of array gives information on the expression level or protein modifications (e.g. phosphorylation) of multiple proteins in one sample, which may shed light on specific pathways represented in the array or on patterns of expression of other cancer-associated proteins. However, the major limitation of antibody microarrays is that its specificity and accuracy is highly dependent on the availability of appropriate antibodies, and because of this, only at most 20–30% of commercially available antibodies are considered usable for antibody microarray.[33] Another problem is sensitivity, for example, one study using prostate cancer patient sera[34] reported that the detection limit of PSA was above the clinically useful range. By using a sandwich immunoassay,[35] the method showed 10-fold better sensitivity.

In RPAs, multiple tissue protein extracts or recombinant proteins are placed on glass slides and probed by a single tagged antibody or antibody mixture. Thus, this analysis evaluates the expression level of a given protein in multiple samples. For potential clinical application, Gulmann *et al*. showed using an RPA with 21 antibodies, the ability of this approach to discriminate between follicular lymphoma and follicular hyperplasia as well as predict survival.[36] However, this RPA is perhaps most interesting as a screening tool to examine protein–protein, protein–nucleic acid and protein–small molecule interactions for drug screening. One recent study demonstrated its capability to examine protein–DNA interactions and protein–antibody binding with an RPA containing p53 mutant recombinant proteins.[37] For reverse phase protein arrays using recombinant proteins a major limitation is in the production of the recombinant proteins, and one report showed only 60% of 300 proteins of interest were successfully produced and adequately purified to be useful for an RPA.[38] Methods have been developed for making recombinant proteins directly on the slides, like oligonucleotide arrays, a procedure referred to as a nucleic acid programmable protein array (NAPPA).[39] This is a very innovative method to make recombinant protein, potentially capable of high throughput and eliminating the necessity of protein purification steps.

In both forward and reverse phase protein microarrays, a good antibody is the key to producing reliable, reproducible data. Although ELISA is the most commonly used technique in clinical protein assays, innovative techniques will be required for good antibody production in the era of large scale proteomics.

SUMMARY

Human genome consortium concluded the number of genes in human genome is around 20 000–25 000.[40] Even though not all genes are expressed in cancer cells, if we consider that after translational modification such as truncation, phosphorylation and ubiquitination, the number of unique protein forms might easily exceed several hundred thousand. In most shotgun proteomics studies cited here, the number of detectable signals is around 1000–3000. It is thus apparent that we have a long way to go before we can profile cancer cells by measuring a significant fraction of the proteome.

While it is clear that proteomics approaches are yielding new insights into carcinogenesis, a potential problem in this field is still in its ability to

detect and quantify proteins. Since there is no technique to amplify proteins analogous to RT-PCR of mRNAs, the development of methods with sufficient sensitivity to measure low abundance proteins is crucial for future progress in this field. FT-ICR is one of the major recent technical advances, which, when combined with classic MS technology, allows dramatically improved mass accuracy and sensitivity and may make significant progress towards this goal.[41,42]

Even with improved instrument technology, the complexity of biological samples is so huge and the dynamic range of expression is so great, that it is likely that any conceivable technology will require some preprocessing to reduce this complexity before MS instrument analysis. MudPIT is one of the approaches to accomplish this, and another is the depletion of extremely abundant proteins such as albumin in serum analysis,[43–45] and Hb in tissue analysis.[15] A recent report shows that elimination of the six most abundant protein dramatically improved the number of signals in 2-D gel analysis of serum and plasma.[45] Unfortunately, this report also demonstrated that elimination of these specific proteins resulted in the unexpected loss of signals and the generation of artefacts. Thus, there is a trade-off of sensitivity for reproducibility and comprehensiveness.

In summary, truly comprehensive proteomic profiling analysis is still not possible today, and even the most advanced technologies available are very early in their clinical development. Each different type of analysis gives us different insights into the proteome to better understand the protein complexities in cancer cells. The clinical usefulness of proteomics analysis reported so far indicates that while a global proteomics approach is promising, future technological innovations in MS instruments, specimen preparation and bioinformatics are urgently needed to make real progress in this challenging field.

ACKNOWLEDGEMENTS

The authors are supported by Lung SPORE P50 CA90949. T.K. was founded by Uehara Memorial Foundation, Tokyo, Japan.

