Oncogene (2006) 25, 3810–3817
© 2006 Nature Publishing Group All rights reserved 0950-9232/06 $30.00
www.nature.com/onc

REVIEW

Proteomics for the early detection and treatment of hepatocellular carcinoma

JT Feng, S Shang and L Beretta

Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

The prognosis for hepatocellular carcinoma (HCC) is poor and has not improved in recent years, largely owing to lack of early diagnosis, frequent recurrence after surgery and resistance to chemotherapy. Proteomics holds the promise of improving our understanding of HCC carcinogenesis and progression as well as of discovering novel diagnostics and therapeutics. Proteomic analyses of HCC cell lines, animal models and serum and tumor tissue from patients with HCC have been performed to date. Proteomic technologies have greatly improved in the past few years as reviewed here. It is anticipated that with the recent development of protein tagging, protein separation methods and mass spectrometry sensitivity, proteomic studies of HCC will allow the identification of diagnostic and prognostic biomarkers as well as therapeutic targets, which could greatly improve the clinical management of HCC patients.
Oncogene (2006) **25,** 3810–3817. doi:10.1038/sj.onc.1209551

Keywords: proteomics; hepatocellular carcinoma; biomarkers; diagnostics; therapeutics

Introduction

Hepatocellular carcinoma (HCC) is the third leading cause of cancer death worldwide (Parkin *et al*., 2001) and the incidence of HCC is increasing in many countries such as France, the United Kingdom and the United States (Taylor-Robinson *et al*., 1997; Deuffic *et al*., 1998; El-Serag and Mason, 1999). Surgery, be it surgical resection or liver transplantation, is to date the only hope for cure because chemotherapeutic intervention appears to be ineffective against HCC. However, less than 15% of patients are currently undergoing surgery because of late clinical presentation and, subsequently, diagnosis. Therefore, the overall 5-year survival rate of HCC remains less than 5% (El-Serag *et al*., 2001). There is an urgent need to develop new therapeutic targets and to identify novel biomarkers for the early detection of HCC before any breakthrough in HCC clinical management can be achieved. Proteomics studies the complete set of proteins expressed in a given cell, tissue or biofluid. Proteomics not only characterizes protein expression profiles but also identifies protein structures, localizations, activities, modifications and interactions in physiological or pathological states. As proteins perform most biological functions, proteomics bridges the gap between the information coded in the genome sequence and cellular behavior. Proteomics studies of HCC may not only elucidate the mechanisms of HCC initiation and progression, but may also have the potential to discover novel diagnostic and prognostic biomarkers as well as therapeutic targets. This review summarizes the proteomic studies of HCC reported to date as well as the recent developments in proteomic technologies.

Correspondence: Dr L Beretta, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, M5-A864, PO Box 19024, Seattle, WA 98109 1024, USA.
E-mail: lberetta@fhcrc.org

Proteomic technologies

No matter which approach is used, the successive steps for proteomic studies include protein preparation, protein separation, protein identification and proteome informatics. The most commonly used proteomic technologies, including those applied to the study of HCC, are briefly described here and summarized in Figure 1.

Protein preparation

It is critical to understand how pre-analytical variables affect proteomic results and, therefore, better understanding of these variables should lead to improved experimental designs and interpretation of data. Sample preparation involves efficient and effective solubilization of proteins. In addition, because of the wide range of protein abundance and of protein properties present in a protein population, it is advantageous to work with enriched sub-proteomes. Sub-proteomic approaches rely on enrichment techniques for isolation of proteins with similar biochemical characteristics, with similar function or within a specific cellular compartment. In most cases, conventional methods (e.g., centrifugation of sucrose gradients for isolating organelles) are used for sample isolation. The tandem affinity purification approach is also widely used to purify multi-protein complexes (Rigaut *et al*., 1999; Puig *et al*., 2001).

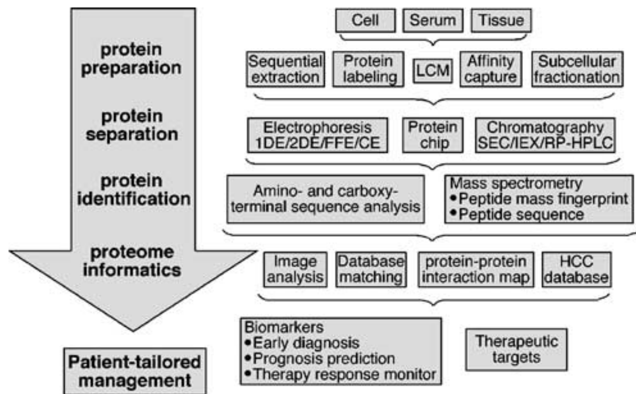

Figure 1 Schematic representation of the most commonly used proteomic approaches. Abbreviations: LCM, laser capture microdissection; 1DE, one-dimensional gel electrophoresis; 2D-PAGE, two-dimensional polyacrylamide gel electrophoresis; SEC, size-exclusion chromatography; IEX, ion-exchange chromatography; RP-HPLC, reversed-phase high-performance liquid chromatography; FFE, free-flow electrophoresis; CE, capillary electrophoresis.

