Biological Forum – An International Journal **15(3): 511-523(2023)**

ISSN No. (Print): 0975-1130
ISSN No. (Online): 2249-3239

Proteomics: An Emerging Tools, Database and Technique for Understanding Biological System

Amanulla Khan[1*], R.L. Pawar[2], M.H. Mali[3], Kamlakar More[4] and Satish V. Deore[5]

[1]Department of Botany, Anjum Islam Janjira Degree College of Science, Murud-Janjira (Maharashtra), India.
[2]Department of Botany, M.T.E.S.',
Doshi Vakil Arts and G.C.U.B. Science & Commence College, Goregaon – Raigad (Maharashtra), India.
[3]Department of Botany, C.H.C. Arts,
S.G.P. Commerce and B.B.J.P. Science College, Taloda, Dist. Nandurbar (Maharashtra), India.
[4]Department of Botany, Sant Gadge Baba Amravati University, Amravati (Maharashtra), India.
[5]Department of Botany, J E S's, Arts, Science and Commerce College, Nandurbar (Maharashtra), India.

(Corresponding author: Amanulla Khan)*
(Received: 19 January 2023; Revised: 14 February 2023; Accepted: 20 February 2023; Published: 22 March 2023)
(Published by Research Trend)

ABSTRACT: Proteomics is an advance tool in protein identification, structures prediction, protein domains and post transcriptional modifications. Proteomics study have several challenges, including the complexity and dynamic nature of the proteome, technical limitations, data analysis, sample preparation, and lack of standardization. Addressing these challenges will require continued innovation and collaboration among scientists, technologists, and bioinformaticians to advance the field of proteomics and realize its full potential in advancing our understanding of biological systems. Various technologies are developed and evolved time to time accordingly to the requirement and availability of the design experiments. Including traditionally used 2DE, immunological to advance modern tandem MS, with more data accuracy and resulted inferences. Information's, reports and data of proteome from these technologies result to build a specific, organized, and formed databases. These biological databases are evolved from primary to secondary databases and many special, accordingly to their specific features. Databases are Improve and develop with relations and collaboration to others database in various forms and formats. These all techniques and databases are having various latest improvements and additions in recent, which help in high level of protein study to understanding biological system level.

Keywords: 2DE, tandem MS, Proteome, Proteomics, Biological Databases, and Biological System.

INTRODUCTION

Biological sciences in the advance post-genomics time are characterized by the famously use of term "omics technologies" (Fernando *et al.,* 2011, Ganesh *et al.,* 2011 and Raquel *et al.,* 2010). These clear trends in the life science details research towards biological entities up to the system level (Marcus *et al.,* 2007; Bilal *et al.,* 2017; Clark *et al.,* 2021). "Omics technologies" is used including Genomics, Trnascriptomics, Proteomics, and Metabolomics (Issa and Grier 2008; Ewa *et al.,* 2003; Lv *et al.,* 2017 ; Beniddir *et al.,* 2021).

From past two decades, "proteomics" are in prominence as a key strategy to investigate systems in biology (Agrawal *et al.,* 2013; Lv *et al.,* 2020; Yan *et al.,* 2022).Where RNA, DNA, Protein, and Metabolites are make a terms Transcriptome, Genome and Metabolom (John *et al.,* 2009; Jon and Visith 2004; Dong *et al.,* 2020) proteome and Metabolome respectively. These omics knowledge evolving a term "Paintomics" which is an tool to visualization of trnascriptomics and metabolomics data from web sources (Fernando *et al.,* 2011; Alseekh *et al.,* 2021). Similarity Human Proteome Organization (HUPO) is developed for the promotion of proteomics with international cooperation with representation of scientific organization for the development of new techniques, new technologies and skilled bases training for proteomics tools (http://www.hupo.org/). As the sequencing of DNA is developed and enhance lager amount of DNA sequence base data is accumulated in databases (Helen, Philip and John 2004; Imanzadeh *et al.,* 2015), meanwhile researchers realized that, the complete known genome sequence is not sufficient to understand the biological system until the expressions of proteins are not recorded at translational level (Akhilesh *et al.,* 2000) and hence it is important to identifying the expressional components of gene in the form of protein (Saligrama *et al.,* 2008, Liu *et al.,* 2020) which leads the great important in proteomics study. Advance technologies, bioinformatics tools and database are now become more complex than genomic information to understood the investigate the comparative proteomic study (Satyanand *et al.,* 2010, Bilal *et al.,* 2017).

The "**Proteomics**" term was firstly coined by Marc Wilkins, in 1994 at Seina Meeting, to simply refer to the "PROTe in component of a genome" (Raquel *et al.,* 2010). This can be defined as the study of systematic and large-scale analysis of proteins (Satyanand *et al.,* 2010; Akhilesh *et al.,* 2000; Ohkmae *et al.,* 2004; Wilkins *et al.,* 1997). It is a rapidly demanding, progressing and developing discipline of this current functional genomics era (Richard and Dalia 2010), will contribute greatly to our understanding of genes function in post-genomic era (Akhilesh *et al.,* 2000; Jorrin *et al.,* 2006), particularly by structure, function and modification of proteins. Due to main catalysts of biological function (Roman *et al.,* 2005; Bilal *et al.,* 2017) several disorders, diseases, abnormalities are caused by improper function of the protein or its expression (Fred *et al.,* 2003), and progressive miss folding of specific protein (Dennis *et al.,* 2003). To better understanding biological systems with the help of protein study various methods, mechanisms technologies and databases (Helen, Philip, and John. 2004, Pouya *et al.,* 2015, Farshid *et al.,* 2015) have evolved. Proteome analysis involves separations, visualization of protein and identification of protein respect to genomic base sequence (Jon and Visith 2004; Francisco *et al.,* 2004; Nutan *et al.,* 2022). Ones this information is collected leads to identification of modified structure on the basis of protein motifs of that protein (Ole *et al.,* 1998). The various practically applied technique like Electrophoretic, Spectroscopic, Chromatographic and Immunological as well as software's and databases are the ways to study proteins which helps to understanding the biological science by protein compositions and its amino acids sequence chain (Nayuni *et al.,* 2022; Palak *et al.,* 2022).

ELECTROPHORESIS TECHNIQUES

1. Two Dimensional Gel Electrophoresis (2DE): Proteomic study 2DE is only the best, sensitive, most powerful and highest resolution analytical method of separation and identification of protein which is widely in use (Arsi *et al.,* 2005; Wenzhu *et al.,* 2000; Scott *et al.,* 1994). 2DE was first introduced in the early 1970s and by using pH based gradient gel experimental problems was overcome (Ganesh *et al.,* 2011; Sarka *et al.,* 2000). The objectives are mostly to know protein that are unregulated or regulated in specific manner with respective to dieses to developed diagnostic markers or therapeutics targets (Akhilesh *et al.,* 2000; Fatemeh *et al.,* 2009). Thousands of proteins can be resolved from single gel electrophoresis (Kose *et al.,* 1995; Waltraud *et al.,* 2010). This technique uses charges which helps to separate different size of proteins in dimension first and further on the basis of molecular weight while in dimension second (Ole *et al.,* 1998). There are several technical challenges like membrane proteins and large hydrophobic protein are not enter in to the dimension second of 2DE (Akhilesh *et al.,* 2000; Kumar *et al.,* 2014). Besides these drawback's, 2DE is still the comparatively good method to separate large and complex protein present in mixture (Arsi *et al.,* 2005). Concentration of protein spot from 2DE ranges from 1ng (nanogram) to 1 µg (microgram) protein (Kose *et al.,* 1995). Identification of spots and image analysis performed with the help of Micro Edman Degradation procedures (Dustin *et al.,* 2005; Stephen *et al.,* 1988), were due to some limited sensitivity of Micro Edman Degradation technique (Ole *et al.,* 1998), each 2D spot is separately digested, extracted and then analyzed, which required large time (Himanshu *et al.,* 2001; Pranav *et al.,* 2022). Now a day's improvements in technologies and various software packages are used for theses analysis, Software like Progenesis (Nonlinear Dynamics, New Castle Upone-Tyne, UK), PDQuset Software, (BioRed, Hercules, USA), etc and digitized images and spot features were automatically detected (Elena *et al.,* 2006, Xiaofeng *et al.,* 2006; Maria *et al.,* 2010, Faraji *et al.,* 2019). Were Propic robot (Genomic Solution Ann Arbor MI. USA), ProGest automated digester (Genomic Solution) etc are also new improvement in 2DE (Dustin *et al.,* 2005). Which all can help in the understanding the complex biological function (Muhammad *et al.,* 2007 and Kumar *et al.,* 2014).

2. Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE): SDS-PAGE is commonly used proteomic tool which separates and identified proteins by the basis of molecular weight of molecular weight (Francisco *et al.,* 2004; Patton, 2002). Sodium Dodecyl Sulfate is an anionic detergent is used to separate proteins to their primary structure and make coat them with uniform negative charge. Despites being introduced more than 30 years ago, there are only been a small development on the basic principle in SDS PAGE (Patton 2002,). The stacking with small quantity of acrylamide concentration and to making wells producing different protein concentration are compared (Ramu *et al.,* 2004). The significant changes are to introduction of gradient gels, while to perform electrophoresis at near neutral pH in SDS-PAGE (Saligrama *et al.,* 2008). These developments make protein separation with wider molecular weight range and decrease protein degradation during preparation of sample (Gisele *et al.,* 2009).

