LECTURE-33

Challenges in Clinical Proteomics

Handout

PREAMBLE

Successful completion of human genome sequence project catalyzed the progress of proteomics research in different disciplines of modern science. Starting from mid 1990s, in due course this promising field has propelled its expansion in nearly every aspect of life science research. The emerging proteomics techniques have incredible potential to offer a plethora of new information to accelerate the pace of basic and applied clinical research; however, there are quite a few basic limitations of proteome level research; mostly due to fragile nature of proteins leading to substantial losses during sample collection and processing steps, post-translational modifications, presence of multiple isoforms of same proteins, and complexity and wide dynamic range of protein concentrations. Moreover, there is no direct amplification strategy in proteome-level research equivalent to PCR, which is used for gene amplification. In this lecture the challenges associated with the commonly used proteomics techniques will be discussed in light of clinical applications.

OUTLINE OF LECTURE

I. Technical challenges associated with different proteomic technologies

- **(a)** Gel-based proteomics
- **(b)** MS-based proteomics
- **(c)** Array-based proteomics

(d) Label-free proteomics

(e) Nanoproteomics

II. Challenges in biomarker discovery; detection of low-abundance proteins

III. Challenges in membrane proteomics

IV. Challenges in analysis of PTM

V. Conclusions

Box 1: Terminologies

Proteomics - Proteomics is comprehensive study of the expression, localization, interaction and post-translational modifications of whole set of proteins encoded by a genome at a specific condition.

Clinical proteomics - The most promising sub-discipline of proteomics, which deals with different possible clinical applications of proteomics, including disease biomarker discovery, study of disease pathogenesis, drug action, host-pathogen interactions and identification of potential drug /vaccine targets.

Gel-based proteomics - Protein separation techniques where polyacrylamide gels are applied for separation of multiple proteins present in complex mixtures. SDS-PAGE, native PAGE, 2DE and DIGE are the widely adopted gel-based proteomics techniques.

MS-based Proteomics - Combination of most useful analytical technologies for accurate mass measurements. Although there are many MS-based tools; the most commonly used MS platforms are ESI-Q-TOFMS/MS, MALDI-TOF and MALDI TOF-TOF.

Protein microarrays - High-throughput proteomics techniques where large number of proteins is concurrently immobilized on a glass or polyacrylamide gel pad surfaces, while the target proteins dissolved in solution allowed to pass through to investigate molecular interactions.

Label-free detection techniques - Label-free methods for quantification eliminates the need for labeling of the query molecules to obtain quantitative information. Such quantitation approaches provide accurate information for the protein abundance changes in complex samples.

Nanoproteomics - Integration of different nanotechnological approaches in proteomics generated this amalgamated analytical platform, which effectively improved the limit of detection, dynamic range, detection speed and multiplexing power of different conventional proteomic techniques.

Biomarkers - Biomarkers are indicator biomolecules that help to detect diseased states at an early stage, make discrimination between different diseases, and provide useful information for monitoring progression/severity of disease.

Membrane proteins - Proteins present in biological membranes are called membrane proteins, which are very difficult to isolate and study because of their hydrophobic properties and comparatively low concentration.

Post-translational modifications (PTMs) - PTM are alterations in the polypeptide chain generated by either the addition or removal of diverse chemical moieties, proteolytic cleavage, or covalent cross-links between different domains of the protein, which can effectively change the molecular functions of the proteins. Most commonly occurring PTMs include, phosphorylation, O-glycosylation, sulfation, nitration, and acylation.

TECHNICAL CHALLENGES ASSOCIATED WITH DIFFERENT PROTEOMIC TECHNOLOGIES

There are some inherent challenges in the field of proteomics due the following reasons:

1. Fragile nature of proteins
2. Lack of any protein amplification method (like PCR used for gene amplification)
3. Difficulties in extraction and isolation of proteins from specific organelles
4. Presences of various isoforms of single protein
5. Massive variation in proteome with time within same individual
6. Variations among the individuals of same or different populations
7. Wide dynamic range of protein concentrations in biological fluids

Technological limitations associated with proteomics approaches are summarized in Table 1.

