Genomic Med. (2008) 2:173–175
DOI 10.1007/s11568-009-9091-8

ABSTRACTS

Disease proteomics

026: Quantitative proteomic approaches to identify biomarkers for hepatocellular carcinoma

[1,2]**Raghothama Chaerkady**, [3]Paul J. Thuluvath, [1,2,3]H. C. Harsha, [4]Vivekanandan Perumal, [4]Michael Torbenson, [2,4,5]Akhilesh Pandey

[1]Institute of Bioinformatics, International Technology Park, Bangalore 560 066, India, [2]Manipal University, Manipal, 576104, Karnataka, India, [3]McKusick-Nathans Institute of Genetic Medicine and Department of Biological Chemistry, Johns Hopkins University, Baltimore, Maryland 21205, USA, [4]Pathology, Johns Hopkins University, School of Medicine, Baltimore, MD, 21205, USA, [5]Oncology, Johns Hopkins University, School of Medicine, Baltimore, MD, 21205, USA

Hepatocellular carcinoma (HCC) is the fifth most common cancer worldwide and is frequently associated with cirrhosis, chronic hepatitis B (HBV) and C (HCV) virus infection. The objective of our study was to identify altered protein expression or glycosylation for use as biomarkers in HCC. Quantitation of proteins in complex biological samples using mass spectrometry has become feasible in recent years using methods that rely on incorporation of stable isotopes into peptides. Our approach was to use multiple technologies including the multiplexed isotope labeling for relative quantitation of HCC associated proteins using iTRAQ method and ^{18}O based quantitation and identification of tumor specific glycosylated proteins for the discovery phase and Western blotting and immunohistochemical labeling of HCC tissues for the validation phase. Two dimensional LC-MS/MS analysis and glycoprotein enrichment strategy generated more than 50,000 MS/MS spectra from which we were able to identify and quantitate more than 1,000 liver proteins in hepatocelluar carcinoma. Over 200 proteins were found to be upregulated in HCC, which included both previously described markers as well as novel ones. Novel overexpressed proteins in HCC included fibroleukin, myeloid-associated differentiation marker, prothymosin alpha, high mobility group AT-hook 1 isoform a and leucine rich repeat containing 59. Over 100 proteins were found to be downregulated in HCC relative to adjacent normal tissues which included urea cycle enzymes, class I and class II alcohol dehydrogenases, fatty acid binding protein 1 and prostatic binding proteins. Western blotting confirmed the results of both upregulated and downregulated proteins. Using immunohistochemical labeling, we were able to validate overexpression of fibroleukin, myeloid-associated differentiation marker and fetuin in HCC and downregulation of urea cycle enzymes, fatty acid binding protein and prostatic binding proteins in HCC using tissue microarrays ($n = 60$). Lectin affinity enrichment was found to be advantageous to quantitate several interesting proteins, which were not detected in the whole proteome screening approach. Using lectin affinity followed by PNGase F digestion coupled to ^{18}O labeling, we identified 34 glycosylation sites. This study indicates that quantitative proteomic profiling of tumor tissue versus non-cancerous tissue is a promising approach for the identification of potential biomarkers for HCC.

027: Embryonic stem cells, their biomedical potential and proteomics approach for quality control

[1]**K. S. Rakesh**, [2]Dawood Dudekula, [2]Nagaraja Ramaiah, [2]S. H. Ko Minoru, [1]Ravi Sirdeshmukh

[1]Centre for Cellular and Molecular Biology, Hyderabad 500 007, India, [2]Laboratory of Genetics, National Institute on Aging, 251, Bayview Blvd., Suite 100, Baltimore, MD, 21224, USA

Stem cells have great potential for developing new approaches for the treatment of degenerative diseases including neurons, skeletal and cardiac muscles, beta cells of pancreas, and cancer cells. Cell replacement therapy with stem cell and their derivatives are thus an important biomedical tool. A better understanding of the molecular pathways, regulatory networks and their dynamics, which determine their diverse differentiation fates, is needed for these therapeutic approaches to be successful. With these objectives, we have been studying protein expression in mouse embryonic stem cell lines (R1-9 and ABI), as model system. We have extensively studied their protein profiles by ESI LC MS/MS approach, integrated the data with transcriptomics studies as well as with proteomics studies from other laboratories. We have thus identified more than 2,000 proteins expressed in stem cells with high confidence. Pathway analysis of these proteins was carried out using KEGG, IPA, GenMAPP and their gene ontology classification revealed transcription regulators, signal transducers, cell cycle and differentiation molecules along with other general classes of proteins. Based on the specific proteins expressed, putative regulatory pathways operational in stem cells could be constructed. Such information on protein expression and plausible biochemical pathways and networks operational in the stem cells and

their derivatives would be useful in defining cell states and developing quality control methods for their biomedical application.

