Available online at http://www.ifgdg.org

Int. J. Biol. Chem. Sci. 17(7): 2943-2957, December 2023

ISSN 1997-342X (Online), ISSN 1991-8631 (Print)

International Journal of Biological and Chemical Sciences

Review Paper http://ajol.info/index.php/ijbcs http://indexmedicus.afro.who.int

Proteomics analysis techniques and Bioinformatics approaches for biomarkers discovery

Ousseynou Ben DIOUF[1,2], Mamadou SOUMBOUNDOU[1,2*] and Cheikh SALL[1]

[1]*Mixed Research Exploration and Diagnosis (UMRED), UFR-Healthy, Iba Der THIAM University of, BP: A967 Thies, Senegal.*

[2]*Diamniadio Children's Hospital (HED), Senegal, BP: 204 Bargny, Senegal.*

[*]*Corresponding author; E-mail: mamadou.soumboundou@univ-thies.sn, Tel: +221 76 156 07 20 / 33 7 44 17 97 59.*

Received: 09-10-2023 Accepted: 11-12-2023 Published: 31-12-2023

ABSTRACT

Proteomics is the study of all the proteins expressed in the proteome and offer new avenues of research in the biomedical field. The recent development of protein analysis tools has continued to improve, offering a wide range of technologies categorised according to their ability to analyse and identify protein such as MS-based technology. These novels tools have been invaluable to the rise of proteomics and permitted significant progress in the identification of clinically applicable biomarkers and new therapeutic targets although detection of very low abundance proteins remains difficult. Due to the high throughput data generated by proteomic mass spectrometric analysis, automatic biomarker identification requires bioinformatics and database screening. Bioinformatics allows the development of new algorithms and software for the interpretation of mass spectrometry data in order to facilitate the exploitation of the data and the validation of potentially discovered biomarkers. The aim of this paper was to discuss the analytical techniques used in protein biomarker discovery and the usefulness of bioinformatics for the interpretation and validation of the data generated.

Keywords: Bioinformatics, Biomarkers, Mass spectrometric, Proteomics.

INTRODUCTION

First used by Marc Wilkins in 1996 to refer to the "PROTein complement of a genOME" (Wilkins et al., 2013), proteomics is by definition the study of the proteome, i.e. all the proteins expressed in the organism, a biological fluid (blood, urine, saliva etc.), a tissue or a cell. It offers new avenues of research in the biomedical field. The development of this field is propelled by a vast international proteomics project, called the Human Proteome Project (HPP), which was launched in 2011 by the world organisation HPO (Human Proteome Organisation) with the aim of establishing a database to describe the proteins corresponding to the 19,800 genes predicted to be coding in humans. Thus, out of 19,778 proteins predicted to be encoded by our genome, 19,357 expressed proteins were discovered, i.e. a rate of 92.8%, considerably reducing the number of missing proteins in the nextProt PE2, PE3 and PE4 databases (Omenn et al., 2020).

DOI: https://dx.doi.org/10.4314/ijbcs.v17i7.26

Although used before 2011, proteomics involves a wide range of processes such as protein expression profiling, post-translational modifications (PTMs), interactions, structures, quantification and functions (Kwon et al., 2021).

Over the past decade, the development of protein analysis tools has continued to improve, offering a wide range of technologies categorised according to their ability to analyse and identify protein. Conventional chromatography-based analytical techniques such as ion exchange chromatography (IEC), size exclusion chromatography (SEC) and affinity chromatography(Aslam et al., 2017; Al-Amrani et al., 2021) are used in individual protein analysis and are distinct from gel electrophoresis analytical techniques.

Gel electrophoresis is most commonly used for the separation of complex protein samples and is distinguished from sodium dodecyl sulphate-polyacrylamide (SDS-Page), two-dimensional gel electrophoresis (2-DE) and two-dimensional differential gel electrophoresis (2D-DIGE).

Alongside these are early analytical techniques such as mass spectrometry (MS) which has been invaluable to the rise of proteomics and has permitted significant progress in the identification of clinically applicable biomarkers and new therapeutic targets Lin et al., 2019) although detection of very low abundance proteins remains difficult (Crutchfield et al., 2016). To address these issues, stable isotope labelling approaches have been developed,

Stable isotope labelling-based approaches fall into three major categories depending on how heavy isotopes are introduced: (1) chemical isotope labelling (ICAT) (Colangelo and Williams, 2006), isobaric Tags for Relative and Absolute Quantitation (iTRAQ) (Ow et al., 2008; Ross et al., 2004a),tandem mass tagging (TMT) (Dayon et al. 2008), in which labels are attached to proteins/peptides through chemical derivatization, (2) enzymatic labelling (e.g. H2 ^{16}O/ H2 ^{18}O) (Tian et al., 2023; Sakai et al., 2005), in which labelling is introduced through enzymatic reaction and (3) metabolic labelling (e.g. SILAC, Stable Isotope Labelling by Amino Acids in Cell culture) (Jiang and English, 2002), in which labelling is incorporated into proteins during in vivo protein synthesis (Hoedt et al., 2019). These technics have significantly improved the sensitivity and detection of proteins by mass spectrometry (Ross et al., 2004; Ong et al., 2002). In addition, these approaches allow simultaneous analysis and quantification of peptides for multiple samples, as well as direct comparison between samples (Hristova and Chan, 2019). In contrast, Label-free quantitative proteomics is useful in searching for disease-associated factors and has been used to investigate the mechanism of TCM in recent years (Xie et al., 2020).

Subsequent analysis MS-based and its derivatives allows the characterization and measurement of the abundance of proteins in a sample and generates a very large amount of data. Bioinformatics allows the development of new algorithms and software for the interpretation of mass spectrometry data in order to facilitate the exploitation of the data and the validation of potentially discovered biomarkers.

In this review, we discuss proteomics analysis techniques and the contribution of bioinformatics to the biomarker discovery process.

PROTEOMICS ANALYSIS METHODS

The early identification of aggressive cancers requires improved sensitivity and the implementation of biomarkers representative of tumor heterogeneity. In order to facilitate the search for biomarkers, various proteomic analysis techniques were set up, in particular one-dimensional (1D) and two-dimensional (2D) gel electrophoresis (2-DE) (Al-Amrani et al., 2021b), thus enabling the first experiment in proteomics to be perfected in 1995. A recent study shows out of 308 biotic stress responsive proteins, 40 proteins were identified as biotic stress responses proteins directly coupled to disease and pathogen infection on wheat (Hena et al., 2010).

Other technologies based on high-throughput gel-free screening, such as

multidimensional protein identification technology, have also emerged and significant advances are being made in this area. The emergence of these has increased the potential for identifying protein biomarkers in cancer.