Electrophoretic separations in proteomics can then be detected with the help of various staining techniques:

(i) Organic dyes, like colloidal Coomassie blue (Ramu *et al.,* 2004; Kaley *et al.,* 2009).

(ii) Zinc-imidazole staining (Raquel *et al.,* 2010).

(iii) Fluorescence-based detection (Ole *et al.,* 1998; Gisele *et al.,* 2009; Raquel *et al.,* 2010)

(iv) Silver staining (Saligrama *et al.,* 2008).

Post-SDS PAGE staining also increase accuracy (Patton, 2002). Molecular mass of 10,000 to 300, 00 soluble proteins are readily covered which also extremely basic as well as acidic proteins can easily visualized (Akhilesh *et al.,* 2000 and Holly *et al.,* 2004). But a major drawback of using 2 Dimensional gels electrophoresis is their incompatibility with membrane proteins due to its hydrophobic nature and play crucial role in cellular mechanism (Kathryn *et al.,* 2007).

3. Differential in Gel Electrophoresis (DIGE): This techniques is introduced by Unlu with his co-workers fellows in 1997 (Unlu *et al.,* 1997). This enables the different protein sample into a single 2 D gel. Labeling each protein sample can possible with separately resolvable florescent cyanine dyes (Cy3 or Cy5) (Ganesh *et al.,* 2011, González *et al.,* 2020). The differentially florescent cyanine dyes (Cy3 and Cy5) labeled protein are mixed and resolved after extractions on the single gel than fluorescence imaging helps to detect difference among these two proteins extracts (Ngan *et al.,* 2011). A third protein was labeled with another florescent cyanine dyes (Cy2) which provides robust statistical information for analysis (Alban *et al.,* 2003). The technique also used to find changes in protein structure in various environmental stresses (Qui *et al.,* 2008). The Differential in Gel Electrophoresis approach has been successfully applied on phosphorylation study in membrane protein as well as to compare dark and light proteome adaptation of thylakiod lumen in chloroplast (Waltraud *et al., 2010*).

4. Blue-Native Polyacrylamide Gel Electrophoresis (BN-PAGE): These are the useful mixture of 1-and 2-DGE which was invented by Schagger and Von Jagow in 1991, to investigate protein complexes and its compositions in respiratory chain from the different organisms (Ganesh *et al.,* 2011). Braun with his co-workers is defined complexes of plant mitochondria with supra molecular organization in chloroplast with this technique. The principle of this technique based on solubilization of protein specially membrane by using mild nonionic detergent, such as triton X100, n-dodecylmaltoside or digitonin. Complex protein solubilizations are critical step after that solublized complex is separated depending upon the mobility of the complex within the gel matrix (Heinemerey *et al.,* 2004). Advances techniques in using the BN-PAGE technology with combination of DIGE as the one dimensional separation gives new prospective and applications (Waltraud *et al.,* 2010*,* González et *al.,* 2020; Melby *et al.,* 2021). Bioinformatics tools and data bases for electrophoresis are available at http://world-2dpage.expasy.org/repository (database), http://expasy.org/melanie and at http://www.2d-gel-analysis.com (image processing tools).

MASS SPECTROSCOPY TECHNIQUES

J J Thomson was fist time built Mass Spectroscopy prototype to measure m/z of electron and awarded with Nobel Prize in year 1906. Where Mass Spectroscopy concept was first introduce by Francis Aston in Cambridge England in year 1919. Mass spectroscopy is the most significant, versatile, powerful and comprehensive tool in large scale protein study in proteomics (John *et al.,* 2009 and David *et al.,* 1998) that allows more accuracy determination of gaseous ion based measurement of its mass to charge ratio (2-5 ppm) (Winston *et al.,* 1997), high resolution and dynamic range more than 10^3 (John *et al.,* 2009 and Winston *et al.,* 1997) and the technique is still evolving (Akhilesh *et al.,* 2000, Timp and Timp G2020). This is preferred method for identification of unknown protein, which can applied to the study of biological problems, such as apoptosis, human cancer and also to elucidate the component of several multi-protein complex (David *et al.,* 1998; Aebersold and Mann 2016; Aabid *et al.,* 2022). Basically mass spectroscopy contains three main components that are a good ionization sources for ionization of peptide, mass analyzer and a detector. With these additionally Inlet system will be the first for protein adding via HPLC, GC, etc and last will be the data analyzing system.

1.Ionization methods: Mass Spectroscopic Ionization are done by Electrospray Ionization (ESI) (Raquel *et al.,* 2010, Ohkmae *et al.,* 2004; Ohkmae *et al.,* 2004; Waltraud *et al.,* 2010; Eugene *et al.,* 2005), Nanoelectro ionization (nanoESI) (Enrique *et al.,* 2005; Setsuko Komatsu 2008), Matrix Assisted Laser Desorotion or Ionization (MALDI) (Marcus *et al.,*2007 and Waltraud *et al., 2010*), Electron Ionization (EI), Chemical Ionization (CI) (Raquel *et al.,*2010), Fast Atom/Ion Bombardment (FAB) (Issa Coulibaly *et al.,* 2008), Atmospheric Pressure Chemical Ionization (APCI) (Robert *et al.,* 2002), Desorption/Ionization on Silicon (DION) (John *et al.,* 2009), Field Desorption/Field Ionization (FD/FI), and Thermo spray ionization. These soft ionization methods (especially ESI and MALDI) allow for the peptide and protein analysis (John *et al.,* 2009). Developments of ionization techniques have driven Mass Spectroscopic build instrument with large range, higher resolution, with mass accuracy and low the coast (Willam *et al.,* 2001).

Electrospray Ionization (ESI): This technique used in MS to produce ions. Is widely in used to mixture of biomolecules into the MS instrument (Willam *et al.,* 2001; Heng *et al.,* 2003), by high voltage (2-6 kV) and applied between separation pipeline and inlet system of the MS. Physicochemical processes of ESI produce electrically charge spray which provides analytic solvent droplets in system. Formation and desolvation is heated by capillary, and inject at the mass spectrometer inlet into system (John *et al.,* 2009). The ESI process of forming gas phase ions are devoted from the work of Dole et al (Willam *et al.,* 2001). The main properties of ESI to allow rapid transfer of analyets at atmospheric pressure from the liquid to the gas phase (Heng *et al.,* 2003; Aebersold and Mann 2016). There are many models of ESI ion formation (Andrej *et al.,* 1996). Similar type of work was started in year 1910, different conformation of same type of protein give rise to specific ESI-MS spectra (Chowdhur *et al.,* 1999). ESI MS approaches added to the complication with multiple peaks, charge state distribution and affecting sensitivity (Kenneth *et al.,* 2006). ESI with tandem MS instruments can used to obtain peptide between 5-15 amino acids. In this fingerprint ion peak is 'captured' and defragmented into amino acid fragment of peptide (Gygi, 1999). Nano ESI is compatible by with capillary reverse phase i-e RP, columns with higher sensitivity than analytical column. The source of ESI is usually coupled with continuous analytical instruments (John *et al.,* 2009). Nano ESI MS/MS is used to localize

phosphorylation sites of solubilized proteins, and recently has been also applied to PAGE isolated phospho-protein (Ole *et al.,* 1998, Aebersold; Mann 2016; Cho *et al.,* 2021).

Matrix-Assisted Laser Desorotion/Ionization (MALDI): Widely used tool for peptide mass mapping in Mass Spectroscopy (Ole *et al.,* 1998, Wenzhu *et al.,* 1994), this technique attracted, attention of the Mass Spectroscopy community during the 11th International Mass Spectroscopy Conference in Bordeaux (1988), when Karas and Hilleenkamp described MALDI of protein with molecular masses exceeding 10 kDa (Willam *et al.,* 2001). MALDI evaluated to directly measuring ratio of substrates and product in enzyme assay (Kenneth *et al.,* 2006), low resolution structure studies of protein and protein complexes and selective radical probe reaction performed on the location of residue on molecular surface (Giovanni *et al.,* 2007). In MALDI, matrix absorbs laser energy and transfers energy to the acidified analyte, whereas the rapid laser heating causes desorption of matrix with $[M + H]^+$ ions into the gas phase (Willam *et al.,* 2001). MALDI ionization needs hundreds of laser shots to achieve acceptable signal with to the noise ratio for ion detection and generate charge. This makes applicable detailed analysis of high molecular weight protein with the help of MALDI (David *et al.,* 1996). It can help to assist various parameters such as linearity and reproducibility and automated multi-enzyme compound screening (Scott *et al.,* 1994, and Kenneth *et al.,* 2006). The drawbacks of MALDI are low in reproducibility and sample preparation. A key development in MALDI is with atmospheric pressure and can be terms as atmospheric pressure MALDI (AP-MALDI) (John *et al.,* 2009), It combines MALDI ion with a highly efficient tandem MS unit that fragment the individual peptide (Akhilesh *et al.,* 2000). The spectra received from AP-MALDI can be analyzed by using the MssLynx 4.0 software (Waters, Milford, MA) and all obtained spectra from AP-MALDI were used to identify proteins in Swiss-Prot protein sequence database by Protein Lynx Global Server 2.0 software (Elena *et al.,* 2006).

