Journal of Pharmacognosy and Phytochemistry
Available online at www.phytojournal.com

E-ISSN: 2278-4136
P-ISSN: 2349-8234
JPP 2019; 8(3): 548-552
Received: 03-03-2019
Accepted: 06-04-2019

Abiraami TV
Division of Microbiology, ICAR-IARI, New Delhi, India

Mayur Naitam
Division of Microbiology, ICAR-IARI, New Delhi, India

Correspondence
Abiraami TV
Division of Microbiology, ICAR-IARI, New Delhi, India

Application of quantitative proteomics in understanding cellular functions in prokaryotes

Abiraami TV and Mayur Naitam

Abstract
The vision to understand the cellular functionalities by using several omics technologies is gaining momentum in the present decades. The proteins are the real key players that perform the cellular machinery, enzyme activities etc. The proteomics, the study of proteins far lagged behind the genomics till 1980 but with the advancement of Mass spectrometry inventories, the genomics era is shifting more towards the proteomics. Within the proteomics also the shift towards quantifying the protein expression becomes the major trend as it has a significant impact on comparative studies. The quantitative proteomic techniques range from classical gel-based to modern MS-based (Mass spectrometry is not inherently quantitative) that includes the introduction of stable isotope commonly by chemical modification or by metabolic labelling and also a label-free approach. Among the many formats for quantitative proteomics, stable-isotope labelling by amino acids in cell culture (Vasileva *et al.* 2018) and isobaric tag for relative and absolute quantification (Kumar *et al.* 2018) has emerged as a simple and most widely used technique. The approaches used in the proteomics with the techniques for quantitative proteomics with the concise applications are mentioned in the review.

Keywords: Quantitative proteomics, cellular functions, prokaryotes

1. Introduction
Proteins are the real key players controlling almost all cellular processes and post-genomic biology will not reach its potential until we have tools to study proteins on a large scale. Sadly, proteomics has lagged far behind DNA-based technologies, mainly because there are no protein-analysis methods similar to oligonucleotide hybridization, amplification and sequencing. This lag is diminishing rapidly, however, due to increasingly powerful mass spectrometry (MS)-based technologies (Gillet *et al.* 2016) [9]. Over the past decade, the application of quantitative proteomics in measuring alterations in varying biological conditions has gained momentum. The quantitative proteomic techniques range from classical gel-based to modern MS-based techniques. Application is highly successful in sophisticated settings and can be extended to understand high complexities such as host-pathogen interactions (Schubert *et al.* 2017) [19], mixed microbial communities. Visual proteomics, a technique using electron cryotomography (Oikonomou and Jensen 2017) [16] extend mass-spectrometry-based inventories and provide a quantitative description of the macromolecular interactions that underlie cellular functions.
Quantitative proteomics is an analytical chemistry technique for determining the number of proteins in a sample. The methods for protein identification are identical to those used in general (i.e. qualitative) proteomics but include quantification as an additional dimension. Rather than just providing lists of proteins identified in a certain sample, quantitative proteomics yields information about the physiological differences between two biological samples, involving the comparative study involving the control and treatment, also within the treatments.

2. Approaches to Metaproteomics
There are two strategies for separating and identifying proteins from complex mixtures (Wilmes, Heintz-Buschart *et al.* 2015) [23]. The proteome can be investigated either by 'top-down' analysis of the intact proteins or by 'bottom-up' analysis of the peptides generated by the trypsin digestion of proteins.

A. Bottom-up approach
The 'bottom-up' approach is suited particularly to identify protein, which is achieved either from mass fingerprinting profiles of the peptides, or from sequences of these peptides obtained

by MS/MS and a subsequent database search (Gillet *et al.* 2016) [9]. In bottom-up' proteomics, the proteins of interest are digested using trypsin and the resulting peptides are analysed by MS. Tandem MS (MS/MS) spectra of the peptides provide information on their amino acid sequences and possible post-translational modifications (PTM). The bottom-up approach begins with the prefractionation of a complex protein mixture by SDS-PAGE or 2-DE, protein bands or spots are excised from the gel, subjected to proteolytic digestion preferably trypsin and identified by MS. Generally, the peptide mixture obtained from the trypsin digestion undergoes an extraction/separation step using reverse-phase liquid chromatography (RP-LC). In the 'bottom-up' approach, the alternative to gel electrophoresis separation is multi-dimensional separation by LC which is called Gelfree approach/Multidimensional Protein Identification (MudPIT). After the protein digestion to peptides, the resulting mixture of peptides is separated by strong cation exchange (SCX) chromatography as the first dimension, followed by reversed-phase chromatography as the second. However, protein identification using 'bottom-up' analysis generally suffers from incomplete protein sequence coverage and the loss of information on post translational modification (PTM) or degradation, as a result of proteolytic digestion, protein interference problem, lack of peptide centric identification (Armengaud *et al.* 2016) [4].