In mass spectrometry, technological advances have increased the capacity for rapid, accurate and efficient proteomic discovery. To date, several proteomic techniques have been applied to identify potential biomarkers of radiosensitivity in cancer (Luo and Ge 2022).

Matrix-Asisted Laser Desorption/ Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF-MS)

In proteomics, mass spectrometry (MS) remains the most widely used technology for protein ionisation. Its applications are very broad and mainly concern the identification of peptides or proteins, the analysis of their amino acid sequence or the detection of post-translational modifications (PTMs). Introduced in 1985 by Karas et al. (1985), MS has been revolutionized in part by matrix-assisted laser desorption mass spectrometry (MALDI-TOF-MS) introduced in 1988 by John Fenn and Koichi Tanaka (Tanaka et al., 1988; Greco et al., 2018). It is often combined with two-dimensional electrophoresis.

Recently, a study on ovarian cancer (OC) showed the performance of MALDI-TOF. Serum proteomic patterns in samples from OC patients were obtained using MALDI-TOF-MS. Eighty-nine serum samples (44 ovarian cancer and 45 healthy controls) were pretreated using solid-phase extraction method. Main outcome of this study was an identification of potential OC biomarkers (complement C3, kininogen-1, inter-alpha-trypsin inhibitor heavy chain H4, and transthyretin) by applying liquid chromatography coupled with tandem mass spectrometry (Swiatly et al., 2017).

MALDI-TOF-MS requires relatively less intense sample preparation, and peaks in a spectrum are generally used as indications of protein content. However, reproducibility of results can be an issue due to the sensitivity of this technique to contaminants from salts. To this end, the emergence of new and more advanced MS techniques could be an alternative to MALDI-TOF-MS.

Liquid chromatography mass spectrometry (LC-MS)

LC-MS is a hyphenated technique, combining the separation power of HPLC, with the detection power of mass spectrometry to accurately identify and/or quantify many ranges of organic compounds, from small-molecule drug metabolites to peptides and proteins, compared to conventional HPLC techniques that do not allow qualitative analysis. LC-MS requires a great deal of methodological development and data processing.

The LC systems up front of the MS can be used to introduce pure analytes, but often are necessary to separate a mixture of analytes prior to MS analysis. It is important to note that certain classes of compounds (e.g. volatile compounds or compounds that tend to hydrolyse under aqueous conditions) are not compatible with LC-MS. The most commonly used LC-MS interface and ionization technique is ESI (De Vijlder et al., 2018).

In LC-MS-based proteomics, protein complexes are first subjected to enzymatic cleavage prior to mass spectrometric analysis of the resulting peptide products, in contrast to "top-down" proteomics, which deals with intact proteins and is limited to mixtures of single proteins (Karpievitch et al., 2010).

Label-based proteomics

The technique of multiplexing in proteomics was first introduced in 1999 (Tian et al., 2023). Several methods based on chemical isotope labelling have been developed and adapted for protein quantification. ICAT was the first chemical isotope labelling introduced and from which other labels have emerged such as SILAC, dimethyl labeling,^{18}O labelling, neutron

encoding (NeuCode) (Rose et al., 2013; Sivanich et al., 2022). These methods introduce small mass differences via heavy isotopologues into proteins or peptides, which can be distinguished from each other in the MS1 precursor spectrum. Several samples can be analysed simultaneously because the respective labelled peptides can be resolved by MS. In addition, pooling of samples prior to LC-MS reduces sample variation in the workflow, signal variation and overall analysis time.

Alongside chemical isotope libelling, isobaric libelling has been developed and includes iTRAQ (Chen et al., 2021), tandem mass tagging (TMT) (Thompson et al., 2003), N,N-dimethyl leucine (DiLeu) (Frost et al., 2020), amine-reactive isobaric deuterium tags (DiART) (Zhang et al., 2010), 10-plex isobaric tags (IBT) (Ren et al., 2018), and a sulfoxide-based isobaric libelling reagent (SOT 2 reagent, SOT) (Stadlmeier et al., 2018).

ITRAQ technology utilizes isobaric reagents to label the primary amines of peptides and proteins (

Figure ***1***). The iTRAQ reagents usually consist of an N-methyl piperazine reporter group, a balance group, and an N-hydroxysuccinimide ester group that is reactive with the primary amines of peptides. The balance groups present in each of the iTRAQ reagents function to make the labelled peptides from each sample isobaric and the quantification is facilitated through analysis of reporter groups that are generated upon fragmentation in the mass spectrometer. There are currently two mainly used reagents: 4-plex and 8-plex, which can be used to label all peptides from different samples/treatments. These samples are then pooled and usually fractionated by nano-liquid-chromatography and analysed by tandem mass spectrometry (MS/MS) (Tian et al., 2019; Rauniyar and Yates, 2014). In 2020, Xia et al. (2020) were able to analyse the effect of metformin on the invasion and migration of CC HeLa and SiHa cell lines in cervical cancer and found that 53 proteins were differentially expressed, including 20 overexpressed and 33 under expressed proteins.

The increased multiplexing capability of TMT tags and demonstrated its application by using 6-plex TMT reagents in relative quantification of standard protein mixtures at various concentrations was shown by Dayon et al. (2008). In this study, TMT 6-plex was also used to assess the differential protein abundance in post-mortem cerebrospinal fluid samples after brain injury vs. antemortem samples.

N, N-Dimethyl leucine (DiLeu) and Deuterium isobaric Amine Reactive Tag (DiART) are seen as alternatives to iTRAQ and TMT. However, the structure of DiLeu reagents consists of a dimethyl leucine as a reporter group, a carbon balancing group and a triazine ester amine reactive group at the C-terminus with an m/z range of 114-119 (Xiang et al., 2010; Rauniyar and Yates, 2014; Sivanich et al., 2022) while DiART contains an NHS amine reactive group similar to TMT and iTRAQ, a β-alanine equilibrator and an N,N-dimethylleucine reporter group with an m/z ratio of 114-119, similar to DiLeu, with the capacity for a 6-complex. DiART has been compared to iTraQ previously and has been reported to have a longer ratio which improves the signal to tone to iso ratio (S/N) as well as a compression without ratio.