Fast Atom/Ion Bombardment (FAB): FAB is another ionization technique introduced by Barber and his co-workers in year 1981 (Barber *et al.,* 1981). First time, routine mass spectrometric analyses with polar thermally labile molecules are practiced in few thousand daltons (Willam *et al.,* 2001). FAB is the suitable method or the analysis of samples with preformed ions in solution, which are protonated or sodiated molecule. A solution of sample containing perhaps $[M + H]^+$ ion and complementary negatively charge counter ions, are dissolved in FAB matrix and introduced into the mass spectrometer vacuum system. The matrix solution of the sample is then bombarded with fast atoms (8 keV, Xe) or ions (20keV, Cs^+) and energy is transferred to matrix sample solution with the result that sample $[M + H]^+$ ions are vaporized along with protonated matrix clusters (Willam *et al.,* 2001; Barber *et al.,* 1981).

2. Mass analyzers: Broadly mass analyzer is categorized into two main categories, which are with the scanning and ion beam MS and trapping MS. Normally the well known ionizers are Time-of-Flight (TOF), Quadropol (Q), Iontrap (IT), Fourier Transform Ion Cyclotron (FTMS), and Orbitrap are the different analyzers.

Time-of-Flight (TOF): Is scanning mass analyzer was, used to measure the time that of travel by the object and detector over a known distance (John *et al.,* 2009). From the year 1948 to 1952, TOF mass analyzers introduced and interfaced with MALDI to performed better pulsed analysis (Kenneth *et al.,* 2006). In protein chemistry filed it is widely in used. TOF shows ability to directly mass analyze for the protein in naturally occurring biological liquids (David *et al.,* 1996). Analysis of integral membrane protein for peptide mass fingerprinting can be done with passive identification of protein by TOF in combination with ESI Q TOF MS (p80). MALDI TOF is good for mass measurement of proteins up to 500 kDa shows accuracy is between 0.1 and 0.01% with fast sampling rates (Kenneth *et al.,* 2006).

Iontrap (IT): This instrument is the high –throughput workhorses in proteomics. Ion trap invents in 1983 by W. Paul (wins 1989 Nobel Prize). Use of this instrument almost exclusively for the identification of proteins from enzymatic digests at high sensitivity form e.g. 1D and 2D gel spots. And versatile instruments features are fast scan that's rates MS^n, for each scans with good resolution, high duty cycle, mass accuracy and high sensitivity (John *et al.,* 2009).

Quadropol (Q): Is a mass analyzer responsible for filtering sample ions where ion filters was also introduced by **W. Paul in 1955**, as per their mass to charge ratio (m/z). Were the ions are separated by stability or trajectories in the oscillation of electric fields applied on rods.

3. The detectors: Is the final element in mass spectrometer which records either charge induced with ion passes that's hits a surface of detector. Normally some kinds of electron multiplier are used and also other detectors are used according to requirements of experiments. Finally data analysis system which involved various algorithms and software which resulted into good data collection of all protein those are tested. Software like MSight by SIB use for image development from ms, SPIDER used ms/ms data for protein identification. The algorithms like PeptIdent (Source: http://www.us.expasy.org/tools/peptident.html), MultiIdent, and ProFound (Source: http://129.85.19.192/profound_bin/WebProFound.exe) are used in proteolytic fingerprinting in protein identification. Where Moscot is an another powerful engine from Matrix Science (Source: http://www.matrixscience.com), that are using mass spectrometry data to identified protein from primary protein sequence database (Gisele *et al.,* 2009; Brigit *et al.,* 2006). Programs operating with ms/ms spectra are Pepsea, and Pep3D (Eugene *et al.,* 2005) are in

commonly used. The algorithm MOWSE (Martin *et al.,* 2005) is more selective and sensitive than other in calculating only number of matching peptides available at Source: http://www.narrador.embl-heildlberq.de/GroupPages/Homepage.html (John *et al.,* 2009).

Multidimensional Protein Identification Technology (MudPIT): This technology is developed to analyze the highly complex samples required for the large-scale analysis of proteome with the help of Electrospray Ionization (EI), database searching and tandem mass spectrometry (MS/MS). Protein is identified by this are statistically significant number of matches (Clair *et al.,* 2007). Recent studies have identified more than the 2,000 proteins in a single fraction with MudPIT provided significant result (Eric *et al.,* 2009). It used with an wide range experiments like large-scale catalogues of protein, organisms protein profiling and organelle or membrane level protein identification (Ganesh *et al.,* 2008). A recent investigation gives four stages of the parasite's life cycle in malaria by these methods (Florens *et al.,* 2002). It helps to gives the stages of human host (Eric *et al.,* 2009). A new strategy, ultra-high-pressure, and advance components of MudPIT (UHP-MudPIT) improve in the resolution of peptides as well as increasing the length of the C_{18} phase were decreasing the particle size (Clair *et al.,* 2007). Technologies such as MudPIT have accelerated biological discovery and leads to openings new ways in research and study of protein and proteomic science (Jorrin *et al.,* 2006).

PepFrag: Another software tool operates with MS/MS which allows searching of protein sequence database (from SWISS-PORT, PIR, GENPEPT) by using a combination of different type of information from mass spectra of peptide maps and fragmentations spectra of peptides. It is publically available on the internet as a section to the PROWL, which is an interactive environment at websites for protein mass spectrometry and protein software like this are available on web for protein identification (David *et al.,* 1998).

Tandem MS: Other approach in spectroscopy utilizations wildly is MS/MS or tandem MS or MS^2 this technique was appears in 1968. In that ions are mass selected with the help of mass analyzer I (MS_1) and then puts into a collision region proceeding with analyzer II (MS_2). Inert gases are generally used to introduce collision (Willam *et al.,* 2001). Application of MS/MS for the structural classification of organic molecule was given by Beynon and his co-workers at Purdue University. These ms/ms can also allow the protein sequencing by various coupling supports (Himanshu *et al.,* 2001). The 1st Triple-Quadropol MS/MS instrument was mad by Yost and Enke in 1970s after twenty years in 1990 the Quadropol-Ion-Trap MS/MS become more popular and many others like hybrid ms/ms instrument. The 1st commercial hybrid MS/MS is the Micro-mass AutoSpect-OATOF, that's having a double focusing type of arrangement in MS_1 and linear TOF arrangement in MS_2 (Willam *et al.,* 2001). Other various development and improvements has been done which become helpful to proteins at system levels.

All described spectroscopic techniques which are experimentally applied with various combinations, like MALDI-MS (S. P. Gygi 1999; Wenzhu *et al.,* 2000, Michael *et al.,* 2000, and Martin *et al.,* 2011), MALDI-TOF-MS (Wenzhu *et al.,* 1994; Sarka *et al.,* 2000; Kris *et al.,* 2000), and ESI-Q-TOF-MS, nanoelectrospray MS/MS (Ole *et al.,* 1998, Holly *et al.,* 2004), LC-ESI-MS (Giovanni *et al.,* 2007) and many others are in uses according to requirements of experiments.

These techniques are basically taken primary data from a technique like Electrophoresis which is described above, Chromatographic and Immunological techniques. All including MS are direct or indirect connected and finally making a huge amount of data of protein information including various protein sequences, structures, domains, modifications of proteins etc leads to increasing depositions in the proteins databases. Which together (different experimental tools and databases) are becoming a good field of study of proteins in broad view of points for biological system understandings.

DATABASES FOR PROTEIN STUDY

(Biological databases). Biological databases are includes databases of protein and nucleic acid but on broad view, biological databases are classified into structural database and sequence databases (Table 1). Sequence database are applicable for both, whereas structure database only applicable for protein. The fist protein sequence database was created after the sequencing of insulin peptide by F. Sanger in year 1953. Insulin is the first protein to be sequenced of 51 amino acids (Helen, Philip, and John. 2004, Wang *et al.,* 2014, Rosenberg and Utz, 2015). After that various development and important in technology for studying protein leads to development of well known protein data bank (PDB) in 1971. This was having only 10 entries. Know exponential development of PDB shows more than 25000 available structures of different proteins (Helen, Philip, and John. 2004) this leads to developments of modern biomedical architecture data bases (Cole and Moore 2018). As well as subsequently development of various other protein databases with different requirements and needs are to be developed, like Swiss-Prot, TrEMBL, CluSTr, UniProt, PRIDE, InterPro, ProToMap, IPI, REIASD and PHYTOPROT which are helping in understanding of various biological pathways and systems in proteomic science.

Table 1: Broad Classification of Protein databases.