B. Top-down approach

In the 'top-down' method, intact protein ions are detected, fragmented and analysed in the mass spectrometer, yielding both the molecular weight of the intact protein and protein fragmentation spectra. This allows the complete primary structure of the protein and all of its PTMs to be deduced. The top-down approach is not widely appreciated like bottom up approach as there is much development needs to be flourished in terms of MS sensitivity, protein extraction protocols, compatible gel electrophoresis method that can be variant of 1-D or 2-D.

3. Quantitative proteomics

Quantitative proteomics can be gel-based or gel-free MS proteomics. The gel-based proteomics includes DIGE, while gel free approach involves the use of labelling or spectral counting. The labelling can be also *in-vitro* or *in-vivo* labelling. The *in-vitro* labelling involves the labelling the proteins after protein digestion that can N-terminal, C-terminal or specific aminoacid labelling. iTRAQ known as isobaric tagging for relative and absolute quantization, TMT's are tandem mass tags and gist - global internal standard technology are N-terminal aminoacid labelling. C-terminal based peptide labelling methods involves esterification and proteolysis using different forms of oxygen. Amino acid labelling approach can include ICAT isotope-coded affinity tag that labels cysteine, QUEST or quantization using enhanced signal tags, VICAT or visible isotope-coded affinity tag, MCAT or mass coded abundance tagging. The *in-vivo* labelling involves the labelling of amino acids and detection in its incorporation to the respective proteins.

A. Gel based proteomics

Differential Gel electrophoresis (DIGE) involves the combination of cyanine dye fluors that are resolvable spectrally and the principle is the same as 2D. The covalent derivatisation of proteins with a fluorophore in complex protein mixtures prior to IEF and SDS-PAGE allows detection and quantification of differences in protein abundance within one single gel. The proteins are separated based on the molecular weight and also the isoelectric point. They modify the amino group of lysine *via* an amide linkage. Each dye adds 450 Da to the mass of protein. This mass shift does not affect the pattern visible on dye as they are matched for mass and charge so that separation will be possible on the single 2D gel. This minimises gel to gel variation and also allows for better spot matching. The unique properties involving in the dye are spectrally distinct, produce discrete signals, size and charge matched, possess multiplexing, photostable, pH insensitive, greater sensitivity: down to 25pg of a single protein. In a protocol followed, the control and treated samples are separately labelled using different dyes (take Cy3 and Cy5, respectively), while a mixture consisting of an equal amount of the control and treated samples are labelled with Cy2. The labelled samples were combined and made to run in a single 2D gel. After the gel loading, for visualisation use different light wavelength to see the samples and control. The images were processed and the quantitative comparison is made to provide a reliable result. DeCyder Software is used for Spot identification, Codetection of spots, spot volume ratio and Quantification calculation.

B. MS-based quantification:

Mass spectrometry plays an important role in structure determination with the prediction of fragmentation spectra for organic compounds. However, it was unable to biomolecule as the hard ionisation methods followed involving chemical ionisation, electron ionisation will cause them to completely decompose. With the emergence of Electrospray ionisation, Matrix-assisted laser desorption technique, fast atom bombardment the soft ionisation was made possible and the identification of protein and metabolites molecular weight determined. MS provides the semiquantitative protein abundance by counting the spectral counts present. However, the shift to relative and absolute abundance can be accompanied by using labelling technique. The representative emerging technique of *in-vitro* labelling iTRAQ, *in-vivo* labelling SILAC have been discussed below.