In order to compare protein expression profiles between two different samples quantitatively, the mass of proteins in one cellular state must be modified so that it is unique from the mass of proteins from another cellular state. Based on the method of quantitative tag incorporation, the protein-labeling techniques can be broadly classified into two groups. One group of techniques is biological labeling, where labeling of the peptide/protein is achieved by growing cells in media enriched in stable isotope-containing anabolites. The representative techniques include stable isotope labeling by amino acids in cell culture (SILAC) (Ong *et al*., 2002) and further evolved into amino-acid-coded mass tags (AACTs) in which isotope-labeled amino acids are used as tag precursors (Gu *et al*., 2004). The other group of techniques is chemical labeling of harvested proteins. The most representative techniques of this group are isotope-coded affinity tagging (ICAT) (Gygi *et al*., 1999) and mass-coded abundance tagging (Cagney and Emili, 2002). The pitfalls of chemical labeling include the relatively low labeling efficiency and the limited abundance of target residues.

Protein separation

The next step after sample preparation is separation of proteins. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) has been so far the most common method for separating complex mixtures of proteins and certainly the most widely used approach in proteomic studies of HCC. Following protein separation based on isoelectric point (p*I*) in the first dimension and molecular weight in the second dimension, 2D-PAGE enables the simultaneous visualization of thousands of protein spots, their semiquantification and the detection of post-translational protein modifications. Although 2D-PAGE provides reasonable protein separation and resolution at a low cost, this method is insufficient for visualizing all protein species (e.g., membrane proteins or basic proteins, and proteins in low abundance). This methodology has recently been improved with regard to reproducibility, resolution and sensitivity. In particular, progress has been made by using ultra-narrow pH-range IPG strips, extreme p*I* IPG strips, multi-gel systems and two-dimensional difference gel electrophoresis (2D-DIGE) (Unlu *et al*., 1997). In 2D-DIGE, protein samples are first labeled using fluorescent dyes such as cyanine (Cy2, Cy3 or Cy5) dyes or Alexa dyes, then mixed equally and subjected to 2D separation. After scanning at different emission wavelengths, multiple images corresponding to different samples are generated from one 2D gel. Two-dimensional difference gel electrophoresis reduces the gel-to-gel variability associated with standard 2D-PAGE and improves accuracy in protein semiquantitation (Van den Bergh and Arckens, 2004).

Other methods employed to separate proteins are high-performance liquid chromatography (HPLC) or two-dimensional liquid chromatography. Combining liquid chromatography (LC) separation and mass spectrometry (MS), the ‘shotgun’ strategy is a gel-free approach in which protein mixtures are digested into peptides, resolved by LC, identified by tandem mass spectrometry (MS/MS) and data-matched by computer algorithms to determine the original content of the mixture (Yates, 1998). Compared with 2D-PAGE, the shotgun strategy has higher sensitivity, higher sample loading capacity and is more amenable to automation. Complex samples containing hundreds of proteins can be sequenced during a single analysis. This approach can be used to characterize a specific cell state by collating lists of identified peptides (cataloguing proteomics), by enumerating differences in peptide composition between samples (subtractive proteomics) or by comparing protein profiles between cell states using stable isotope labeling (quantitative proteomics). When more complex mixtures are analysed, additional separation is required for maximal protein coverage (Wang and Hanash, 2003, 2005). Chromatography separations such as size exclusion, anion exchange, strong cation exchange and reversed phase could be combined and followed with MS to create multi-dimensional protein identification technology (Wolters *et al*., 2001).

Protein arrays have the potential to revolutionize proteomic studies as did DNA microarrays in genomic studies. In the microarray format, protein chips promise high-throughput analysis of very small samples. Protein chips can be used to study protein expression profiles, protein–protein interactions or characterization of post-translational modifications (Liotta and Petricoin, 2000; MacBeath, 2002; Zhu and Snyder, 2003). The development of protein chips is slow, however, because of the complex nature of proteins. Challenges in the fabrication and application of protein chips include efficient protein immobilization, retention of natural epitopes of polypeptides, printing techniques with high spatial resolution, availability of large and diverse sets of proteins or antibodies, probe design and labeling, and development of novel algorithms for data analysis. The

formats of protein chips are diverse. These include macro- and microarrays, microwell chips, microfluidic chips, as well as alternative formats, such as surface-enhanced laser desorption/ionization (SELDI) and surface plasmon resonance. As the most widely used protein chip format, SELDI combines purification of samples on a wide variety of affinity matrices and identification by time-of-flight mass spectrometry (TOF-MS). By TOF-MS, the retained proteins on the modified surface of the array will generate specific signature patterns, which can be compared to distinguish different samples.