Protein Databases		
Structural Databases		**Sequence Database**
Pri-DB	-PDB	-SWISS-PROT -TrEMBL -CluSTr -UniProt
Sec-DB	-SCOP -CATH -RESID	-InterPro

1. SWISS–PROT: A protein sequence database from Department of Medical Biochemistry of the University of Geneva release in year 1986. Now SWISS PROT is a joint enterprise of the European Molecular Biology Laboratory (EMBL) and Swiss Institute of Bioinformatics (SIB) (Michal *et al.,* 2007). It has also joined to the GO Consortium (http://www.geneontology.org) and has adapted its structured and its vocabulary to characterize the activities of protein in TrEMBL (Translations EMBL), Swiss-Prot, and InterPro. It also takes a fist time initiation in Gene Ontology Annotation (GOA) project to provide assignments or checking of GO term to gene products for all organisms with completely sequence genome (Evelyn *el al.* 2003). Sequences of gene stored in this database are having two classes of data. Each and every sequence are apart from its core data, it have annotation with described protein structure, fucation post-translational modification and domains (Michal *et al.,* 2007). This database minimizes entries of protein number sequence to minimize redundancy. Each entry are provided link with other database so one can get more information about the protein which required. Many databases are builds on the basis of SWISS-PROT like UniProt, InterPro, and CluSTr (Nicola *et al.,* 2005, Amos *el a*t. 2005, and Margaret *et al.,* 2002) to provide the user a more precise view. In August 2002 the new release (Release 40.25) was contain 112,657 protein entries (Sarka *et al.,* 2000) and it jumps to 568,363 reviewed SWISS- PROT data of protein data were recorded till November 2022. Now there are bulk entries of several model organisms which represent about 40% of all on protein sequence. The easiest method to access TrEMBL and SWISS-PROT via Expert Protein Analysis System (ExPASy) is from http://www.expasy.org. A standard data mining algorithm is applied on SWISS-PROT to gain knowledge by keywords annotation in to the SWISS-PROT. 11 306 different rules were generated with taxonomy; signature matches and sequences (Philip *et al.,* 2008). An implement of integrated database system known as 'multi-protein survey system' (MPSS), that's provides a platform to retrieve information regarding many proteins. It integrates databases like TrEMBL, SwissProt, InterPro and PDB (Pei Hao *et al.,* 2005). The result can be browsed on http://golgi.ebi.ac.uk:8080/Spearmint/. Swiss-Prot release on March 08, 2011 contains 525997 sequence entries that comprising with 185874894 amino acid abstracts from 196176 different references. The large amount of data from several genome project give rise to TrEMBL which contain entries derived from the translation of coding sequences (CD. Another way to get SWISS-PROT and TrEMBL data is through anonymous ftp from ExPASy (ftp.expasy.org) by European Bioinformatics Institute (EBI) (ftp.ebi.ac.uk/pub/) with latest version which is released in 2020.

2. CluSTr: CluSTr is a Clusters of SWISS-PROT and TrEMBL database for the proteins offers an automatic classification of SWISS-PROT and TrEMBL protein into a group of similar proteins (http://www.ebi.ac.uk/clustr/). These clustering are based on analysis of pair-wise comparisons between proteins using the algorithm (Michal *et al.,* 2007). Protein analysis is carried out by different levels of similarity and hierarchical organization of clusters. While working with the clusters at different level of similarity, biological meaningful clusters selected for groups of proteins with increase the flexibility within database (Rolf *et al.,* 2001).

3. Universal Protein Resource (UniProt): Database is formed in 2002 with combination of, the Swiss-Prot +TrEMBL groups at Swiss Informatics Institute (SIB) and the Protein Identification Resources (PIR) grouped from Georgetown University and National Biomedical Research Foundation in joined forced as the UniProt consortiums, which maintain three DB layer.

The UniProt ARCHIVE (UniParc): It provides comprehensive, a stable, and non-redundant sequence collection by protein sequence data (Nicola *et al.,* 2005). UniParc is designed to capture all protein sequence data from the aforementioned database and from other data sources such as International Protein Index (IPI), Ensemble, RefSeq and FlyBase (Kersey *et al.,* 2004). UniParc represent each protein sequence once only and as it release in 2.6 version in September 2004 contained 4375775 unique sequences from 11978094 original source records. UniParc sequences version is increased each time the underlying sequence changes, making it possible to observed sequence changes in all sources of databases report can found at http:/www.uiprot.org.

The UniProt Knowledgebase (UniProt): A central database with merges annotation and functional information from Swiss-Prot, PIR-International Protein Sequence Database (PIR-PSD) and TrEMBL (Sarah *et al.,* 2009). All suitable PIR-PSD sequences missing from Swiss-Prot and TrEMBL where incorporated into UniProt. Bi-directional cross references were created to allow the easy tracing of PIR-PSD. The UniProt knowledgebase are divided into two parts first is a section of fully literature extraction and second curator evaluated computationally analyzed records. Automatic annotation and classification of UniProt/TrEMBL (for automatic annotation system) (Amos *el at.,* 2005), controlled Integration, vocabularies and ID mapping of new data sources can be found at https://www.uniprot.org/uniprotkb/P57727/entry.

The UniProt Reference (UniRef): This DB parodies non-redundant data collection relies on the UniProt Knowledgebase and UniParc to complete coverage of sequence at several resolution. Automatic procedure has been developed and creates three UniRef database like UniRef90, UniRef100 and UniRef50 (Amos *el a*t. 2005). Where UniRef50 and UniRef90 are build from UniRef100 using the CD-HIT Algorithm. A sample UniRef90 can found at https://www.uniprot.org/help/uniref.

4. PRIDE (Proteomic ID Entification Database): It was developed at the European Bioinformatics Institute (EBI) with Mass Spectral based proteomics

experiments. Because Mass Spectral is currently most preferred technology for the identification of protein (Lennart *et al.,* 2005, Philip *et al.,* 2008) and increased exponentially in last few years. PRIDE stores various kinds of information based on MS and MS/MS mass spectra, Protein Identification (IDs) and any associated meta data of protein (Juan *et al.,* 2010). It also able to capture details about post transcriptional modification found in peptides. Important features of PRIDE is allows data to be in private while anonymously sharing it with journal such as Nature Biotechnology, Nature method. PRIDE is tool for large scale data mining by using the PRIDE BioMart interface available at https://www.ensembl.org/biomart/martview/7b47dd44c579d289aebba4ebfab50a3a. It is possible to retrieve data from PRIDE with other sources. It is combined with database of UniProt (Amos *el a*t. 2005) and IPI database (Kersey *et al.,* 2004). Currently RIDE contains 16,208 Experiments, 4,779,159 Identified Proteins, 23,983,608 Identified Peptides, 3,260,938 Unique Peptides and 137,449,442 Spectra. Most important improvements in PRIDE is that, it is possible to submit files containing fragment, ion annotation on MS/ MS and visualized these annotations by 'PRIDE Spectrum Viewer'. The PRIDE team developed Microsoft Excel workbook which allows the required data to be collated in a series of relatively simple spreadsheets by automatic generation of PRIDE XML at the end of the process (Philip *et al.,* 2008). The more information about are available at https://www.ebi.ac.uk/pride/.

5. InterPro: InterPro is an integrated type of documentation resource for protein functional sites, families, its domains and the major protein signature databases into umbrella of one resource (Lennart *et al.,* 2005; Sarah *et al.,* 2009). These include the database such as PROSITE, Gene3D, PRINTS, TIGRFAMs, ProDom, SMART, PIRSF, PANTHER, SUPERFAMILY and SUPERFAMILY (Lennart *et al.,* 2005). The number of entries and coverage of protein space by InterPro is continuously increasing. The data release in 1999 contained 2423 entries, while the 2004 release of the database contain 11007 entries, that's fivefold increase in five years (Nicola *et al.,* 2005). The application of InterPro ranges of biologically important areas like automatic annotation of protein sequences and genome analysis (Margaret *et al.,* 2002). InterPro covered around 66% of all protein in TrEMBL and Swiss-Prot, and this has increased to over 90% for Swiss-Prot, 76% for TrEMBL for UniProt (Swiss-Prot and TrEMBL). InterPro covers 67 percent of the protein from complete genome database (Margaret *et al.,* 2002).

Major sequences databases like UniProtKB, UniMES and UniParc signature. The number of proteins matching signatures from InterPro and those matching the full set of member database signatures are shown below in Table 2 (Sarah *et al.,* 2009).

Table 2: Different database comparing number of proteins, number of matches, and member database signatures with InterPro.

Sequence	Proteins numbers in database	Number of proteins >0 Matches to InterPro	Protein number with >0 Matches combined Member Database signatures
UniProtKB or Swiss-Prot	397539	369830(93.0%)	379897(95.6%)
UniProtKB/TrEMBL	6212793	4628221(74.5%)	4894258(78.8%)
UniProtKB (Total)	6610332	4998051(75.6%)	5274155(79.8%)
UniParc	17718252	12211006(68.9%)	13290858(75.0%)
UniMES	6028191	4132464(68.6%)	4461935(74.0%)

The InterPro having cross-references with various databases which shows huge data crossing entries is described in Table 3. Many features have been added into the InterPro database since publication in Nucleic Acids Research in year 2003. This includes InterPro Domain Architectures Viewer, additional protein match views, the, taxonomic range information, additional database for 3D structure information (Philip *et al.,* 2008). The release of InterPro version 31.0 is on 9th February 2011 with 21185 entries, representing active sites 97, binding sites 65, conserved sites 615, domains 5936, family 14194, PTM 262 and InterPro cites 32331 publications in PubMed. The recent update of this database with version InterPro 92.0 in 2022 gives additional feature with 291 InterPro entries, 359 new methods integration and 38349 21185 entries. This database is available at https://www.ebi.ac.uk/interpro/release_notes/ .

Table 3: Number of InterPro and cross-references.

Type of Database		InterPro Entries
i.	UniProtKB	13131
ii.	BLOCKS	6134
iii.	CAZy	119
iv.	COMe	204
v.	IntEnz	2336
vi.	IUPHAR receptor	113
vii.	MEROPS	548
viii.	PROSITE doc	1479
ix.	Pfam Clans	1544
x.	PANDIT	7702
xi.	CluSTr	6818
xii.	IntAct	135
xiii.	GO	7131
xiv.	MSD site	1313
xv.	PDB	68021
xvi.	SCOP	6537
xvii.	CATH	6212

6. ProToMap: Is a completely automated method of protein classification by its sequence and searches for similarities by detecting group of homologous proteins and high level structures and group of related clusters. ProToMap not use multiple sequence alignments but need to do algorithm for performance of analysis. The first step for the identification of groups related clusters belongs to same group for strongly connected. ProToMap is strongly connected with the SWISS-PROT and is based on the analysis of the database protein sequence. The analysis based on comparisons of SWISS-PROT and TrEMBL and can accessible https://protomaps.com/. There are many search methods available, from simple keyword through accession number to name of the protein. Essential method is that we can classify protein sequence by submitting and comparing it with existing protein clusters (Michal *et al.,* 2007).