REFERENCES

1 Edwards BK, Brown ML, Wingo PA *et al.* Annual report to the nation on the status of cancer, 1975–2002, featuring population-based trends in cancer treatment. *J. Natl. Cancer Inst.* 2005; **97**: 1407–27.
2 Ries L, Harkins D, Krapcho M *et al. SEER Cancer Statistics Review, 1975–2003*, based on November 2005 SEER data submission, posted to the SEER web site, 2006 edition. National Cancer Institute, Bethesda, MD, 2003.
3 Golub TR, Slonim DK, Tamayo P *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999; **286**: 531–7.
4 Garber ME, Troyanskaya OG, Schluens K *et al.* Diversity of gene expression in adenocarcinoma of the lung. *Proc. Natl. Acad. Sci. USA* 2001; **98**: 13784–9.
5 Beer DG, Kardia SL, Huang CC *et al.* Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat. Med.* 2002; **8**: 816–24.
6 Bhattacharjee A, Richards WG, Staunton J *et al.* Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc. Natl. Acad. Sci. USA* 2001; **98**: 13790–5.
7 Wright GL Jr. Two-dimensional acrylamide gel electrophoresis of cancer-patient serum proteins. *Ann. Clin. Lab. Sci.* 1974; **4**: 281–93.
8 Hirano T, Franzen B, Uryu K *et al.* Detection of polypeptides associated with the histopathological differentiation of primary lung carcinoma. *Br. J. Cancer* 1995; **72**: 840–8,.
9 Bergman AC, Benjamin T, Alaiya A *et al.* Identification of gel-separated tumor marker proteins by mass spectrometry. *Electrophoresis* 2000; **21**: 679–86.
10 Oh JM, Brichory F, Puravs E *et al.* A database of protein expression in lung cancer. *Proteomics* 2001; **1**: 1303–19.
11 Chen G, Gharib TG, Huang CC *et al.* Proteomic analysis of lung adenocarcinoma: identification of a highly expressed set of proteins in tumors. *Clin. Cancer Res.* 2002; **8**: 2298–305.
12 Gharib TG, Chen G, Wang H *et al.* Proteomic analysis of cytokeratin isoforms uncovers association with survival in lung adenocarcinoma. *Neoplasia* 2002; **4**: 440–8.
13 Chen G, Gharib TG, Wang H *et al.* Protein profiles associated with survival in lung adenocarcinoma. *Proc. Natl. Acad. Sci. USA* 2003; **100**: 13537–42.
14 Omenn GS, States DJ, Adamski M *et al.* Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics* 2005; **5**: 3226–45.
15 Amann JM, Chaurand P, Gonzalez AL *et al.* Selective profiling of proteins in lung cancer cells from fine needle aspirates by MALDI-TOF mass spectrometry. *Clin. Cancer Res.* 2006; **12**: 5142–50.
16 Stoeckli M, Chaurand P, Hallahan DE, Caprioli RM. Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues. *Nat Med.* 2001; **7**: 493–6.
17 Caprioli RM. Deciphering protein molecular signatures in cancer tissues to aid in diagnosis, prognosis, and therapy. *Cancer Res.* 2005; **65**: 10642–5.
18 Chaurand P, Sanders ME, Jensen RA, Caprioli RM. Proteomics in diagnostic pathology: profiling and imaging proteins directly in tissue sections. *Am. J. Pathol.* 2004; **165**: 1057–68.
19 Yanagisawa K, Shyr Y, Xu BJ *et al.* Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* 2003; **362**: 433–9.
20 Kikuchi T, Massion PP, Shyr Y *et al.* Proteomic analysis of lung cancer tumors. *J. Clin. Oncol.* 2006; **24**: 371S.
21 Petricoin, EF 3rd, Ornstein DK, Paweletz CP *et al.* Serum proteomic patterns for detection of prostate cancer. *J. Natl Cancer Inst.* 2002; **94**: 1576–8.
22 Adam BL, Qu Y, Davis JW *et al.* Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Res.* 2002; **62**: 3609–14.