Multiple Reaction Monitoring (MRM)

MRM is a highly specific and sensitive mass spectrometry technique that allows the quantification of targeted proteins in complex mixtures. MRM requires that the MS be configured to monitor only specific values of the mass-to-charge ratio (m/z) of interest, therefore the probability of detecting even low levels of a peptide in the presence of a complex mixture of peptides is much higher. It operates in tandem MS (MS/MS) mode coupled with triple quadrupole instrumentation (Beretov et al., 2014; You et al., 2013) which first targets the ion corresponding to the compound of

interest and then fragments the target into several other ions. One or more of the ions resulting from the fragmentation and meeting the mass criterion of the molecule of interest are isolated in the mass spectrometer and subjected to quantification. Shi et al. (2014) and Wu et al. (2017) developed a highly sensitive MRS/MRM-based protocol for the quantification of anterior gradient 2 (AGR2) in urine and serum of prostate cancer (CaP) patients with good correlation of MRS/MRM and ELISA results and a significant difference ($p = 0.026$) was observed in the urine AGR2/PSA concentration ratios between non-cancer and cancer subjects.

BIOINFORMATICS FOR PROTEOMICS

Proteins are the workhorses in living cells and their abnormal abundance is often associated with diseases. For this reason, characterizing and comparing proteomes between patient and control is often more effective in identifying drug targets than genomic or transcriptomic data (Xia, 2017). Most proteomic data are used for comparisons between cohorts of patients and matched normal controls for protein identify by mass spectrometry. Proteomics MS-based generate large amounts of data and their manual integration is impossible to achieve without computer assistance.

In parallel with the rapid advancement of MS techniques and the generation of large amounts of data from them, various bioinformatics analysis methods and dedicated databases are being developed for protein quantification, identification and analysis. These methods can be identified in relation to the approach used in mass spectrometry. In MS1 (first step MS) applications such as MaxQuant (Merrill et al., 2014; Cao et al., 2012; Tyanova et al., 2016), PVIEW (Khan et al., 2009) and XPRESS (Han et al., 2001) are developed for chemical isotope tagging quantification techniques. Most software tools can handle samples from several labelling methods. For example, XPRESS is capable of analysing samples labelled by ICAT, SILAC and ICPL, and it can also calculate the relative abundance of proteins based on the elution profiles of labelled peptide pairs. PVIEW can process SILAC, ICPL and ICAT labelled samples, and it can even perform label-free quantification by non-linear alignment and XIC-based label-free quantification. In contrast, MaxQuant is designed specifically for high-resolution SILAC-labelled data from Thermo Orbitrap and FT mass spectrometers. Using the Perseus framework (Rudolph and Cox, 2019), it is easy to perform a statistical downstream analysis of the raw quantification results from MaxQuant (Chen et al., 2020).

In recent years, the advancement of bioinformatics tools for the effective analysis of the rapidly increasing proteomics data has been a key area of interest (Figure 2). As part of a large interconnected network, protein and peptide expressions are becoming highly useful for the fundamental understanding of disease (Paul et al., 2020). In a recent study, Van et al. (Van et al., 2017) investigated the biological implications of differentially excreted urinary proteins in patients with diabetic nephropathy (DN). Artificially constructed protein-protein interaction (PPI) arrays were used to identify common and stage-specific biological processes in diabetic kidney disease (Chen et al., 2020).

PROTEIN BIOMARKERS

Due to cancer importance, the most proteomics studies in the field of biomarker discovery concerned cancer diseases. Proteomics as one of the modern areas of biochemistry holds great promise in the cancer study. Since proteome represents actual state of the cell, tissue, or organism, there are suitable biomarkers related to the tumours which can be used for diagnostic proposes or follow up of patients (Ghafourian et al., 2013). The use of protein biomarkers provides a better prognosis and increases the sensitivity and specificity for predicting response to patient treatments. To this end, a number of scientific studies have

predicted potential biomarkers that can be used in various diseases due to the new drug discovery strategies MS-based. Biomarker research can be performed at different biological levels such as plasma, serum, tissue cells.

Biomarker discovery in serum or plasma blood

Blood plasma is one of the most useful biological fluids for protein biomarker discovery as well as clinical investigations for diagnostic and therapeutic purposes due to its abundant protein concentration (Mukherjee et al., 2022; Kumar et al., 2020). The major methods employed in plasma proteome-based workflow include separation by 2D-DIGE, MALDI-Q TOF MS/MS analysis (by label-free or label-based quantitation or targeted proteomics), biomarker validation by ELISA, and statistical analysis (Kumar et al., 2020; Martínez-Rodríguez et al., 2021). By comparing a group of invasive cervical cancer patients and a group of normal individuals using the multiplex proximity extension assay (PEA), Berggrund et al. (2019) identified 11 proteins (PTX3, ITGB1BP2, AXIN1, STAMPB, SRC, SIRT2, 4E-BP1, PAPPA, HB-EGF, NEMO and IL27) as signature proteins for cervical cancer. In breast cancer, Skiöld et al. (2015) had investigated the protein expression profiles using isotope-coded protein labelling method (ICPL)and identified 40 proteins differentially regulated by radiosensitivity, a significant fraction of which were proteins regulating responses to oxidative stress. This allowed them to show that the intrinsic response to oxidative stress, as well as additional stress induced by ionizing radiation (IR), could be an important general factor influencing the sensitivity of cancer and normal tissues to radiotherapy. In the same perspective, Drobin et al. (2020) obtained similar results in breast cancer patients highlighting CHIT 1, PDGFB, RP2, SERPINC1, SLC4A, STIM1 and THPO proteins as significant positive predictors of radiosensitivity with a predominance of THPO and STM1 proteins. Another study by Moreno-Acosta et al. (2017) showed that IGF-1R β overexpression and Hb level (≤11 g/dl) were associated with poor prognosis, and thus appear to be possible interesting biomarkers of radiation resistance. These results corroborate previous preclinical studies suggesting IGF-1R and hypoxia to be part of the biological pathways leading to radio-resistance. Likewise, the potential mechanism essential to the regulation and expression of B55α, in addition to the likely role of the B subunit as a tumor suppressor in AML were comprehensively studied. A strong correlation was observed between B55α and several proteins (including MYC, PKCα, and SRC) after reverse phase protein analysis (RPPA) of 230 proteins in 511 AML patients. B55α suppression in OCI-AML3 cells by shRNA indicated that the B subunit is a PKCα phosphatase (Almaiman et al., 2016).

Biomarker discovery in tissue cells

The application of proteomics at the tissue level provides the most accurate reflection of the physiological state of the cancer tumor. Recent advances in MS-based technology have allowed for a continuous increase in proteome coverage with reliable quantification. Tissue-based proteomics is widely applied in different cancer types such as breast (Yanovich et al., 2018), lung (Doll et al., 2018), prostate (Guo et al., 2018).