4. iTRAQ

iTRAQ labels from the applied based system and tandem mass tags labels are available from the thermo fisher are the currently available tagging technologies, where quantitation can be carried out in the MS/MS mode. Both have the same principle differing only in plexing capacity. iTRAQ stands for isobaric tagging for relative and absolute quantization. iTRAQ is N-terminal amino acid labelling mass spectrometry-based technique for relative and absolute quantitation of proteins present up to 4 samples to 8 samples depending upon the type of iTRAQ tags, (Wiese *et al.* 2007) [22]. iTRAQ reagents are set of multiplexed amine specific stable isotope reagents, it enables simultaneous identification and quantization both relative and absolute, there are two different types of iTRAQ reagents currently available,4-plex for sampling up to 4 samples and 8-plex for analysis of up to 8 samples (Aggarwal *et al.* 2006) [2]. The tags are isobaric having the same molecular weight making them coelute in Liquid chromatography and appear as single MS peak in first MS. In the second tandem MS having collision-induced dissociation, the tags component reporter and balancer group separates. The balancer group has no charge and will not be detected in MS. The reporter group has charge and will be detected in MS/MS as signature ions providing quantification.

So, the reporter ions are used to track the quantization and can be used to monitor the relative quantitation for proteins. The advantage includes multiplexing, identification of low-abundance proteins and transcription factors in *E. coli* using iTRAQ labels, expanded coverage of proteome by tagging tryptic peptides, application across diverse MW and pI ranges with increased analytical precision and accuracy. This method has disadvantages in the possibility of making errors in the quantification in the efficiency of enzymatic digestion. The peptide prefractionation step could be another possible way of introducing some variation. High cost per sample also restricts its use. Though multiplexing is the greatest advantage, the high cost restricts to use with more replicate per sample. Further developments in search algorithms and databases are needed for enhanced MS/MS spectra assignment.

5. SILAC

Stable isotope labeling of amino acids in cell culture (SILAC) is a metabolic method of incorporating isotopically labeled amino acids in proteins (Ong *et al.* 2006) [17]. Labeled amino acids are provided in the cell culture media in appropriate concentration. The cells take up the labeled amino acid during log phase and hence proteins will be labelled. The minimum five generations are must consider the protein turn over number. Usually, lysine or arginine amino acids is isotopically labeled with 13C, although other labelling options with 15N in the amino acid with are also available. Labeling lysine or arginine with 13C increases the molecular weight by 6 Da per lysine or arginine (Mann *et al.* 2006) [17]. Hence the shift in the MS peak by m/z of 6 Da denotes the peak for the labeled peptide. The ratio in the abundance of the light peak vs. the heavy peak is a measure of differential expression of that protein, subject to particular external stimuli. It is the powerful method to study post-translation modifications such as phosphorylation, protein–protein interaction removing false positives in protein-interaction studies, cell signalling and also the important method in secretomics, that involves the global study of secreted proteins and secretory pathways. However, the biggest disadvantage of SILAC is its restriction to culturable cells only. SILAC finds its best application in the field of clinical proteomics, where the dynamic change in the concentration of proteins can be studied at each level, for example, at various stages of glioma tissue samples. Many other studies involving post-translational modification, membrane protein dynamics, protein-protein interactions, can also be studied in great details using SILAC, which cannot be done using other quantitative proteomic approaches.