7. International Protein Index (IPI): It was launched in 2001 and it provides better guide to the main databases which describe the proteomes in higher eukaryotic organisms. Important features of International Protein Index are to maintain a database of cross references between the primary and secondary data sources, this provides minimally redundant were that's allow the tracking of sequences in IPI between International Protein Index releases. IPI was updated each month in accordance with the latest data released by the primary data sources till 2011(Lennart *et al.,* 2005). The following Table 4 describes the IPI version (Arabidopsis, 3.77), released on Tue, 18 Jan 2011as in Table 4 and currently the database in merge with UniProt Knowledge Base, Ensembl and Ensembl Genomes projects.

Table 4: Total number of IPI entries and Compositions of IPI entries in *Arabidopsis.*

A.	**IPI entries**	**Numbers**
i.	Number of entries in IPI	37064
ii.	Entries referenced by IPI	118296
iii.	References to UniProtKB	51548
iv.	References to RefSeq	33402
v.	References to TAIR Protein	33346
B.	**Composition of IPI**	**Numbers**
A.	**IPI entries**	
i.	Pointing only to UniProtKB	5805
ii.	Pointing only to RefSeq	12
iii.	Pointing only to UniProtKB, RefSeq	84
iv.	Pointing only to TAIR Protein	24
v.	Pointing only to UniProtKB, TAIR Protein	19
vi.	Pointing only to RefSeq, TAIR Protein	493
vii.	Pointing to UniProtKB and RefSeq and TAIR Protein	30631

For identification and searching IPI in a database identifier (e.g. P50238, IPI00015171 or ENSP00000332449) are to be mentioned to retrieve matching entries from all current IPI dataset's. IPI history search are another features which helps to track deleted and secondary structure identifiers (e.g. IPI00030830) from past version IPI database. The more information can available on ftp://ftp.ebi.ac.uk/pub/ebi/ftp/pub/databases/IPI.

8. RESID: RESID is produced by John S Garovelli. Is a structural database of protein with comprehensive collection of annotations, structures for protein, post translational modifications including C terminal and N terminal modification and peptide chain cross link modifications. It includes systematic and frequently atomic formulas, observed alternate names, and weights, keywords, enzyme activities, taxonomic range, literature citations with database cross-references, molecular models and structural diagrams. The RESID was constructed in 1993 and first time publically released in 1995 with 181 entries (John *et al.,* 1999) where in 1998 it was made available on the web site first with graphical model components. In 2000 production of the RESID database of protein structure modifications also started at the Advanced Biomedical Computing Center of the National Cancer Institute Frederick (John 2004).

RESID is the only publicly available database contains more than 260 structural and regulatory modifications, active site prosthetic modifications, visual display and molecular models of post-translational modifications. RESID Database maintains concurrent cross references to the PIR International Protein Sequence Database, the Chemical Abstracts (CAS), the MEDLINE citation database and PDB (John 2004). This can be searched by entry code or other unique identifier, citation, by name, keyword or feature text search with molecular weight search, or from selection lists based on encoded amino acids. The RESID frequently updated and distributed on at http://www.ebi.ac./RESID/.

9. Phytoprot: Is a advance database of plant proteomics evolved to the number of protein which increases regularly in plant protein sequence, it can promote study and production of functions of different plant base protein, differentially orthologues forms from paralogues, delineating characteristics sequences, building phylogenetic reconstructions and also functional annotation can predict in that clustering. In these clustering databases the complete proteome of *Arabidopsis thaliana* and all the available sequence from the other databases are grouped and build an clusters which make a database called as PHYTOPROT (Mohseni-Zadeh *et al.,* 2004). The protein sequence of *Arabidopsis thaliana* are retrieved from EMBL proteome site http://www.ebi.ac.uk/proteome and other

plants protein sequence information from TrEMBL databases and SWISS-PROT for plant based protein. Before the end of 2003 a new sets of comparisons will be to PHYTOPROT which consist of the *Arabidopsis thaliana* proteome compared against it, which should be useful for the study of the numerous multigenic families in protein plant kingdom.

CONCLUSIONS

Recent developments in protein research databases such as SWISS-PROT, CluSTr, Universal Protein Resource, UniRef, PRIDE, InterPro, ProToMap, International Protein Index, RESID, and PHYTOPROT, have expanded the scope of protein research. These databases provide crucial information on protein structure and function, and have become a valuable resource for protein identification, prediction, and annotation. UniProt has been recently updated to provide more detailed information on protein isoforms and their functions, as well as protein-protein interactions (The UniProt Consortium 2021). In addition, InterPro has been updated to provide more accurate and comprehensive protein domain predictions, while RESID has been updated to include new post-translational modifications (Mitchell, 2019, Huang, 2020). Moreover, new databases have emerged, such as the Protein Data Bank (PDB), which provides high-resolution 3D structures of proteins, and the Human Protein Atlas (HPA), which provides detailed information on the expression and localization of human proteins (Uhlén, 2015 and Burley, 2021). The availability of these databases has led to significant advancements in protein research, and has accelerated the discovery of new drug targets and biomarkers for diseases. In conclusion, these databases continue to play an important role in facilitating protein research and advancing our understanding of the biological world (Burley, 2021).

Authors' contributions. AK, KM and RL are designed concept designed for review topic; were AK, MH and SV provided the preliminary draft of manuscript. AK, KM and RL analyzed and performed statistical analysis of various results given in ligature. AK and SV are supersized the work with its authentication. All authors are carefully read and approved the final manuscript.

Acknowledgements. Authors would like to acknowledge thanks to the to the Principal of the Jijamata College Nandurbar, Head Department of Botany Amravati University Amravati, Principal and staff of the Anjum Islam Janjira Degree College of Science Murud-Janjira, Principal and staff of G.C.U.B. Science College, Goregaon and Principal A. C. S College Taloda for their kind supports during the study of this article.

Conflict of interest. None.

REFERENCES

Aabid Khaliq Tantray, Naveena Nazim, Mohd Shahid Raja, Mohd Farooq Baqual and Seerat Showkat (2021). Quantitative Determination of Protein and Carbohydrate Contents in Fruits of different Mulberry Varieties. *Biological Forum – An International Journal, 13*(4), 1307-1310.

Aebersold, R, Mann M (2016). Mass-spectrometric exploration of proteome structure and function. *Nature, 537*, 347-55.

Agrawal, G. K., Sarkar, A., Righetti, P. G., Pedreschi, R., Carpentier, S. and Wang, T. (2013). A decade of plant proteomics and mass spectrometry: translation of technical advancements to food security and safety issues. *Mass Spectrom Rev. 32*, 335-65.

Akhilesh Pandey and Matthias Mann (2000). Review article Proteomics to study genes and genomes, *Nature, 405*, 837-846.

Alban, A., David, S. O., Bjorkesten, L., Anderssone, C., Sloge, E., Lewis, S. and Carrie, I. (2003). A novel experimental design for comparative two dimensional gel analyses: Two dimensional deference gel electrophoresis incorporating a pooled internal standard. *Proteomics, 3*, 34-44.

Alseekh, S., Kostova, D., Bulut, M., Fernie, A. R. (2021). Genome-wide association studies: assessing trait characteristics in model and crop plants. *CMLS. 78*, 57 43-54.

Amos Bairoch, Rolf Apweiler, Cathy, H. Wu, Winona, C. Barker, Brigitte Boec kmann, Serenella Ferro, Elisabeth Gasteiger, Hongzhan Huang, Rodrigo Lopez Michele Magrane Maria, J. Martin, Darren, A. Natale, Claire, O., Donovan Nicole, and Redaschiand Lai-SuL. Yeh (2005). The Universal Protein Resource (UniProt). *Nucleic Acids Research*, *33*, 154-159.

Andrej Shevcenko, Ole, N. Jensen, Alexander, V. Podtelejnikov, Francis Sagiiocco, Matthias Wilm, Ole Vorm, Peter Mortensen, Anna Shevchenko, Helian Boucherite and Matthias Mann (1996). Linking genome and proteome by mass spectrometry: Large-scale identification of yeast protein from two dimensional gels. *Proc. Natl. Acad. Sci., 93*, 14440-14445.

Arsi, T. Rosengren, Tuula, A. Nyman, Saija Syyrakki, Sampsa Matikainen and Ritta Lahesmaa (2005). Proteomic and trnascriptomics characterization of interferon-α-induce human primary T helper cells. *Proteomics, 5*, 371-379.

Barber, M. Bordoli, R., S., Sedgwick, R. D. and Tyler, A. N. (1981). Fast atom bombardment of solids (FAB): a new ion source for mass spectrometry. *J Chem. Soc. Chem., 21,* 325-327.

Beniddir, M. A., Kang, K. B., Genta-Jouve G., Huber, F., Rogers, S., van der Hooft J. J. J. (2021). Advances in decomposing complex metabolite mixtures using substructure- and network-based computational metabolomics approaches. *Nat Prod Rep.,38*, 19 67-93.