Protein identification

Mass spectrometry is the current method of choice for the identification of proteins, as this method offers high analytical sensitivity and the capacity for high-throughput protein identification. Proteins are first digested with proteases or chemical agents to produce a mixture of peptide fragments. Peptides are then ionized using electrospray ionization (ESI) (Fenn *et al*., 1989) or matrix-assisted laser desorption/ionization (MALDI) (Karas and Hillenkamp, 1988) and analysed by various mass spectrometers. The mass-to-charge (m/z) ratio of these peptide fragments is then determined, which generates a peptide mass fingerprint (PMF). The PMFs from the proteins of interest are then compared to theoretical PMFs of known and DNA-sequence-derived proteins contained within protein databases. Additionally, MS can be employed to obtain amino-acid sequence data to support PMF identification in situations where it is sometimes difficult to assign an unequivocal identity to a protein based on PMF alone or for novel protein identification. Tandem mass spectrometry (MS/MS), which is capable of providing amino-acid sequence information on peptide fragments of the parent protein, is the connection of two mass spectrometers in series. In the application of MS/MS, the first spectrometer generates the PMF spectra; the peptide fragment of interest is selected and isolated by an ion gate voltage that excludes all other fragments but allows the passage of the selected peptide fragment of interest into the second spectrometer where it is fragmented. There are many types of mass spectrometer that can be coupled to MALDI or ESI to generate tandem mass spectra that can be utilized in shotgun proteomic analysis, including quadrupole ion traps, linear ion traps, time-of-flight/time-of-flight and quadrupole time-of-flight (TOF) instruments (Yates, 2004). An additional MS innovation for shotgun proteomics is the recent development of a linear ion trap quadrupole Fourier transform mass spectrometer (LTQ-FT-MS) (Syka *et al*., 2004). This instrument couples ion trap technology with the accurate mass and high-resolution capabilities of an FTMS. The ability to accumulate ions via trap technology and then accurately distinguish small differences in mass will permit improved proteomic informatics. Individual MS/MS spectra are commonly interpreted using search routines such as SEQUEST (Eng *et al*., 1994) or Mascot (Perkins *et al*., 1999). Peptide mass information in conjunction with peptide sequence information obtained from MS/MS analysis provides a strong basis for protein identification.

Proteome informatics

With the accumulation of large sets of proteomic data, proteome informatics becomes a critical aspect of proteomic studies. Both the technologies and experimental designs used to generate and analyse data are becoming increasingly complex. The need for methods by which such data can be accurately described, stored and exchanged between experimenters and data repositories has been recognized. Work by the Proteome Standards Initiative of the Human Proteome Organization (HUPO) has laid the foundation for the development of standards by which experimental designs can be described and data exchange facilitated (Orchard *et al*., 2004). An open generic XML (extensible markup language) representation of MS data, named mzXML, was introduced by Pedrioli *et al*. to facilitate data management, interpretation and data dissemination in proteomic research using different instrumentation platforms (Pedrioli *et al*., 2004).

Hepatocellular carcinoma analysis using proteomics

Cell lines

Because of their relative homogeneity when compared to tumor tissues, HCC cell lines have been extensively used as *in vitro* models for proteomic studies of HCC. This approach was partially validated by comparing the protein expression profiles of normal liver tissue, untransformed liver cell lines and HCC cell lines using 2D-PAGE-MS (Wirth *et al*., 1995). In a more recent study using a 2D-DIGE-MS approach, the protein expression profiles of primary cultured hepatocytes and of HCC cell lines were compared (Fujii *et al*., 2005); in this study, 44 out of 1238 protein spots detected were differentially expressed in HCC cell lines. Proteomic expression studies using a variety of approaches such as 2D-PAGE-MS, 2D-DIGE-MS, AACT-1D PAGE-MALDI-TOF/TOF-MS/MS and 1D PAGE-ICAT-LC-ESI-MS/MS have also been performed to compare protein profiles of alpha-fetoprotein (AFP)-producing and non-AFP-producing HCC cell lines (Yokoo *et al*., 2004), protein profiles of HCC cell lines with different metastatic potentials (Cui *et al*., 2004; Ding *et al*., 2004; Yu *et al*., 2004; Shui *et al*., 2005) or protein profiles of cholangiocarcinoma and HCC cell lines (Srisomsap *et al*., 2004). As a result, 11 proteins involved in apoptosis, glucose metabolism and cytoskeletal organization have been identified as differentially expressed between AFP-producing and non-AFP-producing cell lines. Cytokeratins 7 and 19, U2/2 and galectin-3 expression distinguish cholangiocarcinoma and HCC cell lines and several proteins (e.g., cytokeratin 19, annexin 1, S100 calcium-binding protein A4, cathepsin D, calreticulin precursor) distinguish between HCC cell

lines with high and low metastatic potentials. The secretome of HCC cell lines represents a potential rich source of biomarkers. A novel method combining metabolic labeling of proteins and 2D-PAGE-MS was recently used to characterize the secretome of the HCC cell line HepG2 (Zwickl *et al*., 2005).

Cell lines are also good models to study the effects of modulating the expression of a specific protein or the activity of a specific pathway. This approach was used in particular to investigate the effects of expressing hepatitis B virus (HBV) HBx protein (Tan and Chen, 2005) and *Helicobacter pylori* (Zhang *et al*., 2005) on protein profiles. Using a proteomic approach, it was also reported that hepatitis C virus (HCV) core interacts with microfilament proteins such as cytokeratin and vimentin (Kang *et al*., 2005). Our group recently used two different proteomic approaches in order to characterize modifications in the lipid raft proteome upon HCV replication in Huh7 cell lines (Mannova and Beretta, 2005). Using 2D gel separation of lipid raft proteins and Q-TOF-MS, we identified upregulation in lipid raft fractions from HCV-replicating cells of proteins involved in cell signaling, protein trafficking, vesicle formation and transport. Using Western blot and immunofluorescence assays, we demonstrated that upregulation of the majority of proteins was owing to their subcellular redistribution into lipid rafts in HCV-replicating cells. The second proteomic approach was based on recently emerged techniques that allow quantification and analysis of a complex protein mixture after fractionation by HPLC or sodium dodecyl sulfate (SDS)–PAGE. We used a SILAC approach to label proteins differentially in lipid rafts in HCV replicon and control cell lines. Detergent-insoluble fractions were analysed using LTQ-FT-MS/MS after SDS–PAGE separation and in-gel trypsin digestion. This approach allowed us to identify and quantitate approximately 500 proteins, including both proteins in low abundance and hydrophobic proteins, which were not detected by 2-D gel separation.