Table 1: Limitations associated with different proteomics technologies

Type	Commonly used techniques	Limitations	Recent advancements
(a) Gel-based proteomics	1D SDS-PAGE, native PAGE, 2DE	• Poor reproducibility • Inadequate sensitivity and dynamic range (10^3-10^4) • Insufficient coverage of complex proteome • Low-throughput • Biasness in analysis	• 2D-DIGE: Better reproducibility and sensitivity (Unlu et al., 1997) • Post-electrophoresis Epicoccone fluorescent dyes like Lightning Fast and Deep purple: Increased sensitivity, dynamic range and coverage (Miller et al., 2006) • Activity-based protein profiling (enzyme-targeting probes)

		process, • Lengthy experimental process • Reliance on performer's technical skill	information regarding protein activity (Hu et al., 2003)
(b) MS-based proteomics	MALDI-TOF/TOF, ESI-Q-TOF, ESI-TRAP, MALDI-QUAD-TOF	• Narrow dynamic ranges (10^2-10^4) • Inadequate coverage • Low-throughput • Overfitting the data • Machine fluctuation • Instrument noise and contaminants in spectrum • Dearth of standard procedure for analysis and interpretation of MS and MS/MS spectrum (Patterson 2003; deVera et al., 2006)	• CDIT and Super-SILAC: large scale quantitative proteomics •TMT and iTRAQ (4 and 8 plex): multiplexing • Label-free LC-MS/MS: superior quantitative accuracy • Chip-based and Nano-LC-MS: low sample consumption • FTICR, LTQ-FT; improved sensitivity • MRM MS: large-scale biomarker discovery (Qian et al., 2006)
(c) Array-based proteomics	Protein microarrays, Antibody microarrays, Reverse phase microarrays	• Protein array designing difficulties • Problems in acquisition, arraying, and stable attachment of proteins to array surfaces • Inadequate sensitivity	• On-chip synthesis of protein; protein arraying by cell-free expression • Combination with label-free detection techniques: real-time sensitive detection • Use of nanoparticles for signal

		to detect very weak interactions and low-abundance analytes • Miniaturization of assays and protein dehydration • Non-specific binding • Unavailability of highly specific antibodies • Lack of direct correlation between protein abundance and activity (Kodadek 2001; Talapatra et al., 2002)	amplification: Better sensitivity
(d) Label-free proteomics	SPR, SPRi, Ellipsometry- and Interference-based techniques, Microcantilevers	• Reduced sensitivity and specificity when complex samples are analyzed (Yu et al., 2006; Ray et al., 2010a) • Costly fabrication techniques • Morphological anomalies of sample spots • Insufficient knowledge regarding the precise working mechanism	New label-free methods; SPR–MS, backscattering interferometry, brewster angle straddle interferometry, UV fluorometry, tagged-internal standard assay and spectral-domain optical coherence phase microscopy: Improvement of sensitivity and HT capability (Ray et al., 2010a)

(e) Nano-proteomics	Carbon nanotubes, nanowires, silicon nanowire field effect transistor, quantum dots, gold nanoparticles and nanocages	• Toxicity, biosafety and biocompatibility issues • Inadequate knowledge on the precise mechanism of action • Insolubility in biologically compatible buffers • Short life-time • Presence of metallic impurities • Lack of standard protocol for determining degree of purity of synthesized nanotubes and nanowires (Ray et al., 2010b; Ray et al., 2011a)	• Encapsulating shell and capping materials: longer life-time • Combinations with immunoassays • New cost-effective fabrication techniques (Ray et al., 2011a)

II. CHALLENGES IN BIOMARKER DISCOVERY; DETECTION OF LOW-ABUNDANCE PROTEINS

Biomarkers are indicator biomolecules that help to detect diseased states at an early stage, make discrimination between different diseases, and useful for monitoring progression/severity of disease. In spite of diverse advancements, even now there are several biological and technological limitations for the present proteomics technologies regularly applied for discovery of disease related marker proteins. Pre-analytical variations introduced during sample collection, handling and storage process, are also detrimental for screening of true biomarkers. Additionally, complexity of biological sample, very dynamic range of protein concentrations, presence of high-abundance proteins masking low-abundance marker proteins, high levels of salts and other interfering compounds in most of the biological specimens, insufficient sensitivity of the detection technology and paucity of throughput and multiplexed detection ability are the major obstacle for the direct application of proteomics technologies in clinics (Ray et al., 2011b).

With passage of time different combination of separation, detection and labeling strategies such as SCX (separation), ICAT, iTRAQ, TMT (labeling), nanoparticles like nanowires and nanotubes and quantum dots (signal amplification and enrichment of low-abundance proteins) etc. to surmount the basic technological limitations associated with existing proteomics approaches. Selective enrichment of low-abundance biomarkers and protection of degradable proteins are fascinated by the application of core shell hydrogel particles functionalized with various affinity selector baits and size exclusion exteriors. When the hydrogel particles are kept in biological fluids, those particles selectively entrap

low-abundance biomarkers in their baits and increase the effective concentration of those analytes (Longo et al., 2009) (Fig 1).