Bilal Aslam, Madiha Basit, Muhammad Atif Nisar, Mohsin Khurshid, Muhammad Hidayat Rasool (2017). Proteomics: Technologies and Their Applications. *Journal of Chromatographic Science*, *55*(2), 182-196.

Brigit Schilling, James Murrey, Chris, B. Yoo, Richard, H. Row, Michael, P., Cusack, Roderick, A. Capaldi, and Bradford, W. Gibson (2006). Proteomic analysis of succinate dehydrogenase and ubiquinol-cytochrome c reductase (Complex II and III) isolated by immunoprecipitation from bovine and mouse heart mitochondria. *Biochimica et Biophysica Acta., 1762,* 213-222.

Burley, S. K., Berman, H. M., Bhikadiya, C., Bi, C., Chen, L., Di Costanzo, L., and Zardecki, C. (2021). RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental

biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Research, 49*(D1), D437-D451.

Ch. Muhammad Ishtiaq, Q. Ile, J. P. Huang, Yi. Wang, P.G. Xiao and Yi Yu Cheng (2007). Biosystematics and Plant Proteomics: Role of Proteomics in Plant Phylogenetic Analysis. *Pakistan Journal of Biological Sciences, 10*(20), 3487-3496.

Cho, K., Schwaiger-Haber M., Naser F. J., Stanclife E., Sindelar M. and Patti G. J. (2021). Targeting unique biological signals on the fy to improve MS/MS coverage and identification efficiency in metabolomics. *Anal Chim Acta., 5*(11), 33-42.

Chowdhur, S. K., Katta, V. and Chait, B. T. (1999). Probing Conformational Changes in Protein by Mass Spectrometry. *J. Am. Chem. Soc.*, *112*, 9012-9013.

Clair, M. Delahunty and John R. Yates III. (2007). MudPIT: multidimensional protein identification technology. *BioTechniques*, *43*, 563-569.

Clark, N. M., Elmore J. M. and Walley J. W. (2021). To the proteome and beyond: advances in single-cell omics profiling for plant systems. *Plant Physiol., 188,* 726-37.

Cole, B. S. and Moore J. H., (2018). Eleven quick tips for architecting biomedical informatics workflows with cloud computing. *PLoS Comput Biol., 14*(3), 89-94.

David, C. Schriemer and Liang, Li. (1996). Detection of high Molecular Weight Narrow Polydisperse Polymers up to 1.5 Million Dalton by MALDI Mass Spectrometry. *Anal. Chem., 68*, 2721-2725.

David Eenyo, Jun Qin, and Brain T. Chait (1998). Protein identification using mass spectrometric information. *Electrophoresis*, *19*, 998-1005.

Dennis, J. Selkoe. (2003). Folding proteins in fatal ways. *Nature, 426*, 900-904.

Dong, Y., Sonawane, P, Cohen, H., Polturak, G, Avivi, S. and Rogachev, I. (2020). High Mass Resolution, Spatial Metabolite Mapping Enhances the Current Plant Gene and Pathway Discovery Toolbox. *New Phytol., 228*,1986-2002.

Dustin Lippert, Jun Zhuang, Steven Ralph, Dave, E. Ellis, Margarita Gilbert, Robert Olafson, Kermit Ritland, Brian Ellis, Carl, J. Dougals and Jorg Bohlmann (2005). Proteome analysis of early somatic embryogenesis in *Picea glauca*. *Proteomics, 5*, 461-473.

Elena Silvestri, Maria Moreno, Lulgi Schiavo, Pleter de Lande, Assunta Lombardi, Angela Chambery, Augusto Parente, Antonla Lanni and Fernando Goglia (2006). A Proteomic Approach to Identify Protein Expression Changes in Rat Liver Following Administration of 3, 5, 3'-Triido-L-thyronine. *Journal of Proteome Research*, *5*, 2317-2327.

Enrique Calvo, M. Graciela Pucciarelli, Helene Bierne, Pascale Cossart, Juan Pablo Albar, and Francisco Garcia-del Portillo (2005). Analysis of the Listeria cell wall proteome by two-dimensional nanoliquid chromatography coupled to mass spectrometry. *Proteomics*, *50*, 433-443.

Eugene, C. Yi, Xiao-jun Li, Kelly Cooke, Hookeun Lee, Beian Raught, Anrew Page, Victoria Aneliunas, Phil Hieter, David, R. Goodle, and Ruedi Aebersold (2005). Increased quantitative proteome coverage with C_{13}/C_{12}-based, acid-cleavable isotope-coded affinity tag reagent and modified data acquisition scheme. *Proteomic, 5*, 380-387.

Evelyn Camon, Michele Magrane, Daniel Barrell, David Binns, Wolfgang Fleischmann, Paul Kersey, Nicola Mulder, Tom Oinn John Maslen, Anthony Cox, and Rolf Apweiler (2003). The Gene Ontology Annotation (GOA) Project: Implementation of GO in SWISS-PROT, TrEMBL, and InterPro. *Genome Research*, *13*, 662-672.

Ewa Urbanczyk-Wochniak, Alexander Luedemann, Joachim Kops, Joachim Selbig, Ute Roessner-Tunali, Lothar Willimitzer and Alisdair R. Fernie (2003). Parallel analysis of transcript and metabolic profile: a new approach in system biology. *EMBO reports*, *4*, 989-993.

Faraji M., Yamini Y. and Gholami M., (2019). Recent Advances and Trends in Applications of Solid-Phase Extraction Techniques in Food and Environmental Analysis. *Chromatographia*, *82*, 12 17-49.

Farshid Golshani, Barat Ali Fakheri, Elham Behshad and Roghayeh Mohammadpour Vashvaei (2015). PRs proteins and their Mechanism in Plants. *Biological Forum – An International Journal, 7*(1), 477-495.

Fatemeh Maghuly, Gorji Marzban and Margit Laimer (2009). Functional Genomics of Allergen Gene Families in Fruits. *Nutrients*, *1*, 119-132.

Fernando Garcia-Alcalde, Federico Garcia-Lopeze, Joaquin Dopazo and Ana Conesa (2011). Paintomics: a web based tool for the joint visualization of trnascriptomics and metabolomics data. *Bioinformatics Advance Access published, 27*(1), 137-139.

Francisco, M. Canovas, Eliane Dumas-Gaudot, Ghislaine Recorbet, Jesus Jorrin, Hans-Peter Mock and Michel Rossignol (2004). Plant proteome analysis. *Proteomics, 4*, 285–298.

Fred, E. Cohen and Jeffery, W. Kelly (2003). Therapeutic approaches to protein-misfolding diseases. *Nature, 426*, 905-909.

Ganesh Kumar Agrawal, Jacques Bourguignon, Norbert Rolland, Genevieve Ephritikhine, Myriam Faerro, Michel Jaquinod, Kondtantinos, G. Alexious, Thierry Chardot, Niranjan Chakraborty, Pascale Jolivet, John, H. Doonan and Randeep Rakwal (2011). Plant Organelle Proteomics: Collaborating for Optimal Cell Function. *Published online in Wiley Online Library, 30,* 772-785.

Ganesh Kumar Agrawal, Martin Hajduch, Katherine Graham and Jay, J. Thelen (2008). In-Depth Investigation of the Soybean Seed-Filling Proteome and Comparison with a Parallel Study of Rapeseed. *Plant Physiology, 148*, 504-518.

Giovanni Renzone, Anna Maria Salzano, Simona Arena, Chiara, D. Ambrosio, and Andrea Scaloni (2007). Mass Spectrometry-Based Approach for Structural Studies on Protein Complexes at Low-Resolution. *Current Proteomics, 4*, 1-16.

Gisele, G. Tomazella, Idalete da Silva, Helen, J. Laure, Jose, C. Rosa, Roger Chammas, Harald, G. Wiker, Gustavo, A. de Souza and Lewis, J. Greene (2009). Proteomic analysis of total cellular proteins of Human neutrophils. *Proteome Science*, 7-32.

González-Riano, C., Dudzik, D., Garcia, A., Gil-de-la-Fuente, A., Gradillas, A., Godzien, J., (2020). Recent Developments along the Analytical Process for Metabolomics Workfows. *Anal Chem*., *92,* 203-26.

Heinemerey, J., Eubel, H., Wehmhnoer, D., Jansch, L. and Braun, H. P. (2004). Proteomic approach to characterize the supramolecular organization of photosystems in to higher plants. *Phytochemisry, 65*, 1683-1692.

Helen, M., Berman, Philip, E. Bourne and John Westbrook (2004). The Protein Data Bank: A Case Study in Management of Community Data. *Current Proteomics*, *1*, 49-57.

Heng Zhu, Metin Bilgin and Michal Snyder (2003). PROTEOMICS. *Annu. Rev. Biochem., 72*, 783-812.

Himanshu Dubey and Anil Grover (2001). Current initiatives in proteomics research: The plant perspective. *Current Science, 2*(80), 262-269.

Holly, D. Coughenour, Reggie, S. Spaulding and M. Thompson (2004). The synaptic vesicle proteome: A comparative study in membrane protein identification. *Proteomics, 4*, 3141-3155.

Huang, H., McGarvey, P. B., Suzek, B. E., Mazumder, R., Zhang, J., Chen, Y. and Wu, C. H. (2020). UniProtKB/Swiss-Prot: the manually annotated section of the UniProt Knowledge Base in 2020. *Nucleic Acids Research, 48*(D1), D570-D576.