Expression maps of proteins expressed in hepatoma cell lines have been constructed (Lee *et al*., 2003; http://proteome.btc.nus.edu.sg/hccm/). These maps will facilitate integration of studies performed in different groups worldwide.

Tumor tissues

To date, clinical tissue samples have been the most extensively studied samples in HCC proteomic studies and most studies compared protein expression profiles between tumor tissue and adjacent non-tumor tissue (Kim *et al*., 2002; Lim *et al*., 2002; Takashima *et al*., 2003, 2005; Yokoyama *et al*., 2004; Zeindl-Eberhart *et al*., 2004; Li *et al*., 2005). One study compared the protein profiles among HBV+, HCV+ and HBV−/HCV− HCC tumor tissues using 2D-PAGE-MS and found that among the 60 differentially expressed proteins when comparing tumor and non-tumor tissues, 46 proteins were specific to the underlying viral etiology (Kim *et al*., 2003). This study, therefore, confirmed previous observations from our group (using DNA microarrays) that the mechanisms of hepatocarcinogenesis vary according to the underlying liver disease. In a study to investigate the molecular basis for iron depletion in human HCC, 19 paired HCC tumor and adjacent non-tumor tissue samples were compared using 2D-PAGE-MS. The expression of an iron-storage protein, tissue ferritin light chain, was identified as severely reduced, whereas the corresponding transcript was not changed (Park *et al*., 2002b). Our group has applied proteomic tools to the comparative analysis of protein profiles between HCC and adjacent non-tumor liver tissues as a means of discovering novel molecular markers (Chignard *et al*., 2006). Forty-seven protein spots that showed reproducible variation were identified by MS, corresponding to proteins encoded by 23 distinct genes. A positive correlation between transcript and protein level variations was observed for only seven out of the 23 genes. Proteolytic cleavage accounted for the discrepancies between mRNA and protein level changes for seven genes, including calreticulin, PDIA3, PDI and GRP78. We detected at least one fragment of each of these four endoplasmic reticulum proteins in the culture supernatant of the PLC-PRF5 hepatoma cell line, suggesting that their cleavage leads to the release of selected cleaved products into the extracellular compartment. We also detected calreticulin and PDIA3 cleavage products in the sera of patients with HCC, and a statistically highly significant difference was observed in the levels of calreticulin and PDIA3 fragments in the sera of patients with HCC when compared with the sera of healthy individuals. Amounts of calreticulin and PDIA3 fragments were also significantly different between patients with HCC and at-risk patients (patients with chronic hepatitis and patients with cirrhosis). This study suggests that specific isoforms in general and cleavage products in particular should be further evaluated as new markers for HCC.

Some studies used laser capture microdissection (LCM) in order to characterize isolated tumor cell populations from heterogeneous tissue sections. By combining LCM with 2D-DIGE-MS for more accurate comparison, Liang found that the protein profiles of well- and poorly differentiated HCC tissues, all of which exhibited HBV infection, are significantly different, highlighting in particular metabolic enzymes involved in the methylation cycle (Liang *et al*., 2005). Laser capture microdissection coupled with ICAT and 2D-LC-MS was also used for quantitative analysis of tumor and non-tumor cells (Li *et al*., 2004). The protein expression profiles of different regions in one tumor node are different, as demonstrated by a comparison of protein profiles between central tumor and peripheral tumor tissues by combining LCM with SELDI-TOF-MS (Melle *et al*., 2004). This study showed that the heterogeneity of HCC is manifest not only between patients, but also in one tumor node of the same patient, which suggests that different tumor cell clones may be growing in one node at the same time and that microenvironmental factors (e.g., hypoxia, host responses) may play critical roles in the differentiation

and progression of HCC tumor cells. Another study that combined LCM with protein microarrays identified 32 out of 83 tested proteins expressed differentially between purified HCC and normal liver tissue; in this study, cyclin D1 and SOCS1 were further validated, using Western blot and tissue microarrays, as associated with tumor prognosis (Tannapfel *et al*., 2003).

An HCC proteome database (http://yprcpdb.proteomix.org/) which stores 2D-PAGE images and spot identifications, together with the corresponding clinical information associated with the samples, has been recently established (Cho *et al*., 2002). These efforts in database generation will greatly facilitate integration of studies performed in different groups worldwide.