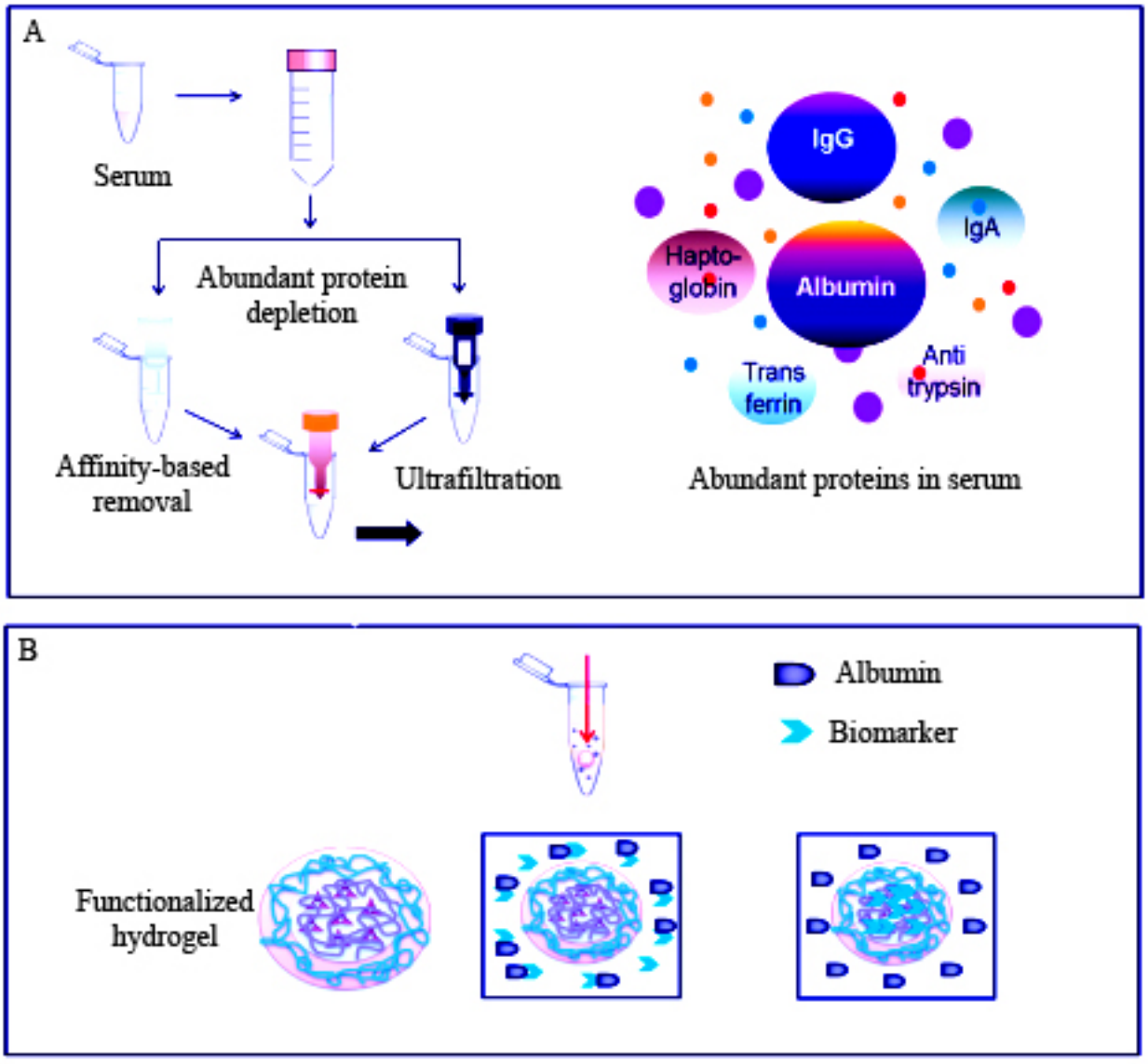

Fig 1: (A) Serum sample preparation for proteomic analysis. Inset shows the six major high abundance proteins, namely albumin, IgG, IgA, haptoglobin, transferrin and anti-trypsin, which makes it very challenging to detect the relatively low abundance protein biomarkers.

(B) Application of hydrogel particles for selective enrichment and preservation of low abundance biomarkers.

III. CHALLENGES IN MEMBRANE PROTEOMICS

Investigation of the proteins present in biological membranes is very informative but challenging due to the hydrophobic properties of the membrane proteins. Additionally, most of the membrane proteins are very low-abundance in nature. For proteome level analysis membrane proteins must be solubilized from the lipid layers, but usually solubility of membrane proteins is low at their isoelectric points. Due to the presence of multiple hydrophobic domains, membrane proteins tend to aggregate and subsequently precipitate during commonly used proteome sample preparation approaches (Helbig et al., 2010).