Using cervical cancer cell lines and iTRAQ-based quantitative proteomics analysis, Xia et al. (2019) and Martínez-Rodríguez et al. (2021) analyzed the effect of metformin on invasion and migration of the CC cell lines HeLa and SiHa. The mechanism by which metformin inhibits the proliferation and invasion of CC cells was analyzed and the authors found 53 differentially expressed proteins, 20 overexpressed proteins, and 33 under-expressed proteins. Proteomic analysis, complemented with tumor xenograft modelling, showed that the expression of nine

proteins was decreased in cells treated with metformin, namely TGFβ-1, CCPG1, LGMN, SLC38A2, TRIM26, MTR, ATP6AP1, CIRBP, and PTP4A1, while the expression of CYR61 and IGFBP7 was increased compared to control cells. The authors concluded that metformin was capable of inhibiting the proliferation and invasion of CC cells in this proteomic assay.

A 2022 study showed that Copa could be used as a prognostic biomarker for cervical cancer. In this study, the authors investigated the protein expression profiles of cervical cancer of 28 fresh frozen tissue samples (11 adenocarcinoma (AC), 12 squamous cell carcinoma (SCC) and 5 normal cervixes (HC)) were included in discover cohort; 45 fresh frozen tissue samples (19 AC, 18 SCC and 8 HC) were included in verification cohort; 140 paraffin-embedded tissues samples of cervical cancer (85 AC and 55 SCC) were used for immunohistochemical evaluation (IHC) of coatomer protein subunit alpha (COPA) as a prognostic biomarker for cervical cancer; how deficiency of COPA affects cell viability and tumorigenic ability of cervical cancer cells (SiHa cells and HeLa cells) were evaluated by cell counting kit-8 and clone formation in vitro. The author identified COPA as a potential prognostic biomarker for cervical cancer in quantitative proteomics analysis (Bao et al., 2022).

In South Africa, a recent study searching for protein markers that are related to tenderness in the retail sector showed that WBS decreases with ageing, with a value of 6.37 kg on day 3 and 4.67 kg on day 14 after slaughter. The same trend was observed for MFL, which was 35.37 µm on day 3 and 22.53 µm on day 14. The authors concluded that WBS and MFL can be used as predictors of meat quality (Moloto et al., 2015).

Biomarker discovery in urinary

MS-based urine proteomics has great potential for the development of diagnostic and prognostic tests. Urine is an extremely complex biological fluid containing a wide variety of proteins, and although its dynamic range is lower than that of plasma, it is still important for identifying biomarkers of urological malignancies as it can be easily and non-invasively obtained and contains cells and proteins that originate from urogenital system (Decramer et al., 2008). Furthermore, as a glomerular filtrate of plasma, the urine proteome can reflect physiological and pathological status of the human body (Wu and Gao, 2015). In this sense, Lin et al. (2018) suggested the feasibility of applying the high throughput workflow in extensive urinary proteome profiling and clinically relevant biomarker discovery.

In the case of prostate cancer, a study using DDA-MS to characterize cat urine showed that all three proteins (β2M, PGA3 and MUC3) were potentially capable of distinguishing prostate cancer from benign prostatic hyperplasia when used individually or in combination (Jedinak et al., 2015).

To identify biomarkers of immunosuppression in healthy men, Xu et al. (2020) explored the urine proteome using the iTRAQ proteomics techniques and four proteins in the urine (SEMG-1, PIP, PDGFRL, and NDPK) were found to increase incrementally with the increased exercise intensity, which have the potential to be used as non-invasive immune biomarkers of exercise-induced immunosuppression. For the "co-upregulation" and "co-downregulation" of differential proteins in incremental treadmill running exercises, comprehensive biological functional annotations were performed using the Uniprot database website and DAVID data analysis software.

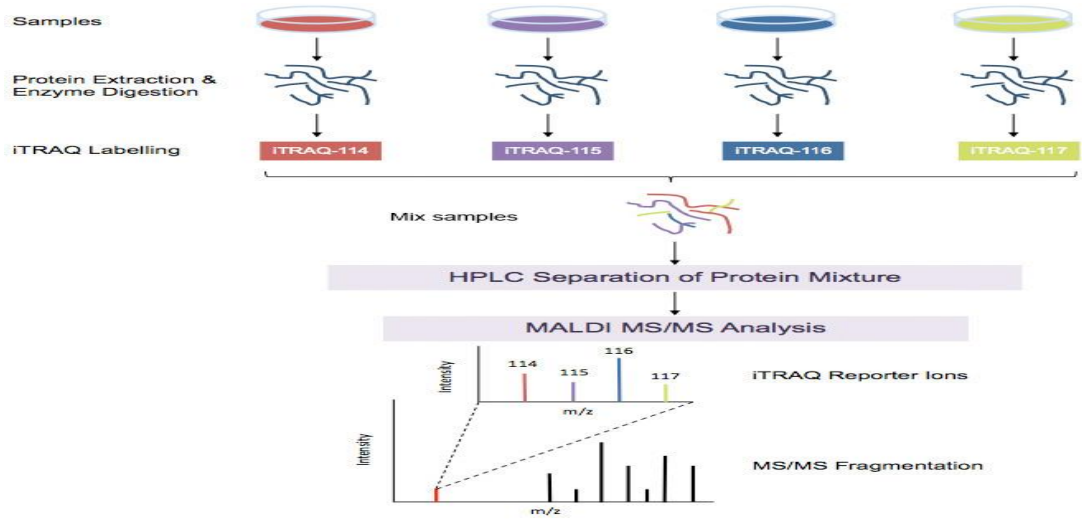

Figure 1 : iTRAQ workflow (4-plex) is shown above. Samples to be quantified are prepared under various treatment conditions followed by cell lysis to extract proteins. After using a standard protein assay to estimate the protein concentration of each sample, proteins are digested using an enzyme, such as trypsin, to generate proteolytic peptides. Each peptide digest is labelled with a different iTRAQ reagent and then the labelled digests are combined into one sample mixture. The combined peptide mixture is analysed by LC-MS/MS for both identification and quantification. A database search is then performed using the fragmentation data to identify the labelled peptides and hence the corresponding proteins. The fragmentation of the attached tag generates a low molecular mass reporter ion that can be used to relatively quantify the peptides and the proteins from which they originated.