Sera

Alpha-fetoprotein is the only HCC serum diagnostic biomarker in common use, although its sensitivity is only 50–70% (Daniele *et al*., 2004). Therefore, the discovery of novel HCC serum biomarkers with greater sensitivity and specificity remains a major challenge for the early detection of HCC. As serum has high concentrations of proteins and includes thousands of proteins that are simply reflective of physiological or pathological processes in the human organism, analysis of serum samples is still a daunting task for proteomic studies. In addition, 22 of the most abundant proteins account for >99% of total serum proteins (Anderson and Anderson, 2002). To detect the disease-associated proteins present in very low abundance using currently available technologies, the most abundant proteins have to be removed first, by techniques such as immunodepletion. Our group is using a combination of immunodepletion of the six most abundant proteins (albumin, IgG, IgA, antitrypsin, transferrin and haptoglobin), fluorophore labeling, 3D protein separation (on-line anion exchange and reversed-phase LC separation, followed by 1D electrophoresis) to identify differential protein expression in serum from patients with HCC and patients with liver cirrhosis (Figure 2). This approach uniquely allows the quantification as well as the identification of specific isoforms such as cleavage products. Analysis of serum by 2D-PAGE-MS following immunodepletion of albumin and IgG has also been performed, identifying heat-shock protein 27 as a potential HCC biomarker (Feng *et al*., 2005).

Highly abundant proteins are known to function as carriers and transporters of other proteins within the blood, and therefore their depletion may result in the loss of potentially important proteins (Mehta *et al*., 2003). Alpha-fetoprotein was found to be retained in the albumin/IgG-containing fraction (Feng *et al*., 2005). By using highly sensitive MS, our group has identified 83 proteins bound to the six most abundant proteins (unpublished data). Glycoproteomics is also a promising approach in the effort to reduce the complexity of the serum proteome for the identification of novel biomarkers (Comunale *et al*., 2004; Block *et al*., 2005).

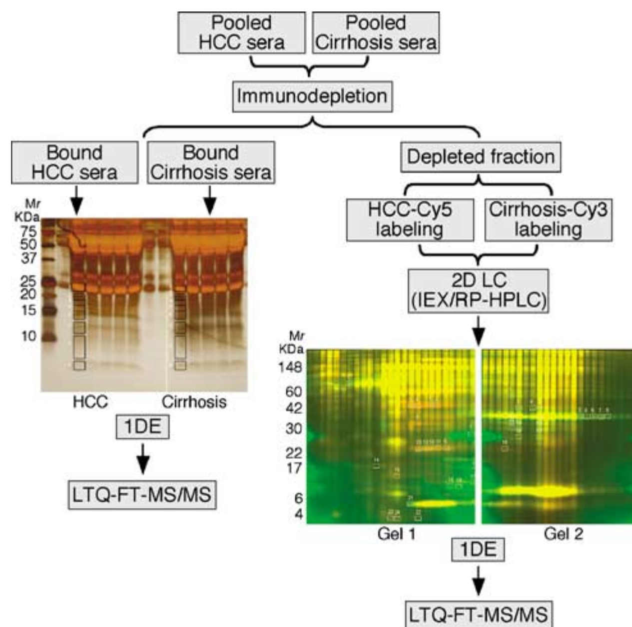

Figure 2 Flowchart of the multi-dimensional serum protein identification technology adopted by our group for the discovery of hepatocellular carcinoma biomarkers. After immunodepletion of the six most abundant proteins, pooled serum samples labeled with fluorophore dyes were subjected to on-line two-dimensional liquid chromatography followed by SDS-PAGE for separation and LTQ-FT-MS/MS for protein identification. Proteins bound to the six removed proteins were also separated and identified by LTQ-FT-MS/MS.

SELDI-TOF-MS has also been extensively used in serum proteomics studies; using this approach, proteomic features that can distinguish between HCC and chronic liver diseases with high sensitivity and specificity and proteomic features that can further divide HCC with lymph node invasion or distant metastasis were identified (Poon *et al*., 2003; Paradis *et al*., 2005; Schwegler *et al*., 2005). One of the most discriminating features was identified as the C-terminal fragment of vitronectin, which may be cleaved by metalloproteases that are enhanced in HCC (Paradis *et al*., 2005).

There is increasing evidence for an immune response to cancer in humans, demonstrated in part by the identification of autoantibodies to tumor antigens (Shalhoub *et al*., 2001). The identification of panels of tumor antigens that elicit a humoral response may have utility in cancer screening, diagnosis or prognosis. Several approaches are currently available for the identification of tumor antigens. We have used a proteomics-based approach for the identification of tumor antigens that induce an antibody response in HCC. Proteins eliciting a humoral response in HCC were identified by 2D Western blotting using sera from patients with HCC, followed by MS analysis and database searches. The proteins identified were members of diverse groups including chaperones, structural proteins and enzymes (Le Naour *et al*., 2002). In particular, antibodies against epitopes at the C-terminal end of calreticulin exhibited the highest frequency of

autoantibodies in sera from patients with HCC. We now have evidence that the presence of these autoantibodies is a consequence of a specific cleavage of calreticulin in HCC (unpublished data).