For enrichment of membrane proteins prior to the proteomic analyses different fractionation methods and free-flow electrophoresis and two-phase partitioning of the membrane vesicles are employed. Additionally, different commercially available surfactants are being used for improving membrane protein solubilization to increase the coverage of total membrane protein identification.

IV. CHALLENGES IN ANALYSIS OF PTM

Biological activities and molecular function of majority of the eukaryotic proteins depend on the post-translational modifications (PTMs). Efficient and sensitive methods for large-scale study of PTMs are still lacking (Mann and Jensen 2003).

Most commonly occurring PTMs include phosphorylation (+80 Da), O-glycosylation (>203 Da), sulfation (+80 Da), nitration (+45 Da), and acylation (>200 Da). Although analysis of PTMs in proteomics is a challenging task, since most of the PTMs are low abundance and/or sub-stoichiometric; MS-based analysis of PTM is very promising. However, the chemical stability of PTM is crucial for its efficient detection since many of the PTMs are very labile during MS and MS/MS analysis. Additionally, many PTMs are hydrophilic in nature, which makes PTM sample handling and purification extremely difficult prior to MS (Larsen et al., 2006).

V. CONCLUSIONS

Findings obtained from proteome-level research in the last ten years, have contributed significantly to unravel various unexplored secrets of human diseases and paved the way of proteomics in different applications of clinical research; including elucidation of mechanism of drug action, identification of drug and vaccine targets, establishment of diagnostic and prognostic biomarkers. However, different technological limitations associated with most common candidates available in proteomic tool-box are hindering the bed-side translation of proteomics approaches in real-life. Impending future of this highly promising research field will depend on successful solution of the existing limitations and collaborative initiatives at global level to prepare standard protocols for clinical proteomics research to avoid pre- and post-analytical variations.

REFERENCES FOR FURTHER READING

deVera IE, Katz JE, Agus DB. Clinical proteomics: the promises and challenges of mass spectrometry-based biomarker discovery. Clin Adv Hematol Oncol. 2006; 4(7):541-9.

Helbig AO, Heck AJ, Slijper M. Exploring the membrane proteome--challenges and analytical strategies. J Proteomics. 2010; 73(5):868-78.

Hu Y, Huang X, Chen GY, Yao SQ. Recent advances in gel-based proteome profiling techniques. Mol Biotechnol. 2004; 28(1):63-76.

Kodadek T. Protein microarrays: prospects and problems. Chem Biol. 2001;8(2):105-15.

Larsen MR, Trelle MB, Thingholm TE, Jensen ON. Analysis of posttranslational modifications of proteins by tandem mass spectrometry. Biotechniques. 2006; 40(6):790-8.

Longo, C., Patanarut, A., George, T., Bishop, B. et al., Coreshell hydrogel particles harvest, concentrate and preserve labile low abundance biomarkers. PLoS One 2009, 4, e4763.

Mann M, Jensen ON. Proteomic analysis of post-translational modifications. Nat Biotechnol. 2003; 21(3):255-61.

Miller, I., Crawford, J., Gianazza, E., Protein stains for proteomic applications: which, when, why? Proteomics 2006, 6, 5385–5408.

Patterson SD. Data analysis--the Achilles heel of proteomics. Nat Biotechnol. 2003; 21(3):221-2.

Qian WJ, Jacobs JM, Liu T, Camp DG 2nd, Smith RD. Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. Mol Cell Proteomics. 2006; 5(10):1727-44.

Ray S, Mehta G, Srivastava S. Label-free detection techniques for protein microarrays: prospects, merits and challenges. Proteomics 2010a; 10:731–48.

Ray, S., Chandra, H., Srivastava, S., Nanotechniques in proteomics: current status, promises and challenges. *Biosens. Bioelectron.* 2010b, *25*, 2389–2401.

Ray S, Reddy PJ, Choudhary S, Raghu D, Srivastava S. Emerging nanoproteomics approaches for disease biomarker detection: A current perspective. J Proteomics. 2011, 74, 2660-2681.

Ray S, Reddy PJ, Jain R, Gollapalli K, Moiyadi A, Srivastava S. Proteomic technologies for the identification of disease biomarkers in serum: advances and challenges ahead. Proteomics 2011b; 11: 2139–61.

Talapatra A, Rouse R, Hardiman G. Protein microarrays: challenges and promises. Pharmacogenomics. 2002; 3(4):527-36.

Unlu, M., Morgan, M. E., Minden, J. S., Difference gel electrophoresis: a single gel method for detecting

changes in protein extracts. Electrophoresis 1997, 18, 2071–2077.

Yu, X., Xu, D., Cheng, Q., Label-free detection methods for protein microarrays. *Proteomics* 2006, *6*, 5493–5503.