028: Functional characterization of putative disease proteins in Type 2 Diabetes Mellitus

Amitabh Sharma, Sreenivas Chavali, Rubina Tabassum, Mukesh Lalwani, Sridhar Sivasubbu, Dwaipayan Bharadwaj

Functional Genomics Unit, Institute of Genomics and Integrative Biology, CSIR, Mall Road, Delhi-110007, India

Type 2 Diabetes Mellitus (T2DM) is a polygenic disorder with both gene–environment and protein–protein interactions influencing the disease risk. A huge effort to unravel the risk factors has been undertaken worldwide for over a decade but with little success. On these lines, we designed a system biology approach to identify plausible disease candidates by assigning a score called weight value (Wv) ranging from 0 to 1 based on domain interactions and subsequent network analysis. Using this approach we identified certain already known T2DM proteins and also captured proteins of unknown functions and several new candidates for this dreaded disease. Our analysis revealed Lipin family member: Lipin-2 and Lipin-3 to be probable disease candidates in T2DM. Hence, to validate our approach, the functional characterization and identification of Lipin family was done in zebrafish model system. There are three members of lipin family having Phosphatidic acid phosphate (PAP) enzyme activity, responsible for generating diacylglycerol utilized in the synthesis of triacylglycerol (TAG). The regulation of TAG storage is important in human disease because both excessive and inadequate food storage might lead to abnormalities like insulin resistance and diabetes. Protein sequence alignment and phylogentic analysis of three unknown proteins in zebrafish showed high similarity to that of three human lipin family members. Expression analysis using Reverse transcriptase PCR of total RNA obtained from different tissues of zebrafish revealed the expression of Lipin-1, 2 and 3 proteins in intestine, gills, muscle, ovary, liver and kidney. In situ hybridization of Lipin members (Lipin-1, 2 and 3) revealed that all the three had expression pattern centered in eye and head region of 24 h post fertilization (hpf), 36hpf, 48hpf and 72hpf embryos. Thus, further studies assessing the roles of these three proteins in zebrafish would assist in defining their explicit role in human diseases.

029: Development of Novel Biomarkers for Breast Cancer in North Indian Population

[1]**Yogeshwer Shukla**, [1]Jasmine George, [1]Preeti Roy, [1]Pranav K. Chaurvedi, [2]Shailendra Kumar, [2]Sandeep Kumar

[1]Proteomics Laboratory, Indian Institute of Toxicology Research, Mahatma Gandhi Marg, Lucknow 226001, India, [2]Dept. of Surgery, King George Medical University, Lucknow, India

Breast cancer is one of the most common malignancies among women worldwide. Despite tremendous advances in screening, diagnosis, and treatment, the causes of this disease remain elusive and complex. Protein-based breast cancer biomarkers are a primary resource for breast cancer detection at the earliest and most treatable stages of the disease. Blood is an easily accessible source of proteins, which have diagnostic value, as it is in contact with practically all tissues in the human body. The vast dynamic range of protein abundance in blood plasma represent a major challenge in applying a proteomic based strategy for their identification. From an experimental design point of view, most cancer biomarker studies, including those aimed at identifying markers for early detection, are initiated with analysis of specimens from newly diagnosed subjects. The discovered candidate biomarkers are subsequently investigated for their utility for early cancer diagnosis. Differential protein expression in the breast cancer and healthy individual's plasma samples was investigated by two dimensional gel electrophoresis (2-DE), and spots showing a significantly expression between the two were analyzed. Image matching derived from two sets showed total of 112 spots. Out of 112, 2 proteins were up-regulated or present in patient's plasma samples as compared to control and 3 protein spots were found to be differentially expressed by means of quantitative analysis. This investigation can lead to the development of potential biomarker that may have clinical utility in discovering biomarkers of breast cancer.