Animal models

Animal models of HCC have been extensively studied. These models may have utility for the study of the mechanisms of carcinogens in HCC initiation and progression. To investigate proteins involved in peroxisome proliferation, Chu *et al*. (2002) compared the protein profiles of acyl-CoA oxidase-deficient and peroxisome proliferator-treated mouse liver with SELDI-TOF-MS. With 2D-PAGE-MS, Zeindl-Eberhart *et al*. (2001, 2004) compared the protein profiles of rat HCC tissues induced by genotoxic or non-genotoxic carcinogens and identified aldose reductase-like variants among HCCs induced by carcinogens of different categories. Protein profiles of mouse HCCs induced by two different drugs of the same non-genotoxic category were distinct, demonstrating the great heterogeneity of pathways activated in HCC (Iida *et al*., 2003). Differentially expressed proteins in animal models can possibly be used as potential biomarkers for HCC diagnosis or targets for HCC chemotherapy as demonstrated in Zeindl-Eberhart's study; variants of aldose reductase-like proteins, which were found differentially expressed in rat HCC, were further validated as potential biomarkers for human HCC by immunohistology (Zeindl-Eberhart *et al*., 2004). A 2D LC and capillary electrophoresis approach coupled with MALDI-TOF-TOF-MS was used to identify approximately 300 proteins in tumor tissue from a high metastatic HCC model in nude mice (Zhang *et al*., 2004).

Concluding remarks and future directions

To date, both proteomic and genomic studies have highlighted the heterogeneity of HCC. This heterogeneity emphasizes the complex nature of HCC carcinogenesis and disease progression in which multiple pathogenesis mechanisms seem to be involved (Chignard and Beretta, 2004; Suriawinata and Xu, 2004; Parent and Beretta, 2005; Zhang and Ji, 2005). Proteomics can provide a molecular characterization of HCC, which may allow for individualized molecular therapy. Laboratory tests used in the clinical setting are targeted at expressed proteins. Therapeutics are now being developed that work at the level of the proteome, such as targeted small molecules and recombinant humanized monoclonal antibodies. In addition, proteomic studies of HCC to date have highlighted that some information in the proteome cannot be predicted simply from its related nucleic acid sequence. A number of alterations or modifications can occur at translational and post-translational levels to affect protein function. Several proteomic studies of HCC, including ours, have found that changes in protein expression often do not correlate with changes at the corresponding transcript level. In particular, protein cleavage seems to play an important role in HCC. Many protein cleavage fragments, including the N-terminal end of glypican-3 (Hippo *et al*., 2004), the C-terminal end of vitronectin (Paradis *et al*., 2005), a fragment of complement C3 (Steel *et al*., 2003) and the N-terminal ends of calreticulin and of PDIA3 (Chignard and Beretta, 2004), have been reported to be specific to HCC. Also reported in HCC were specific isoforms of apolipoprotein A1 (Steel *et al*., 2003; Fernandez-Irigoyen *et al*., 2005), variants of aldehyde dehydrogenase isozymes (Park *et al*., 2002a) and hyperfucosylation of Golgi Protein 73 (Block *et al*., 2005). It is estimated that there are more than 100 different post-translational modifications (e.g., glycosylation, phosphorylation, oxidation) (Krishna and Wold, 1988; Yang, 2005), although only a few have yet been studied in HCC.

Although proteomic studies of HCC are still in their infancy, the pioneering studies reviewed herein have proven that proteomics is beginning to impact human disease diagnosis and therapeutic intervention. Further work is required to enhance the performance and reproducibility of established proteomic tools, and issues regarding pre-analytical variables, analytical variability and biological variation must be addressed if further progress is to be made. In addition, the revolution in clinical management of HCC by proteomics will require close collaboration and cooperation among researchers, clinicians, biostatisticians and engineers. Fortunately, the Human Liver Proteome Project (HLPP) was launched by HUPO in 2002 (Hanash and Celis, 2002; Hanash, 2004) and, to date, HLPP members have identified about 5000 proteins in normal human liver tissue and many standard operating procedures have been established (He, 2005; Parent and Beretta, 2005).

Because of the heterogeneity of HCC, it is likely that proteomic patterns or panels of biomarkers will be identified for early detection of HCC. To achieve higher efficacy with lower toxicity, individualized therapeutic combinations that best target the patient's specific protein signature should prevail over present HCC therapies that are directed at a single molecular target. In the near future, there is the potential to apply proteomic techniques to real-time assessment that will permit redirection of therapy based on changes in a patient's protein signature that may indicate drug resistance. With the rapid development of proteomics and its extensive application in the study of HCC, clinical management of HCC is poised to benefit enormously, with the potential to develop better diagnostic and prognostic tests, to identify new therapeutic agents and ultimately to allow individualized therapy.

Acknowledgements

We are grateful to Paul Farley for his expert editorial assistance.

References

Anderson NL, Anderson NG. (2002). *Mol Cell Proteomics* **1**: 845–867.

Block TM, Comunale MA, Lowman M, Steel LF, Romano PR, Fimmel C *et al*. (2005). *Proc Natl Acad Sci USA* **102**: 779–784.

Cagney G, Emili A. (2002). *Nat Biotechnol* **20**: 163–170.

Chignard N, Beretta L. (2004). *Gastroenterology* **127**: S120–S125.

Chignard N, Shang S, Wang H, Marrero J, Bréchot C, Hanash S *et al*. (2006). *Gastroerterology* in press.

Cho SY, Park KS, Shim JE, Kwon MS, Joo KH, Lee WS *et al*. (2002). *Proteomics* **2**: 1104–1113.