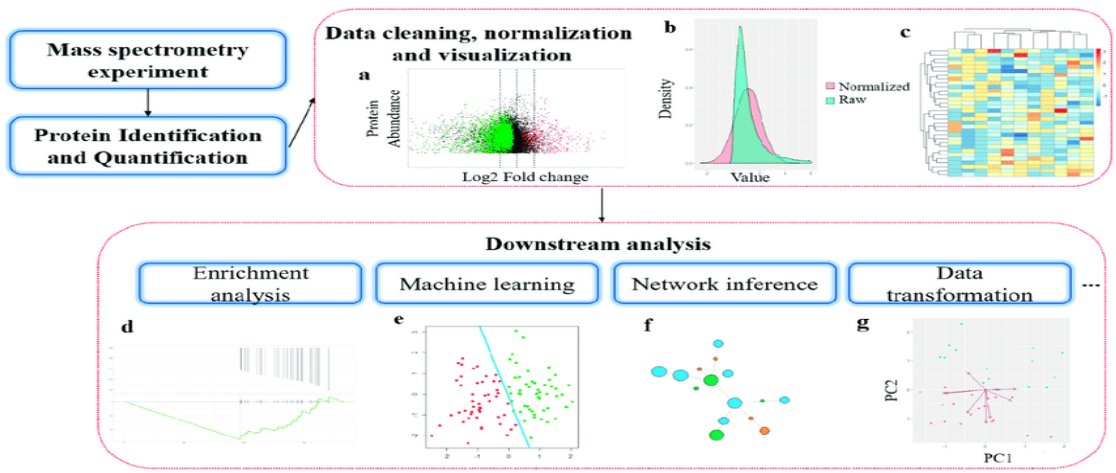

Figure 2: General workflow of bioinformatics analysis in mass spectrometry-based proteomics. (a) MA-plot from protein differential abundance analysis. X-axis is the log2 transformed fold change and Y-axis is the average protein abundance from replicates. (b) Distribution of protein abundance data before and after normalization. (c) Heatmap for protein abundance with clustering; (d) Protein set enrichment analysis, Y-axis in the above plot shows the ranked list metric, and in the bottom, plot shows the running enrichment score. X-axis is the ranked position in protein list. (e) Machine learning-based sample clustering. (f) Illustration of a network inferred from proteomics data. (g) Dimensionality reduction of proteomics expression profile.

CONCLUSION

The use of proteomics and bioinformatics for biomarker research has shown great promise. Its application in cervical cancer radiotherapy could remove some of the barriers to precision treatment of tumours in developing countries such as Senegal. The use of mass spectrometry and bioinformatics technology in our research laboratories would make it possible to both analyse the mechanisms underlying radiation resistance and to select patients suitable for radiotherapy treatment based on the biomarkers potentially discovered in these areas. These aspects could be used for the implementation of precision radiotherapy in cervical cancer, the prevalence of which is very high in Senegal.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

All authors contributed to the realization of this work and to the preparation of the manuscript.

REFERENCES

Al-Amrani, Safa, Zaaima Al-Jabri, Adhari Al-Zaabi, Jalila Alshekaili, Murtadha Al-Khabori. 2021a. Proteomics: Concepts and Applications in Human Medicine. *World Journal of Biological Chemistry*, **12**(5): 57. DOI : https://doi.org/10.4331/WJBC.V12.I5.57 .

Almaiman AA, Rasedee A, Ahmad BA, Zeenathul A, Eltayeb EE, Sultan AMS. 2016. Proteomic Profile of Acute Myeloid Leukaemia : A Review Update. *Tropical Journal of Pharmaceutical Research*. University of Benin. DOI : https://doi.org/10.4314/tjpr.v15i7.29.

Aslam B, Madiha B, Muhammad AN, Mohsin K, Muhammad HR. 2017. Proteomics : Technologies and Their Applications. *Journal of Chromatographic Science*, **55**(2): 182–96. DOI : https://doi.org/10.1093/CHROMSCI/BMW167.

Bao H, Xiaobin L, Zhixing C, Zhihong H, Li C, Mingbing W, Jiali H, Wenting L, Hongwei S, Xue J, Ping M, Huawen L. 2022. Identification of COPA as a Potential Prognostic Biomarker and Pharmacological Intervention Target of Cervical Cancer by Quantitative Proteomics and Experimental Verification. *Journal of Translational Medicine* **20**(1): 18. DOI : https://doi.org/10.1186/s12967-021-03218-1.

Beretov J, Valerie CW, Graham PH, Ewan KM, Kearsley JH, Yong L. 2014. Proteomics for Breast Cancer Urine Biomarkers. 123–67. DOI : https://doi.org/10.1016/B978-0-12-800094-6.00004-2.

Berggrund M, Stefan E, Lundberg M, Assarsson E, Stålberg K, Lindquist D, Göran H, Kjell G, Matts O, Gyllensten U. 2019. Identification of Candidate Plasma Protein Biomarkers for Cervical Cancer Using the Multiplex Proximity Extension Assay. *Molecular & Cellular Proteomics : MCP* **18** (4) : 735. DOI : https://doi.org/10.1074/MCP.RA118.001208.

Cao R, Ke C, Qin S, Yi Z, Jia L, Xianchun W, Ping C, Songping L. 2012. Quantitative Proteomic Analysis of Membrane Proteins Involved in Astroglial Differentiation of Neural Stem Cells by SILAC Labeling Coupled with LC–MS/MS. *Journal of Proteome Research*, **11**(2): 829–38. DOI : https://doi.org/10.1021/pr200677z.

Chen C, Jie H, John JT, Jianlin C. 2020. "Bioinformatics Methods for Mass Spectrometry-Based Proteomics Data Analysis." *International Journal of Molecular Sciences*, **21**(8). DOI : https://doi.org/10.3390/IJMS21082873.

Chen X, Yaping S, Tingting Z, Lian S, Roepstorff P, Fuquan Y. 2021. Quantitative Proteomics Using Isobaric Labeling: A Practical Guide. *Genomics, Proteomics & Bioinformatics*, **19** (5) : 689–706. DOI : https://doi.org/10.1016/j.gpb.2021.08.012.

Colangelo CM, Kenneth RW. 2006. Isotope-Coded Affinity Tags for Protein Quantification. *Methods in Molecular Biology (Clifton, N.J.)*, **328** : 151–58. DOI : https://doi.org/10.1385/1-59745-026-X:151.

Crutchfield CA, Stefani NT, Lori JS, Chan DW. 2016. Advances in Mass Spectrometry-Based Clinical Biomarker Discovery. *Clinical Proteomics*, **13** (1). DOI : https://doi.org/10.1186/S12014-015-9102-9.

Dayon L, Hainard A, Licker V, Turck N, Kuhn K, Hochstrasser DF, Burkhard PR, Jean-Charles Sanchez. 2008. "Relative Quantification of Proteins in Human Cerebrospinal Fluids by MS/MS Using 6-Plex Isobaric Tags." *Analytical Chemistry* **80**(8): 2921–31. DOI : https://doi.org/10.1021/ac702422x.