Imanzadeh Reza, Soori Rahman and Faraji Gholamreza (2015). The Effect of Intermittent Training on C-reactive Protein in Young Women. *Biological Forum – An International Journal, 7*(1), 625-628.

Issa Coulibaly and Grier, P. Page (2008). Bioinformatics Tools for Inferring Functional Information from Plant Microarray Data II: Analysis Beyond Single Gene. *International Journal of Plant Genomics,* Article ID 893941, 11-13.

J. Kose and U. Kobalz (1995). Two-dimensional Electrophoresis of Protein: An Updated Protocol and Implications for Functional Analysis of the Genome. *Electrophoresis, 16*, 1034-1059.

J. V. Jorrin, D. Rubiales, E. Dumas-Gaudot, G. Recorbet, A. Maldonado, M.A. Castillejo and M. Curto (2006). Proteomics: a promising approach to study biotic interaction in legumes. A review. *Euphytica, 147*, 37-47.

John, R. Yates, Cristian, I. Ruse, and Aleksey Nakorchevsky (2009). Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. *Annu. Rev. Biomed., 11*, 49-79.

John, S. Garavelli (1999). The RESID Database of protein structure modifications. *Nucleic Acids Research, 27*, 198-199.

Jon B. Klein and Visith Thongboonkerd (2004). Overview of Proteomics. Proteomics in Nephrology. *Contrib Nephrol. Basel, Karger, 141*, 1-10.

Juan Antonio Vizcaıno, Richard Cote, Florian Reisinger, Harald Barsnes, Joseph, M. Foster, Jonathan Rameseder, Henning Hermjakob and Lennart Martens (2009). The Proteomics Identifications database: 2010 update. *Nucleic Acids Research*, *38*, 738-742.

Kaley Morris, Marce, D. Lorenzen, Yasuaki Hiromasa, John, M. Tomich, Cris Oppert Elena, N. Elpidina, Konstantin Vinokurov, Junan Luis Jurat-Funentes, Jeff Fabrick, and Brenda Oppert (2009). Tribolium castaneum Larval Gut Transcriptome and Proteome: A Resource for the Study of the Coleopteran Gut. *Journal of Proteome Research*, *8*, 3889-3898.

Kathryn, S. Lilley and Paul Dupree (2007). Plant organelle proteomics. *Current Opinion in Plant Biology*, *10*, 594-599.

Kenneth, D. Greis, Songtao Zhou, Thomas, M. Burt, Andrew, N. Carr, Elizabeth Dolan, Vijayasurian Easwaran, Artem Evdokimov, Richard Kawamoto, Jeff Roesgen, and Gregory, F. Davis (2006). MALDI-TOF MS as a Label-Free Approach to Rapid Inhibitor Screening. *J Am Soc Mass Spaectrom, 17*, 815-822.

Kersey, P. J., Duarte, J., Williams, A., Karavidopoulou, Y., Birney, E. and Apweiler, R. (2004). The International Protein Index: An integrated database for proteomics experiments. *Proteomics 4*(7), 1985-1988.

Kris Gevaert and Joel Vandekerckhove (2000). Protein identification methods in proteomics. *Electrophoresis, 21*, 1145-1154.

Kumar, A., Srivastava, N. C., Singh, V. P., & Sunder, J. (2014). Electrophoretic analysis of Indian isolates of Mycoplasma agalactiae and Mycoplasma bovis by SDS-PAGE and immunoblotting. *Veterinary Medicine International*, *2014*.

Lennart Marten, Henning Hermjakob, Philip Jones, Marcin Adamski, Chris Taylor, David States, Kris Gevaert, Joël Vandekerckhove and Rolf Apweiler (2005). PRIDE: The proteomics identifications database. *Proteomics, 5*, 3537-3545.

Liu, X., Qin, D., Piersanti, A., Zhang, Q., Miceli, C. and Wang, P. (2020). Genome-wide association study identifes candidate genes related to oleic acid content in soybean seeds. *BMC Plant Biol*., *20*, 39-43.

Luis Berrade, Angia, E. Garcia, Julio, A. Camarero (2011). Protein Microarrays: Novel Developments and Applications. *Pharm Res.* (7) 815-822.

Lv, Q. Q., Yang, X. N., Yan, D. M., Liang, W. Q., Liu, H. N. and Yang, X. W. (2017). Metabolic profiling of dehydrodiisoeugenol using xenobiotic metabolomics. *J Pharm Biomed Anal*., *145* 7 25-33.

Lv, W., Wang, L., Xuan, Q., Zhao, X., Liu, X., Shi, X., & Xu, G. (2020). Pseudotargeted method based on parallel column two-dimensional liquid chromatography-mass spectrometry for broad coverage of metabolome and lipidome. *Analytical chemistry*, *92*(8), 6043-6050.

Marcus Bantscheff, Markus Schirle, Gavain Sweetman, Jens Rick and Bernhard Kuster (2007). Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Cham., 389*, 1017-1031.

Margaret Biswas, John, F., OÕRourke, Evelyn Camon, Gill Fraser, Alexander Kanapin Youla Karavidopoulou, Paul Kersey, Evgenia Kriventseva, Virginie Mittard, Nicola Mulder, Isabelle Phan, Florence Servant and Rolf Apweiler (2002). Application of InterPro in protein annotation and genome analysis. *Briefings in Bioinformatics, 3*, 285-295.

Maria Pleguezuelo, Laura, M, Lopez-Sanchez, Antonio Rodriguez-Ariza, Jose, L. Montero, Javier Briceno, Ruben Ciria, Jordi Muntane, and Manuel de la Mata (2010). Proteomic analysis for developing new biomarkers of hepatocellular carcinoma. *World Journal of Hepatology*, *2*(3), 127-135.

Martin, H. Maurer (2011). Proteomic Definitions' of Mesenchymal Stem Cells. *Stem Cells International*, 1-9.

Martin Hajduch, Ashwin Ganapathy, Joel, W. Stein, and Jay, J. Thelen (2005). A Systematic Proteomic Study of Seed Filling in Soybean. Establishment of High-Resolution Two-Dimensional Reference Maps, Expression Profile, and an Interactive Proteome Database. *Plant Physiology*, *137*, 1397-1419.

Melby, J. A., Roberts, D. S., Larson, E. J., Brown, K. A., Bayne, E. F. and Jin, S. (2021). Novel Strategies to Address the Challenges in Top-Down Proteomics. *J Am Soc Mass Spectrom*, *32*(12), 78-94.

Melissa, M. Dix, Gabriel, M. Simon, and Benjamin, F. Cravatt (2000). Global Mapping of the Topography and Magnitude of Proteolytic Events in Apoptosis. *The Cell, 134*, 679-691.

Michael, J. Dutt and Kelvin, H. Lee. Proteomic analysis. *Current Opinion in Biotechnology*, *11*, 176-179.

Michal Pist, Irena Kustrzeba-Wojcicka, Malgorzata Matusiewicz, Malgorzata Krzystek-Koepocka, and Teresa Banas (2007). Bioinformatics: From arduous

beginning to molecular database. *Adv. Chin Exp Med.*, *16*, 85-93.

Mitchell, A. L., Attwood, T. K., Babbitt, P. C., Blum, M., Bork, P., Bridge, A., and Finn, R. D. (2019). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Research, 47*(D1), D351-D360.

Nayuni Vandana, Esther Magdalene Sharon M., Sivakumar G. M. and Geetha P. (2022). Optimization of Whey Protein Concentrate-Natural Deep Eutectic Solvents Based Edible Coating of Paneer. *Biological Forum – An International Journal, 14*(4a), 551-558.

Ngan, F. Huang, Kyle Kurpinski, Qizhi Fabg, Randall J. Lee, and Song Li (2011). Proteomic identification of biomarkers of vascular injury. *Am J Transl Res., 3*(2) 139-148.

Nicola, J. Mulder, Rolf Apweiler, Teresa, K. Attwood, Amos Bairoch, Alex Bateman, David Binns, Paul Bradly, Peer Bork, Phillip Bucher, Lorenzo Cerutti, Richard Copley, Emmanuel Courcelle, Ujjwal Das, Richard Durbin, Wolfgang Fleischmann, Julian Gough, Daniel Haft, Nicola Harte, Nicolas Hulo, Daniel Kahn, Alexander Kanapin, Maria Krestyaninova, David Lonsdale, Rodrigo Lopez, Ivica Letunic, Martin Madera, John Maslen, Jennifer McDowall, Alex Mitchel, Anastasia N., Nikolskaya, Sandra Orchard, Marco Pagni, Chris P. Ponting, Emmanuel Quevillon, Jeremy Selengut, Christian, J., A. Sigrist, Ville Silventoinen, David, J. Studholme, Robert Vaughan and Cathy, H. Wu (2005). InterPro, Progress and status in 2005. *Nucleic Acids Research, 33*, 201-205.

Nutan Moharana, Bansidhar Pradhan, Krishnendu Chattopadhyay, Shuvendu Shekhar Mohapatra, Tapash Kumar Mishra, Mihir Ranjan Mohanty and Rajesh Kumar Kar (2022). Analysis of Marker-trait Association of Backcross Population in Rice using SSR Markers for High Grain Protein Content. *Biological Forum – An International Journal*, *14*(4a), 656-661.