Chu R, Zhang W, Lim H, Yeldandi AV, Herring C, Brumfield L *et al*. (2002). *Gene Expr* **10**: 165–177.

Comunale MA, Mattu TS, Lowman MA, Evans AA, London WT, Semmes OJ *et al*. (2004). *Proteomics* **4**: 826–838.

Cui JF, Liu YK, Pan BS, Song HY, Zhang Y, Sun RX *et al*. (2004). *J Cancer Res Clin Oncol* **130**: 615–622.

Daniele B, Bencivenga A, Megna AS, Tinessa V. (2004). *Gastroenterology* **127**: S108–S112.

Deuffic S, Poynard T, Buffat L, Valleron AJ. (1998). *Lancet* **351**: 214–215.

Ding SJ, Li Y, Shao XX, Zhou H, Zeng R, Tang ZY *et al*. (2004). *Proteomics* **4**: 982–994.

El-Serag HB, Mason AC. (1999). *N Engl J Med* **340**: 745–750.

El-Serag HB, Mason AC, Key C. (2001). *Hepatology* **33**: 62–65.

Eng JK, McCormack AL, Yates III JR. (1994). *J Am Soc Mass Spectrom* **5**: 976–989.

Feng JT, Liu YK, Song HY, Dai Z, Qin LX, Almofti MR *et al*. (2005). *Proteomics* **5**: 4581–4588.

Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. (1989). *Science* **246**: 64–71.

Fernandez-Irigoyen J, Santamaria E, Sesma L, Munoz J, Riezu JI, Caballeria J *et al*. (2005). *Proteomics* **5**: 4964–4972.

Fujii K, Kondo T, Yokoo H, Yamada T, Iwatsuki K, Hirohashi S. (2005). *Proteomics* **5**: 1411–1422.

Gu S, Du Y, Chen J, Liu Z, Bradbury EM, Hu CA *et al*. (2004). *J Proteome Res* **3**: 1191–1200.

Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. (1999). *Nat Biotechnol* **17**: 994–999.

Hanash SM, Celis JE. (2002). *Mol Cell Proteomics* **1**: 413–414.

Hanash SM. (2004). *Mol Cell Proteomics* **3**: 298–301.

He F. (2005). *Mol Cell Proteomics* **4**: 1841–1848.

Hippo Y, Watanabe K, Watanabe A, Midorikawa Y, Yamamoto S, Ihara S *et al*. (2004). *Cancer Res* **64**: 2418–2423.

Iida M, Anna CH, Hartis J, Bruno M, Wetmore B, Dubin JR *et al*. (2003). *Carcinogenesis* **24**: 757–770.

Kang SM, Shin MJ, Kim JH, Oh JW. (2005). *Proteomics* **5**: 2227–2237.

Karas M, Hillenkamp F. (1988). *Anal Chem* **60**: 2299–2301.

Kim J, Kim SH, Lee SU, Ha GH, Kang DG, Ha NY *et al*. (2002). *Electrophoresis* **23**: 4142–4156.

Kim W, Oe Lim S, Kim JS, Ryu YH, Byeon JY, Kim HJ *et al*. (2003). *Clin Cancer Res* **9**: 5493–5500.

Krishna RG, Wold F. (1988) In: Angeletti RH (ed). Academic Press: San Diego, CA, pp. 121–206.

Le Naour F, Brichory F, Misek DE, Brechot C, Hanash SM, Beretta L. (2002). *Mol Cell Proteomics* **1**: 197–203.

Lee CL, Hsiao HH, Lin CW, Wu SP, Huang SY, Wu CY *et al*. (2003). *Proteomics* **3**: 2472–2486.

Li C, Hong Y, Tan YX, Zhou H, Ai JH, Li SJ *et al*. (2004). *Mol Cell Proteomics* **3**: 399–409.

Li C, Tan YX, Zhou H, Ding SJ, Li SJ, Ma DJ *et al*. (2005). *Proteomics* **5**: 1125–1139.

Liang CR, Leow CK, Neo JC, Tan GS, Lo SL, Lim JW *et al*. (2005). *Proteomics* **5**: 2258–2271.

Lim SO, Park SJ, Kim W, Park SG, Kim HJ, Kim YI *et al*. (2002). *Biochem Biophys Res Commun* **291**: 1031–1037.

Liotta L, Petricoin E. (2000). *Nat Rev Genet* **1**: 48–56.

MacBeath G. (2002). *Nat Genet* **32**(Suppl): 526–532.

Mannova P, Beretta L. (2005). *J Virol* **79**: 8742–8749.

Mehta AI, Ross S, Lowenthal MS, Fusaro V, Fishman DA, Petricoin III EF *et al*. (2003). *Dis Markers* **19**: 1–10.

Melle C, Kaufmann R, Hommann M, Bleul A, Driesch D, Ernst G *et al*. (2004). *Int J Oncol* **24**: 885–891.

Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A *et al*. (2002). *Mol Cell Proteomics* **1**: 376–386.

Orchard S, Taylor C, Hermjakob H, Zhu W, Julian R, Apweiler R. (2004). *Expert Rev Proteomics* **1**: 179–183.

Paradis V, Degos F, Dargere D, Pham N, Belghiti J, Degott C *et al*. (2005). *Hepatology* **41**: 40–47.