030: Use of In-silico approaches to identify promoter elements and transcription factors associated with malignancy in ovarian cancer followed by validation of potential targets in tissues and serum samples of patients with ovarian cancer using proteomic approaches

[1]**Narasimhan Kothandaraman**, [2]Bajic Vladimir, [3]Razvi Khalil, [1]Stephen Koh, [4]Manuel Salto Tellez, [4]Brendon Pang, [5]Yiong Huak Chan, [1]Illancheran Arunachalam, [1]Jeffrey Low, [1]Arijit Biswas, [1]Mahesh Choolani

[1]National University Health System, Department of Obstetrics and Gynaecology, Diagnostic Biomarker Discovery Laboratory, Yong Loo Lin School Of Medicine, Singapore, [2]University of the Western Cape, South African National Bioinformatics Institute (SANBI), South Africa, [3]Southend University Hospital, NHS Foundation Trust, UK, [4]National University Health System, Department of Pathology, Singapore, [5]National University of Singapore, National University Medical Institutes, Singapore

Aims: Epithelial ovarian cancer (EOC) presents in the later stages of the disease. Regulatory genes/transcription factors (TFs) control the functions of various genes associated with cancer development. We hypothesized that over expressed TFs associated with EOC pathogenesis could be useful tissue and serum markers for malignancy associated with EOC.

Methods: Using microarray data we shortlisted candidate genes associated with apoptosis and cell structure integrity. Using promoter and TF binding analysis we shortlisted target TFs that could play a key role in malignancy associated with EOC. Using immunoblotting and IHC we validated the in silico predictions and associated these markers with CA125 for their clinical utility.

Results: Using a combined genomic-bioinformatics and proteomic approach we identified potential regulatory genes associated with the pathogenenicity of EOC. EOC-specific microarray data were examined for over-expressed genes. Promoter models of over-expressed genes were determined and bioinformatic analysis implicated transcription factors ZEB1, E2F5, PAX8 and ELF3, involved in cell proliferation, as promising regulatory targets in early stage disease. Additional analysis found E2F5 as a prime candidate for further study. Our in silico observation was supported by tissue microarray experiments that showed E2F5 expression only in EOC samples and not in normal and benign tissues as observed in a total of 135 (111 tumours and 24 normal) tissue array (45.6% of EOC tissue samples) and 144 serum (Normal: 56, benign: 40 and malignant: 48) 81.25%, $p = 0.0001$ patients with EOC). Analysis of clinical characteristics of this protein used in different combinations with CA125 for distinguishing malignant cyst from benign cyst shows increased sensitivity (97.9%; increase

from 87.5% if CA125 only is used) and more importantly increased specificity (72.5%; increase from 55% if CA125 only is used). This significantly improved accuracy suggests improved diagnostics of EOC. Overall, our findings provide evidence that some of the cell-cycle regulatory proteins might play a significant role in EOC pathogenesis. Conclusion: Our findings show involvement of E2F genes in EOC development and progression. Larger prospective studies are essential to validate our findings. We have demonstrated the application of microarray analysis could facilitate the identification of genes, genetic pathways, and proteins involved in the pathogenesis of EOC and their utility as potential serum markers.

031: BIOBASE Knowledge Library™ (BKL) Disease View and HGMD®: A comprehensive literature-derived disease information management system

[1]**Jesintha Maniraja**, [2]Jennifer Hogan, [2]Laurie DiDonato, [3]Ellen Fricke, [1]Sekhar Muppavarapu, [1]Mahesh Mahadevan, [1]Arul Prakasam Peter, [3]Olga Kel-Margoulis, [4]David N. Cooper, [3,5]Edgar Wingender

[1]Biobase Databases India Private Limited, 32/1 Crescent Road, Bangalore 560 001, India, [2]Biobase Corporation, 100 Cummings Center, Beverly 01915, MA, USA, [3]BIOBASE GmbH, Halchtersche Str. 33, D-38304, Woflenbuettel, Germany, [4]Institute of Medical Genetics, Cardiff University, Health Park, Cardiff CF14 4XN, UK, [5]Department of Bioinformatics, University of Goettingen, Goldschmidtstr. 1, D-37077, Goettingen, Germany

The **Biobase Knowledge Library™ Disease View** presents detailed information on 1988 different human diseases with the aim of helping researchers to understand the molecular mechanisms underlying human disease. More than 5,234 proteins are associated with disease terms resulting about 30,000 protein-disease links. Disease View provides manually curated links between genes and MeSH terms in four major categories: biomarker, therapeutic target, molecular mechanism and negative correlation; as well as information on expression of gene or protein in a particular disease condition and knockout models. Each disease report provides detailed information about the effects of molecular alterations on the biological processes associated with the respective gene. Information is provided for 4,284 biomarkers, 1,311 therapeutic targets, and 94,461 annotations comprising 4,937 polymorphisms, 12,749 mutations, etc.