C. Protein identification: mass spectrometry

Once proteins have been extracted, another challenge is their identification and characterization. Mass spectrometry combined with database searches has become the preferred method for identifying the proteins present in cells or tissue and has been used also to identify the soil metaproteome (Schulze *et al.*, 2005) [20]. This technique makes it possible to execute large-scale proteome analyses of species whose genomes have been sequenced. The problem resides in the fact that many sequences of soil microorganisms have not been included in databases and no significant match is found when they are studied. In this sense, de novo sequencing could be used to assign protein functions, as has proved possible for complex matrices. MS provides information on the protein molecule, such as the mass of the peptide obtained from the protein molecule and its amino acid sequence. With this information, the original protein can be identified by a database search. The database search engine most commonly used are Sequest and Mascot. Many new search engines are emerging in the market purely academic also commercial providing advantageous in time, accuracy and sensitivity. The database mostly will be a public database such as Uni Port, RefSeq, SwissProt etc. The extracted protein samples are trypsin-digested to the peptides and fractioned by Reverse phase LC, then they are passed through Mass spectrometry. To analyse peptides by MS, it is necessary for the molecules to be dry and charged and preferable in the desolvated ions. The most common methods for ionization are ESI and MALDI (Aebersold *et al.* 2003) [1]. In both methods, the peptides are converted into ions by the addition or elimination of one or more protons. ESI employs the principle of nebulisation of electrospraying of peptides along with the solvent under high potential difference causing the solvent to evaporate, allowing to form stable droplets which are detected by the mass analyser. Normally, the ESI apparatus is connected in-line with a chromatograph and peptides are separated automatically and purified before being injected in the form of ions into the mass spectrometer. In the MALDI system, the sample is incorporated in a matrix and then exposed to laser radiation, which leads to the formation of molecular ions (Ong *et al.* 2005) [18]. The MALDI ionization system can be automated and ionization can be carried out directly, without purification, which represents an advantage over ESI. After ionization of the peptide molecules, their mass is analysed by a mass analyser, which separates the molecular ions in a vacuum chamber according to their charge and mass. Two types of mass analysers are common, the quadrupole and the time of flight (TOF). In quadrupole, the ions are conducted through an electric field created by a system of four parallel metal rods. The quadrupole can act as a mass filter that permits ions with a given m/z ratio to be transmitted (Fenn, Mann *et al.* 1990, [17]; Banerjee and Mazumdar 2012) [5]. In Time of flight: m/z ratio of an ion and the necessary time to pass through a flight tube are determined. Mostly specific ionisation and the mass analyser will be present in the mass spectrometer such as ESI-quadrupole, MALDI-TOF.

6. Application of quantitative proteomics

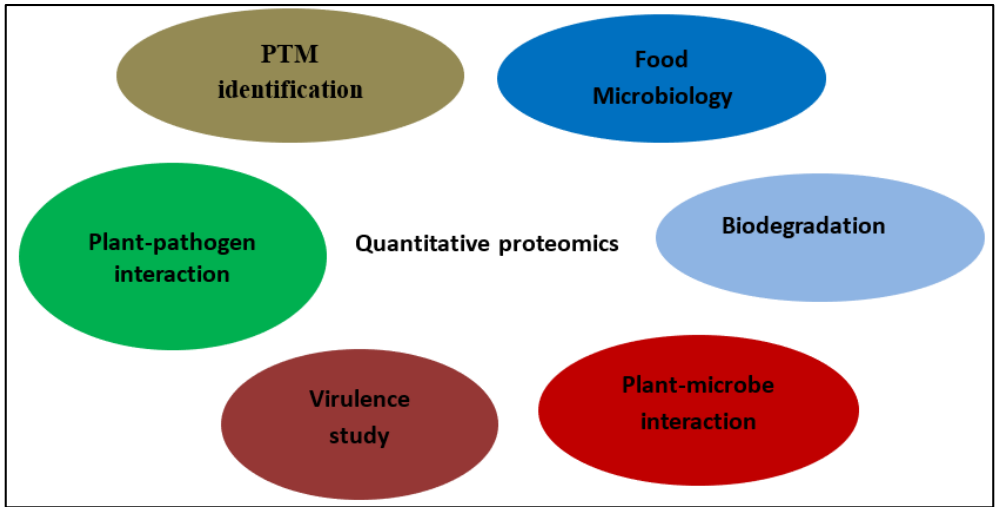

Fig 1: Overall application of quantitative proteomics

Xanthomonas campestris Pathovar vesicatoria (*Xcv*) uses the type III secretion system (TTSS) to inject effector proteins into cells of Solanaceous plants during pathogenesis. A number of *Xcv* TTSS effectors have been identified; however, their function *in planta* remains elusive. The *Xanthomonas* type III effector XopD was demonstrated to deSUMOylate SUMO-conjugated proteins both *in vitro* and *in planta.* Hotson *et al.* 2003 [10] studied this plant-pathogen interaction using DIGE. Here, they provide direct evidence phytopathogenic bacterial TTSS effector *in planta* by demonstrating that the XopD encodes an active cysteine protease with plant-specific SUMO substrate specificity. Divergent metabolism of phenol and succinate in *P. putida* KT2440 was studied using iTRAQ. They induce a global response to aromatic hydrocarbon sources (phenol or benzoate) by up- or down-regulating series of enzymes.