Parent R, Beretta L. (2005). *J Hepatol* **43**: 177–183.

Park KS, Cho SY, Kim H, Paik YK. (2002a). *Int J Cancer* **97**: 261–265.

Park KS, Kim H, Kim NG, Cho SY, Choi KH, Seong JK *et al*. (2002b). *Hepatology* **35**: 1459–1466.

Parkin DM, Bray F, Ferlay J, Pisani P. (2001). *Int J Cancer* **94**: 153–156.

Pedrioli PG, Eng JK, Hubley R, Vogelzang M, Deutsch EW, Raught B *et al*. (2004). *Nat Biotechnol* **22**: 1459–1466.

Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. (1999). *Electrophoresis* **20**: 3551–3567.

Poon TC, Yip TT, Chan AT, Yip C, Yip V, Mok TS *et al*. (2003). *Clin Chem* **49**: 752–760.

Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E, Bragado-Nilsson E *et al*. (2001). *Methods* **24**: 218–229.

Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B. (1999). *Nat Biotechnol* **17**: 1030–1032.

Schwegler EE, Cazares L, Steel LF, Adam BL, Johnson DA, Semmes OJ *et al*. (2005). *Hepatology* **41**: 634–642.

Shalhoub P, Kern S, Girard S, Beretta L. (2001). *Dis Markers* **17**: 217–223.

Shui W, Liu Y, Fan H, Bao H, Liang S, Yang P *et al*. (2005). *J Proteome Res* **4**: 83–90.

Srisomsap C, Sawangareetrakul P, Subhasitanont P, Panichakul T, Keeratichamroen S, Lirdprapamongkol K *et al*. (2004). *Proteomics* **4**: 1135–1144.

Steel LF, Shumpert D, Trotter M, Seeholzer SH, Evans AA, London WT *et al*. (2003). *Proteomics* **3**: 601–609.

Suriawinata A, Xu R. (2004). *Semin Liver Dis* **24**: 77–88.

Syka JE, Marto JA, Bai DL, Horning S, Senko MW, Schwartz JC *et al*. (2004). *J Proteome Res* **3**: 621–626.

Takashima M, Kuramitsu Y, Yokoyama Y, Iizuka N, Fujimoto M, Nishisaka T *et al*. (2005). *Proteomics* **5**: 1686–1692.

Takashima M, Kuramitsu Y, Yokoyama Y, Iizuka N, Toda T, Sakaida I *et al*. (2003). *Proteomics* **3**: 2487–2493.

Tan TL, Chen WN. (2005). *J Clin Virol* **33**: 293–298.

Tannapfel A, Anhalt K, Hausermann P, Sommerer F, Benicke M, Uhlmann D *et al*. (2003). *J Pathol* **201**: 238–249.

Taylor-Robinson SD, Foster GR, Arora S, Hargreaves S, Thomas HC. (1997). *Lancet* **350**: 1142–1143.

Unlu M, Morgan ME, Minden JS. (1997). *Electrophoresis* **18**: 2071–2077.

Van den Bergh G, Arckens L. (2004). *Curr Opin Biotechnol* **15**: 38–43.

Wang H, Hanash S. (2003). *J Chromatogr B* **787**: 11–18.

Wang H, Hanash S. (2005). *Mass Spectrom Rev* **24**: 413–426.

Wirth PJ, Hoang TN, Benjamin T. (1995). *Electrophoresis* **16**: 1946–1960.

Wolters DA, Washburn MP, Yates III JR. (2001). *Anal Chem* **73**: 5683–5690.
Yang XJ. (2005). *Oncogene* **24**: 1653–1662.
Yates III JR. (1998). *J Mass Spectrom* **33**: 1–19.
Yates III JR. (2004). *Annu Rev Biophys Biomol Struct* **33**: 297–316.
Yokoo H, Kondo T, Fujii K, Yamada T, Todo S, Hirohashi S. (2004). *Hepatology* **40**: 609–617.
Yokoyama Y, Kuramitsu Y, Takashima M, Iizuka N, Toda T, Terai S *et al*. (2004). *Proteomics* **4**: 2111–2116.
Yu Y, Cui J, Wang X, Liu Y, Yang P. (2004). *Proteomics* **4**: 3112–3120.
Zeindl-Eberhart E, Haraida S, Liebmann S, Jungblut PR, Lamer S, Mayer D *et al*. (2004). *Hepatology* **39**: 540–549.
Zeindl-Eberhart E, Klugbauer S, Dimitrijevic N, Jungblut PR, Lamer S, Rabes HM. (2001). *Electrophoresis* **22**: 3009–3018.
Zhang J, Hu H, Gao M, Yang P, Zhang X. (2004). *Electrophoresis* **25**: 2374–2383.
Zhang LH, Ji JF. (2005). *World J Gastroenterol* **11**: 463–468.
Zhang Y, Fan XG, Chen R, Xiao ZQ, Feng XP, Tian XF *et al*. (2005). *World J Gastroenterol* **11**: 3485–3489.
Zhu H, Snyder M. (2003). *Curr Opin Chem Biol* **7**: 55–63.
Zwickl H, Traxler E, Staettner S, Parzefall W, Grasl-Kraupp B, Karner J *et al*. (2005). *Electrophoresis* **26**: 2779–2785.