Disease View is closely integrated with TRANSFAC®, a database on gene regulation and transcription factors, and TRANSPATH®, a database on signal transduction and metabolic pathways. Integrated databasing allows the user to make sophisticated queries and to generate new information based upon multiple facts brought together from the dispersed literature. As an example workflow through the integrated BKL™, one could identify all proteins associated with metabolic disorders, as well as those involved in signaling and metabolic pathways and gene regulatory networks. With the help of the accompanying tools, one could (1) retrieve the promoters of the corresponding genes, (2) identify putative TF binding sites and (3) analyze networks to find potential key nodes.

An exclusive feature of the BKL™ is the interlinking with the internationally acclaimed **Human Gene Mutation Database (HGMD®)** curated at Cardiff University. The HGMD® represents a comprehensive core collection of germ-line mutations in nuclear genes underlying or associated with human inherited disease. The database contains over 79,000 different lesions detected in 3,000 different human genes, with new entries currently accumulating at a rate exceeding 9,000 per annum. HGMD® records the first report of a disease-causing mutation or disease-associated/functional polymorphism. The data comprise single base-pair substitutions in coding, regulatory and splicing-relevant regions of human nuclear genes, micro-deletions and micro-insertions, indels, repeat expansions, gross lesions and complex rearrangements.

032: Methylation status of EpCAM in y79 retinoblastoma cell line: Proteomic analysis shows modification of various other proteins

Moutushy Mitra, Mallikarjuna Kandalam, Krishnakumar Subramanian

Department of Pathology, Sankaranethralaya, 18, College Road, Nungambakkam, Chennai, India

Retinoblastoma (RB) is the most common intraocular malignancy of infancy and childhood in which epithelial cell adhesive molecules (EpCAM) was found to be over expressed and associated with invasion of the choroids and optic nerve. But surprisingly Y79 RB cell line did not show any EpCAM expression. However the molecular mechanisms responsible for the downregulation of EpCAM in Y79 RB cell line was not known. DNA methylation is an important mechanism for inactivating various genes during tumourigenesis and progression. The presence of a CpG island in the TACSTD1 gene promoter and first exon, encoding EpCAM, led us to investigate in this study whether EpCAM expression can be influenced by DNA methylation. We examined the methylation status of the TACSTD1 promoter region RB cell lines using bisulfite sequencing. We found the promoter of EpCAM-negative Y79 RB cell line to be methylated to a higher degree. Demethylation of cell lines was performed using 5azacytidine (AZC). TaqMan gene expression assays were used to quantify the mRNA expression of EpCAM normalized against two endogenous controls using geNorm software. Immunofluorescence (IF) was carried out using EpCAM antibody. EpCAM RNA and protein expression could be partially restored by treating cells with AZC. By real-time PCR the EpCAM RNA expression increased more than 250-fold after 5 days of incubation with AZC compared to the untreated cell line, and by IF it was resulted in the expression of detectable amounts of EpCAM protein on the cell surface. We have also attempted to determine the changes in the other protein profile changes in AZC treated Y79 cell line. The strategy applied makes use of proteomics technologies to reveal and identify other proteins that are differentially regulated in control and in AZC treated cells. We have identified and analyzed 16 differentially expressed proteins in Y79 using 2DE-MS approach. The identity of most of these differentially expressed proteins was determined by 4,800 MALDI TOF TOF. We found 13 proteins are up regulated and 3 proteins were downregulated in AZC treated cells compared to untreated Y79. The functional significance of all the proteins was identified from swiss-prot. Taken together, these data suggest that epigenetic inactivation of genes by DNA methylation can be reversed by treatment with the DNA methylation inhibitor AZC. Our study will provide a basis for further investigation into metabolic pathways affected by demethylating drugs and their mode of action.