Decramer S, Gonzalez de Peredo A, Breuil B, Mischak H, Monsarrat B, Bascands J-L, Schanstra JP. 2008. Urine in Clinical Proteomics. *Molecular & Cellular Proteomics*, **7**(10): DOI : 1850–62. https://doi.org/10.1074/mcp.R800001-MCP200.

Doll S, Kriegmair MC, Santos A, Wierer M, Coscia F, Neil HM, Porubsky S, Geyer PE, Mund A, Nuhn P, Mann M. 2018. "Rapid Proteomic Analysis for Solid Tumors Reveals <scp>LSD</Scp> 1 as a Drug Target in an End-stage Cancer Patient." *Molecular Oncology*, **12**(8) : 1296–1307. DOI : https://doi.org/10.1002/1878-0261.12326.

Drobin K, Marczyk M, Halle M, Danielsson D, Papiez A, Sangsuwan T, Bendes A, Hong M-G, Qundos U, Harms-Ringdal M, Wersall P, Plonska J, Schwenk JM, Haghdoost S. 2020. Molecular Profiling for Predictors of Radiosensitivity in Patients with Breast or Head-and-Neck Cancer. *Cancers*, **12**(3): 753. DOI : https://doi.org/10.3390/cancers12030753

Frost DC, Feng Y, Li L. 2020. 21-Plex DiLeu Isobaric Tags for High-Throughput Quantitative Proteomics. *Analytical Chemistry*, **92**(12): 8228–8234. DOI : https://doi.org/10.1021/acs.analchem.0c00473.

Ghafourian S, Sekawi Z, Raftari M, Mohd Shukuri MA. 2013. Application of Proteomics in Lab Diagnosis. *Clinical Laboratory*, **59**(5–6): 465–74. DOI : https://doi.org/10.7754/Clin.Lab.2013.130124.

Greco, Viviana, Cristian Piras, Luisa Pieroni, Maurizio Ronci, Lorenza Putignani, Paola Roncada, Andrea Urbani. 2018. Applications of MALDI-TOF Mass Spectrometry in Clinical Proteomics. *Expert Review of Proteomics*, **15**(8) : 683–96. DOI : https://doi.org/10.1080/14789450.2018.1505510.

Guo T, Li L, Qing Z, Niels JR, Konstantina C, Wong CE, Wagner U, Rueschoff JH, Jochum W, Frankhauser CD, Saba K, Poyet C, Wild PJ, Aebersold R, Beyer A. 2018. Multi-Region Proteome Analysis Quantifies Spatial Heterogeneity of Prostate Tissue Biomarkers. *Life Science*

Alliance, **1**(2) : e201800042. DOI : https://doi.org/10.26508/lsa.201800042.

Han DK, Eng J, Zhou H, Aebersold R. 2001. Quantitative Profiling of Differentiation-Induced Microsomal Proteins Using Isotope-Coded Affinity Tags and Mass Spectrometry. *Nature Biotechnology*, **19**(10): 946–51. DOI : https://doi.org/10.1038/nbt1001-946.

Hena A, Kamal M, Ki-Hyun K, Kwang-Hyun S, Da-Eun K, Myeong-Won O, Jong-Soon C, Hisashi H, Hwa-Young H, Sun-Hee W. 2010. Proteomics-Based Dissection of Biotic Stress Responsive Proteins in Bread Wheat (Triticum Aestivum L.). *African Journal of Biotechnology*, **9** (43) : 7239–55. DOI : https://doi.org/10.5897/AJB10.257.

Hoedt E, Zhang G, Neubert TA. 2019. Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC) for Quantitative Proteomics. 531–39. DOI : https://doi.org/10.1007/978-3-030-15950-4_31.

Hristova VA, Chan DW. 2019. Cancer Biomarker Discovery and Translation : Proteomics and Beyond. *Expert Review of Proteomics*, **16**(2): 93–103. DOI : https://doi.org/10.1080/14789450.2019.1559062.

Jedinak A, Curatolo A, Zurakowski D, Dillon S, Bhasin MK, Libermann TA, Roy R, Sachdev M, Loughlin KR, Moses MA. 2015. Novel Non-Invasive Biomarkers That Distinguish between Benign Prostate Hyperplasia and Prostate Cancer. *BMC Cancer*, **15**(1): 259. DOI : https://doi.org/10.1186/s12885-015-1284-z.

Jiang H, English AM. 2002. Quantitative Analysis of the Yeast Proteome by Incorporation of Isotopically Labeled Leucine. *Journal of Proteome Research*, **1**(4): 345–50. DOI : https://doi.org/10.1021/pr025523f.

Karas M, Bachmann D, Hillenkamp F. 1985. Influence of the Wavelength in High-Irradiance Ultraviolet Laser Desorption Mass Spectrometry of Organic Molecules. *Analytical Chemistry*, **57**(14): 2935–39. DOI : https://doi.org/10.1021/ac00291a042.

Karpievitch YV, Polpitiya AD, Anderson GA, Smith RD, Dabney AR. 2010. Liquid Chromatography Mass Spectrometry-Based Proteomics: Biological and Technological Aspects. *The Annals of Applied Statistics*, **4**(4). DOI : https://doi.org/10.1214/10-AOAS341.

Khan Z, Bloom JS, Garcia BA, Singh M, Kruglyak L. 2009. Protein Quantification across Hundreds of Experimental Conditions. *Proceedings of the National Academy of Sciences*, **106** (37) : 15544–15548. DOI : https://doi.org/10.1073/pnas.0904100106

Kumar V, Ray S, Ghantasala S. 2020. An Integrated Quantitative Proteomics Workflow for Cancer Biomarker Discovery and Validation in Plasma. *Frontiers in Oncology*, **10** (September) : 1840. DOI : https://doi.org/10.3389/FONC.2020.543997/BIBTEX.

Kwon YW, Han Seul J, Sungwon B, Youngsuk S, Parkyong S, Minseok S, Jong Hyuk Y. 2021. Application of Proteomics in Cancer : Recent Trends and Approaches for Biomarkers Discovery. *Frontiers in Medicine*, **8** (September). DOI : https://doi.org/10.3389/FMED.2021.747333.

Lin L, Quan Y, Jiaxin Z, Zonglong C, Ruijun T. 2018. Fast Quantitative Urinary Proteomic Profiling Workflow for Biomarker Discovery in Kidney Cancer. *Clinical Proteomics*, **15** (1) : 42. DOI :

https://doi.org/10.1186/s12014-018-9220-2.