Ohkmae, K. Park (2004). Proteomic Studies in Plants. *Journal of Biochemistry and Molecular Biology*, *37*, 133-138.

Ole Norregaard Jensen, Martin, R. Larsen and Peter Roepstorff (1998). Mass Spectrometric Identification and Microcharacterization of Proteins from Electrophoretic Gels: Strategies and Applications. PROTEINS: Structure, *Function and Genetics Suppl*., *2*, 74-89.

Palak, P. Vaidh, Pravin, M. Lunagariya, Jigar, H. Patel, Poojan, H. Pandya, Malhar, R. Khant, Niti Sharma and Kishan, N. Wadhwani (2022). Growth, Reproductive Performance, and Health Attributes of Crossbred Heifers fed High Plane of Energy and Protein. *Biological Forum – An International Journal, 14*(4): 579-585.

Patton, W. (2002). Detection technologies in proteome analysis. *Journal of Chromatography,* (771) 3-31.

Pei Hao, Wei-Zhong He, Yin Huang, Liang-Xiao Ma, Ying Xu, Hong Xi, Chuan Wang, Bo-Shu Liu, Jin-Miao Wang, Yi-Xue Li, and Yang Zhong (2005). MPSS: an integrated database system for surveying a set of proteins. *Bioinformatics*, *21*(9), 2142-2143.

Pouya Yari , Akbar Yaghobfar , Habib Aghdam Shahryar , Yahya Ebrahimnezhad and Sara Mirzaie Goudarzi (2015). Effects of Diets Formulation Based on Different Methods of Metabolisable Energy (AMEn & TMEn) and Amino Acids (TAA & DAA) Expression on Performance, Energy and Protein Efficiency Ratio and Productive Efficiency Factor of Broiler Chicks. *Biological Forum – An International Journal 7*(1): 73-78.

Pranav Wadje, Akshay R. Patil and R. Meenatchi (2022). Effect of pre-treatments on proximate composition, protein extraction yield and extraction rate of Superworm protein. *Biological Forum – An International Journal, 14*(3), 288-294.

Qui, Q. S., Huber J. L., Booker, F. L., Jain V., Leakey A. D., Fiscus E. L., Yau, P. M., Ort, D. R. and Huber, S. C. (2008). Increased protein carbonylation in leaves of Arabidopsis and Soybean in response to elevated (CO_2). *Photosynth Resh*, *97*, 155-166.

Ramu, S. Saravanan and Jocelyn, K. C. Rose (2004). A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant pant tissues. *Proteomics, 4*, 2522-2532.

Raquel González-Fernández, Elena Prats, and Jesús V. Jorrín-Novo (2010). Review Article Proteomics of Plant Pathogenic Fungi. *Journal of Biomedicine and Biotechnology*, 1-36.

Richard Kramer and Dalia Cohen (2010). Functional Genomics to new drug targets. *Nature Reviews Drug Discovery, 3*, 965-972.

Robert Hall, Mike Beal, Oliver Fiehn, Nigel Hardy, Lloyd Sumner and Raoul Bino (2002). Plant Metabolomics: The Missing Link in Functional Genomics Strategies. *Plant Cell, 14*, 1437-1440.

Rolf Apweiler, Margaret Biswas, Wolfgang Fleischmann, Alexander Kanapin, Youla Karavidopoulou, Paul Kersey, Evgenia, V. Kriventseva, Virginie Mittard, Nicola Mulder, Isabelle Phan and Evgeni Zdobnov (2001). Proteome Analysis Database: online application of InterPro and CluSTr for the functional classification of proteins in genomes. *Nucleic Acids Research*, *29*, 44-48.

Roman Kaliszan, Tomasz Baczek, Anna Cimochowska, Paulina Juszczyk, Kornelia Wisniewska and Zbigniew Grzonka (2005). Prediction of high-performance liquid chromatograpy retention of peptides with the use of quantitative structure-retain relationships. *Proteomics, 5*, 409-415.

Rosenberg, J. M. and Utz, P. J. (2015). Protein microarrays: a new tool for the study of autoantibodies in immunodeficiency. *Frontiers in Immunology, 6*, 138.

S. P. Gygi (1999). Quantitative Analysis of Complex Protein Mixtures Using Isotop-coded Affinity Tags. *Nature Biotechnology*, *17*, 994-999.

S. Mohseni-Zadeh, A. Louis, P. Brezellec and J. L. Risler (2004). PHYTOPROT: a database of clusters of plant proteins. *Nucleic Acids Research*, *32*, 351-353.

Saligrama, A. Deepak, Junko Shibato, Hideo Ishii, Yoko Ogawa, Yasukazu Yoshida, Hitoshi Lwahashi, Yoshinori Masuo, Ganesh Kumar Agrawal, and Randeep Rakwal (2008). Proteomics Approach for Investigating the Disease Resistance Using Cucumber as a Model Plant. *American Journal of Biochemistry and Biotechnology, 4*(3), 231-238.

Sarah Hunter, Rolf Apweiler, Teresa K. Attwood, Amos Bairoch, Alex Bateman, David Binns, Peer Bork, Ujjwal Das, Louise Daugherty, Lauranne Duquenne, Robert D. Finn, Julian Gough, Daniel Haft, Nicolas Hulo, Daniel Kahn, Elizabeth Kelly, Aurelie Laugraud, Ivica Letunic, David Lonsdale, Rodrigo Lopez, Martin Madera, John Maslen, Craig McAnulla, Jennifer McDowall, Jaina Mistry, Alex Mitchell, Nicola Mulder, Darren Natale, Antony, F. Quinn, Jeremy, D. Selengut , Christian, J. A. Sigrist, Manjula Thimma, Paul, D. Thomas, Franck Valentin, Derek Wilson, Cathy, H. Wu, and Corin Yeats (2009).

InterPro: the integrative protein signature database. *Nucleic Acids Research*, *37*, D211-D215.

Sarka Beranova-Giorgianni (2003). Proteome analysis by two-dimensional gel electrophoresis and mass spectrometry: strength and limitations. *Trends in Analytical Chemistry*, *22*, 273-281.

Satyanand Tyagi, Raghvendra, Usha Singh, Taruna Kalra, Kavita Munjal, and Vikas (2010). Practical Application of Proteomics-A Technique for Large-Scale Study of Proteins: An Overview. *International Journal of Pharmaceutical Science, 3*, 86-90.

Scott, D. Patterson (2000). Proteomics: the industrialization of protein chemistry. *Current Opinion in Biotechnology*, *11*, 413-418.

Setsuko Komatsu (2008). Plasma membrane proteome in Arabidopsis and rice. *Proteomics, 8*, 4137-1445.

Stephen B. H. Kent (1988). Chemical Synthesis of Peptides and Proteins. *Ann. Rev. Biochem*, *57*, 957-989.

The UniProt Consortium. (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research*, *49*(D1), D480-D489.

Timp ,W., Timp, G. (2020). Beyond mass spectrometry, the next step in proteomics. *Sci Adv., 6*(2), 89-78.

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., and Pontén, F. (2015). Tissue-based map of the human proteome. *Science*, *347*(6220), 1260419.

Unlu, M., Morgan, M. E. and Minden, J. S. (1997). Differential gel electrophoresis: A single gal method for detecting changes in protein extracts. *Electrophoresis*, *18*, 2071-2077.

Waltraud, X. Schulze and Bjorn Usadel (2010). Quantitation in Mass-Spectrometry-Based Proteomics. *Annu. Rev. Plant Biol.*, *61*, 491-516.

Wang, Z. Q., Xu, X. Y., Gong, Q. Q., Xie, C., Fan, W., Yang, J. L., (2014). Root proteome of rice studied by iTRAQ provides integrated insight into aluminum stress tolerance mechanisms in plants; *Journal of Proteomics, 98*, 189-205.

Wenzhu Zhang, Andrew J. Czernik, Tom Yungwirth, Ruedi Aebersold and Brian, T. Chait (1994). Matrix-assisted laser desorption mass spectrometric peptide mapping of proteins and separated by two-dimensional gel electrophoresis: Determination of phosphorylation in synapsin I. *Protein Science, 3*, 677-686.

Wilkins, M. R. Williams, K. L. Apple, R. D. and Hochstrasser D. F. (1997). Proteome Research: New Frontiers in Functional Genomics. *Springer, Berlin*, *7*, 12-43.

Willam, J. Griffiths, Andreas, P. Jonsson, Suya Liu, Dilip, K. Rai, and Yuqin Wahg (2001). Electrospray and tandem mass spectrometry in biochemistry. *Biochem., 355*, 545-561.

Winston, R. L. and Fitzgerald, M. C. (1997). Mass spectrometry as readout of protein structure and function. *Mass Spectrom. Rev*., *16*, 165-179.

Xiaofeng Wang, Weimin Ni, Xiaochun Ge, Jiejin Zhang, Hong Ma, Kaiming Cao (2006). Proteomic identification of potential target proteins regulated by an ASK1-mediated proteolysis pathway. *Cell Research, 16*, 489-498.

Yan, S., Bhawal, R. and Yin, Z. (2022). Recent advances in proteomics and metabolomics in plants. *Mol Horticulture*, *2*, 17-22.

How to cite this article: Amanulla Khan, R.L. Pawar, M.H. Mali, Kamlakar More and Satish V. Deore (2023). Proteomics: An Emerging Tools, Database and Technique for Understanding Biological System. *Biological Forum – An International Journal, 15*(3): 511-523.