23 Mian S, Ball G, Hornbuckle J *et al.* A prototype methodology combining surface-enhanced laser desorption/ionization protein chip technology and artificial neural network algorithms to predict the chemoresponsiveness of breast cancer cell lines exposed to Paclitaxel and Doxorubicin under *in vitro* conditions. *Proteomics* 2003; **3**: 1725–37.
24 Yang SY, Xiao XY, Zhang WG *et al.* Application of serum SELDI proteomic patterns in diagnosis of lung cancer. *BMC Cancer* 2005; **5**: 83.
25 Zhukov TA, Johanson RA, Cantor AB, Clark RA, Tockman MS. Discovery of distinct protein profiles specific for lung tumors and pre-malignant lung lesions by SELDI mass spectrometry. *Lung Cancer* 2003; **40**: 267–79.
26 Venter JC, Adams MD, Myers EW *et al.* The sequence of the human genome. *Science* 2001; **291**: 1304–51.
27 Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* 2001; **19**: 242–7.
28 Chen EI, Hewel J, Felding-Habermann B, Yates JR 3rd. Large scale protein profiling by combination of protein fractionation and multidimensional protein identification technology (MudPIT). *Mol. Cell Proteomics* 2006; **5**: 53–6.
29 Jessani N, Niessen S, Wei BQ *et al.* A streamlined platform for high-content functional proteomics of primary human specimens. *Nat. Methods* 2005; **2**: 691–7.
30 Kislinger T, Cox B, Kannan A *et al.* Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. *Cell* 2006; **125**: 173–86.
31 Mauri P, Scarpa A, Nascimbeni AC *et al.* Identification of proteins released by pancreatic cancer cells by multidimensional protein identification technology: a strategy for identification of novel cancer markers. *FASEB J.* 2005; **19**: 1125–7.
32 Tyan YC, Wu HY, Lai WW, Su WC, Liao PC. Proteomic profiling of human pleural effusion using two-dimensional nano liquid chromatography tandem mass spectrometry. *J. Proteome Res.* 2005; **4**: 1274–86.
33 Haab BB, Dunham MJ, Brown PO. Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biol*. 2001, **2**: R4.
34 Miller JC, Zhou H, Kwekel J *et al.* Antibody microarray profiling of human prostate cancer sera: antibody screening and identification of potential biomarkers. *Proteomics* 2003; **3**: 56–63.
35 Nielsen UB, Cardone MH, Sinskey AJ, MacBeath G, Sorger PK. Profiling receptor tyrosine kinase activation by using Ab microarrays. *Proc. Natl. Acad. Sci. USA* 2003; **100**: 9330–5.
36 Gulmann C, Espina V, Petricoin E 3rd *et al.* Proteomic analysis of apoptotic pathways reveals prognostic factors in follicular lymphoma. *Clin. Cancer Res.* 2005; **11**: 5847–55.
37 Boutell JM, Hart DJ, Godber BL, Kozlowski RZ, Blackburn JM. Functional protein microarrays for parallel characterisation of p53 mutants. *Proteomics* 2004; **4**: 1950–8.
38 Braun P, Hu Y, Shen B *et al.* Proteome-scale purification of human proteins from bacteria. *Proc. Natl. Acad. Sci. USA* 2002; **99**: 2654–9.
39 Ramachandran N, Hainsworth E, Bhullar B *et al.* Self-assembling protein microarrays. *Science* 2004; **305**: 86–90.
40 International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* 2004; **431**: 931–45.
41 Reid GE, McLuckey SA. 'Top down' protein characterization via tandem mass spectrometry. *J. Mass Spectrom.* 2002; **37**: 663–75.
42 Frahm JL, Howard BE, Heber S, Muddiman DC. Accessible proteomics space and its implications for peak capacity for zero-, one- and two-dimensional separations coupled with FT-ICR and TOF mass spectrometry. *J. Mass Spectrom.* 2006; **41**: 281–8.
43 Govorukhina NI, Reijmers TH, Nyangoma SO, van der Zee AG, Jansen RC, Bischoff R. Analysis of human serum by liquid chromatography-mass spectrometry: improved sample preparation and data analysis. *J. Chromatogr. A* 2006; **1120**: 142–150.
44 Sun W, Wu S, Wang X, Zheng D, Gao Y. An analysis of protein abundance suppression in data dependent liquid chromatography and tandem mass spectrometry with tryptic peptide mixtures of five known proteins. *Eur. J. Mass Spectrom. (Chichester, Eng)* 2005; **11**: 575–80.
45 Echan LA, Tang HY, Ali-Khan N, Lee K, Speicher DW. Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics* 2005; **5**: 3292–303.