Lin Y-H, Eguez RV, Torralba MG, Singh H, Golusinski P, Golusinski W, Masternak M, Nelson KE, Freire M, Yu Y. 2019. Self-Assembled STrap for Global Proteomics and Salivary Biomarker Discovery. *Journal of Proteome Research*, **18**(4): 1907–1915. DOI : https://doi.org/10.1021/acs.jproteome.9b00037.

Luo H, Ge H. 2022. Application of Proteomics in the Discovery of Radiosensitive Cancer Biomarkers. *Frontiers in Oncology*, **12**(February). DOI : https://doi.org/10.3389/fonc.2022.852791.

Martínez-Rodríguez F, Limones-González JE, Mendoza-Almanza B, Esparza-Ibarra EL, Gallegos-Flores PI, Ayala-Luján JL, Godina-González S, Salinas E, Mendoza-Almanza G, Kalyuzhny E. 2021. Cells Understanding Cervical Cancer through Proteomics. DOI : https://doi.org/10.3390/cells10081854.

Merrill AE, Hebert AS, MacGilvray ME, Rose CM, Bailey DJ, Bradley JC, Wood WW, El Masri M, Westphall MS, Gasch AP, Coon JJ. 2014. NeuCode Labels for Relative Protein Quantification. *Molecular & Cellular Proteomics*, **13** (9): 2503–12. DOI : https://doi.org/10.1074/mcp.M114.040287.

Moloto KW, Frylinck L, Strydom PE, Koorsen G. 2015. Proteomics Approach as a New Way to Predict Tenderness as Compared to the Classical South African Beef Carcass Classification System. *South African Journal of Animal Science*, **45** (3): 249–54. DOI : https://doi.org/10.4314/sajas.v45i3.3.

Moreno-Acosta P, Vallard A, Carrillo S, Gamboa O, Romero-Rojas A, Molano M, Acosta J, Mayorga D, Rancoule C, Garcia MA, Cotes Mestre M, Magné N. 2017. Biomarkers of Resistance to Radiation Therapy: A Prospective Study in Cervical Carcinoma. *Radiation Oncology*, **12** (1) : 120. DOI : https://doi.org/10.1186/s13014-017-0856-2.

Mukherjee A, Pednekar CB, Kolke SS, Kattimani M, Duraisamy S, Burli AR, Gupta S, Srivastava S. 2022. Insights on Proteomics-Driven Body Fluid-Based Biomarkers of Cervical Cancer. *Proteomes*, **10**(2). DOI : https://doi.org/10.3390/PROTEOMES10020013.

Omenn GS, Lane L, Overall CM, Cristea IM, Corrales FJ, Lindskog C, Paik YK, Van Eyk JE, Liu S, Pennington SR, Snyder MP, Baker M, Bandeira N, Aebersol R, Moritz RL, Deutsch EW. 2020. Research on the Human Proteome Reaches a Major Milestone : >90% of Predicted Human Proteins Now Credibly Detected, According to the HUPO Human Proteome Project. *Journal of Proteome Research*, **19**(12) : 4735–4746. DOI : https://doi.org/10.1021/ACS.JPROTEOME.0C00485/SUPPL_FILE/PR0C00485_SI_004.XLS.

Ong SE, Blagoev B, Kratchmarova I. Kristensen DB, Steen H, Pandey A, Mann M. 2002. Stable Isotope labelling by Amino Acids in Cell Culture, SILAC, as a simple and Accurate Approach to Expression Proteomics. *Molecular & cellular proteomics*, **1** (5): 376-386. DOI : https://doi.org/10.1074/mcp.M200025-MCP200.

Ow SY, Cardona T, Taton A, Magnuson A, Lindblad P, Stensjö K, Wright PC. 2008. Quantitative Shotgun Proteomics of Enriched Heterocysts from Nostoc sp. PCC 7120 using 8-plex Isobaric Peptide

tags. *Journal of Proteome Research*, **7**(4): 1615–1628. DOI : https://doi.org/10.1021/PR700604V.

Paul P, Antonydhason V, Gopal J, Haga SW, Hasan N, Oh JW. 2020. Bioinformatics for Renal and Urinary Proteomics: Call for Aggrandization. *International journal of molecular sciences*, **21**(3): 961. DOI : https://doi.org/10.3390/ijms21030961.

Navin R. 2014. Isobaric Labeling-Based Relative Quantification in Shotgun Proteomics. *Journal of Proteome Research*, **13**(12): 5293–5309. DOI : https://doi.org/10.1021/pr500880b.

Ren Y, He Y, Lin Z, Zi J, Yang H, Zhang S, Lou X, Wang Q, Li S, Liu S, Liu S. 2018. Reagents for Isobaric Labeling Peptides in Quantitative Proteomics. *Analytical Chemistry*, **90** (21): 12366–12371. DOI : https://doi.org/10.1021/acs.analchem.8b00321.

Rose CM, Merrill AE, Bailey DJ, Hebert AS, Westphall MS, Coon JJ. 2013. Neutron Encoded Labeling for Peptide Identification. *Analytical Chemistry*, **85**(10): 5129–5137. DOI : https://doi.org/10.1021/ac400476.

Ross PL, Yulin NH, Jason NM, Brian W, Kenneth P, Stephen H, Nikita K, Sasi P, Subhakar D, Scott D, Subhasish P, Peter J, Stephan M, Michael B-J, Allan J, Darryl JP. 2004. Multiplexed Protein Quantitation in Saccharomyces Cerevisiae Using Amine-Reactive Isobaric Tagging Reagents. *Molecular & Cellular Proteomics : MCP*, **3**(12): 1154–1169. DOI: https://doi.org/10.1074/MCP.M400129-MCP200.

Rudolph JD, Cox J. 2019. A Network Module for the Perseus Software for Computational Proteomics Facilitates Proteome Interaction Graph Analysis. *Journal of Proteome Research*, **18** (5): 2052–2064. DOI : https://doi.org/10.1021/acs.jproteome.8b00927.

Sakai J, Kojima S, Yanagi K, Kanaoka M. 2005. 18O-Labeling Quantitative Proteomics Using an Ion Trap Mass Spectrometer. *Proteomics*, **5** (1): 16–23. DOI : https://doi.org/10.1002/pmic.200300885.

Sivanich MK, Gu T-J, Tabang DN, Li L. 2022. "Recent Advances in Isobaric Labeling and Applications in Quantitative Proteomics." *Proteomics*, **22** (19–20): 2100256. DOI : https://doi.org/10.1002/pmic.202100256.

Skiöld S, Azimzadeh O, Merl-PhamJ, Naslund I, Wersall P, Lidbrink E, Tapio S, Harms-Ringdahl M, Haghdoost S. 2015. Unique Proteomic Signature for Radiation Sensitive Patients; a Comparative Study between Normo-Sensitive and Radiation Sensitive Breast Cancer Patients. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **776**: 128–35. DOI : https://doi.org/10.1016/j.mrfmmm.2014.12.002.

Stadlmeier M, Bogena J, Wallner M, Wühr M, Carell T. 2018. A Sulfoxide-Based Isobaric Labelling Reagent for Accurate Quantitative Mass Spectrometry. *Angewandte Chemie International Edition*, **57** (11): 2958–2962. DOI : https://doi.org/10.1002/anie.201708867.

Swiatly A, Horala A, Hajduk J, Matysiak J, Nowak-Markwitz E, Kokot ZJ. 2017. MALDI-TOF-MS Analysis in Discovery and Identification of Serum Proteomic Patterns of Ovarian Cancer. *BMC Cancer*, **17**(1): 472. DOI : https://doi.org/10.1186/s12885-017-3467-2.

Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, Yoshida T, Matsuo T. 1988. Protein and Polymer Analyses up Tom/z 100 000 by Laser Ionization Time-of-Flight Mass Spectrometry. *Rapid Communications in Mass Spectrometry*, **2**(8): 151–153. DOI : https://doi.org/10.1002/rcm.1290020802.

Thompson A, Schäfer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, Hamon C. 2003. Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. *Analytical Chemistry*, **75**(8): 1895–1904. DOI : https://doi.org/10.1021/ac0262560.

Tian L, You H-Z, Wu H, Wei Y, Zheng M, He L, Liu J-Y, Guo S-Z, Zhao Y, Zhou R-L, Hu X. 2019. ITRAQ-Based Quantitative Proteomic Analysis Provides Insight for Molecular Mechanism of Neuroticism. *Clinical Proteomics*, **16**(1): 38. DOI : https://doi.org/10.1186/s12014-019-9259-8.

Tian X, Permentier HP, Bischoff R. 2023. Chemical Isotope Labeling for Quantitative Proteomics. *Mass Spectrometry Reviews*, **42** (2): 546–576. DOI : https://doi.org/10.1002/mas.21709.

Tyanova S, Temu T, Cox J. 2016. The MaxQuant Computational Platform for Mass Spectrometry-Based Shotgun Proteomics. *Nature Protocols*, **11** (12): 2301–19. DOI : https://doi.org/10.1038/nprot.2016.136.

Van JAD, Scholey JW, Konvalinka A. 2017. Insights into Diabetic Kidney Disease Using Urinary Proteomics and Bioinformatics. *Journal of the American Society of Nephrology*, **28** (4): 1050–61. DOI : https://doi.org/10.1681/ASN.2016091018.

Vijlder TD, Valkenborg D, Lemière F, Romijn EP, Laukens K, Cuyckens F. 2018. A Tutorial in Small Molecule Identification via Electrospray Ionization-Mass Spectrometry: The Practical Art of Structural Elucidation. *Mass Spectrometry Reviews*, **37** (5): 607–29. DOI : https://doi.org/10.1002/mas.21551.

Wilkins MR, Sanchez JC, Gooley AA, Appel RD, Humphery-Smith I, Hochstrasser DF, Williams KL. 1996. Progress with Proteome Projects: Why All Proteins Expressed by a Genome Should Be Identified and How To Do It. *Biotechnology and Genetic Engineering Reviews*, **13**(1): 19–50. DOI: https://doi.org/10.1080/02648725.1996.10647923.

Wu D, Ni J, Beretov J, Cozzi P, Willcox M, Wasinger V, Walsh B, Graham P, Li Y. 2017. Urinary Biomarkers in Prostate Cancer Detection and Monitoring Progression. *Critical Reviews in Oncology/Hematology*, **118**: 15–26. DOI : https://doi.org/10.1016/j.critrevonc.2017.08.002.

Wu J, Gao Y. 2015. Physiological Conditions Can Be Reflected in Human Urine Proteome and Metabolome. *Expert Review of Proteomics*, **12** (6): 623–36. DOI : https://doi.org/10.1586/14789450.2015.1094380.

Xia Ch, Yang F, He Z, Cai Y. 2019. ITRAQ-Based Quantitative Proteomic Analysis of the Inhibition of Cervical Cancer Cell Invasion and Migration by Metformin. *Biomedicine & Pharmacotherapy*, **123**: 109762–109762. DOI : https://doi.org/10.1016/J.BIOPHA.2019.109762.

Xiang F, Ye H, Chen R, Fu Q, Li N. 2010. N, N-Dimethyl Leucines as Novel Isobaric Tandem Mass Tags for Quantitative Proteomics and Peptidomics. *Analytical*

Chemistry, **82**(7): 2817–2825. DOI : https://doi.org/10.1021/ac902778d.

Xia X. 2017. Bioinformatics and Drug Discovery. *Current Topics in Medicinal Chemistry*, **17** (15): 1709–1726. DOI : https://doi.org/10.2174/1568026617666161116143440.

Xie Y, Qian J, Lu Q. 2020. Proteomics Analysis of Differentially-Expressed Proteins in Uterus of Primary Dysmenorrhea Mice Following Administration of Nuangong Zhitong. *Tropical Journal of Pharmaceutical Research*, **19**(2): 265–76. DOI : https://doi.org/10.4314/tjpr.v19i2.8.

Xu G, Wentao W, McAinch AJ, Yan X, Weng X. 2020. Identification of Urinary Biomarkers for Exercise-Induced Immunosuppression by ITRAQ Proteomics. *BioMed Research International*, **2020**: 1–13. DOI : https://doi.org/10.1155/2020/3030793.

Yanovich G, Agmon H, Harel M, Sonnenblick A, Peretz T, Geiger T. 2018. Clinical Proteomics of Breast Cancer Reveals a Novel Layer of Breast Cancer Classification. *Cancer Research*, **78** (20): 6001–6010. DOI : https://doi.org/10.1158/0008-5472.CAN-18-1079.

You J, Willcox MD, Madigan MC, Wasinger V, Schiller B, Walsh BJ, Graham PH, Kearsley JH, Li Y. 2013. Tear Fluid Protein Biomarkers. *Advances in Clinical Chemistry*, **62**: 151–196. https://doi.org/10.1016/B978-0-12-800096-0.00004-4.

Zhang J, Wang Y, Li S. 2010. Deuterium Isobaric Amine-Reactive Tags for Quantitative Proteomics. *Analytical Chemistry*, **82**(18): 7588–95. DOI : https://doi.org/10.1021/ac101306x.