The study was conducted by Surget *et al.* 2012 [13] to find the mechanism involved in UV tolerance using DIGE and also ICAT. The marine bacterium *Photobacterium angustum* S14 on exposing to UVB seen almost a 3-fold change in RecA and a high number of antioxidants, transport proteins. metabolism-related proteins, transcription/translation regulators, chaperonins and proteases. The researchers also compared the copiotroph P. *angustum* and oligotroph *Sphingopyxis alaskensis* on UVB response and global protein expression profiles.

Zimaro *et al.* 2013 [25] used DIGE to compare biofilm and planktonic cells virulence in *Xanthomonas axonopodis.* The researchers found that biofilm formation was associated with major variations in the composition of outer membrane proteins including receptors and transporter proteins. Okanishi *et al.* 2014 [15] used SILAC for studying PTM-Lysine propionylation in *Thermus thermophiles.* Lysine propionylation, a newly discovered type of acylation, occurs in several proteins, including some histones. In this study, they identified 361 propionylation sites in 183 mid-exponential phase and late stationary phase proteins from *Thermus thermophilus* HB8, an extremely Thermophilic eubacterium. Functional classification of the propionyl proteins revealed that the number of propionylation sites in metabolic enzymes increased in late stationary phase, irrespective of protein abundance. Formation of the SICs with low metabolic activity and high survival ability was a survival strategy for *E. coli* O157:H7 against HPCD was illustrated by Bi *et al.* 2017 using ITRAQ.

Kumar *et al.* 2018 [11] used iTRAQ for global proteome expression involving the aerobic unusual '-CoA'-mediated degradation pathway of phenylacetate and benzoate (reported only in 16 and 4–5% of total sequenced bacterial genomes) for lignin degradation. Both *ortho* and *meta* ring cleavage pathways enzymes were detected.

Endocrine disrupting compounds are the predominant environmental contaminants exist in different environments and have significant adverse effects on the reproductive system of animals and humans. Biodegradation using microorganisms as one efficient strategy to remove Endocrine disrupting compounds and *P. putida* SJTE-1 isolated from sludge was able to degrade multiple estrogens efficiently, including estrone, 17β –estradiol and other estrogenic chemicals and bio-transform them into non-estrogenic products. Xu *et al.* 2017 [24] using ITRAQ identified 78 proteins with significant changes in expression; 45 proteins and 33 proteins were up-regulated and down-regulated, respectively

7. Visual Proteomics

Electron Cryotomography is an emerging imaging technique that has unique potential for molecular cell biology (Förster *et al.* 2011). At the present resolution of 4–5 nm, large supramolecular structures can be studied in cellular environments and, in the future, it is possible to map molecular landscapes inside cells in a brief manner. 'Visual proteomics' aims to complement and extend mass-spectrometry-based inventories, and to provide a quantitative description of the macromolecular interactions that underlie cellular functions. The principle involves Cryogenic EM where the sample is cooled to cryogenic temperatures to convert into non-crystalline ("vitreous") ice (Oikonomou *et al.* 2017) [16], tomography involving the imaging of the sample as they are tilted, resulting in a series of 2D images that can be combined to produce a 3D reconstruction

The aim of visual proteomics is to map all of the macromolecular complexes that are found inside a cell in a comprehensive manner. The structures of individual macromolecular complexes are determined by single-particle electron-microscopy, X-ray crystallography or hybrid approaches. Selected structures to resolve reliably using current cryo-electron-tomography methods will be created in

the 'template library' (Nickell *et al.* 2006) [14]. The result will be the protein atlas, not only defining the orientation of each particular complete and spatial distribution but also its coordinates in relation to other structures — the cellular 'Interactome'.

8. Future prospects

The unbiased protein extraction from the cell is pre-requisite to undergo the effective functionality study. Additionally, the direct analysis of protein complexes is yet another, more ambitious step up in the biological hierarchy. There are many insights needed for selective peptide identification like selective reaction Monitoring (SRM). The multiplexing and cost-effective technology are needed in advance when compared to genomics. A long way to go for the field of absolute quantification of membrane proteins on a gel-based approach (Schubert *et al.* 2017) [19].

9. Reference

1. Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature. 2003; 422(6928):198.
2. Aggarwal K, Choe LH, Lee KH. Shotgun proteomics using the iTRAQ isobaric tags. Briefings in Functional Genomics. 2006; 5(2):112-120.
3. Arentz G, Weiland F, Oehler MK, Hoffmann P. State of the art of 2D DIGE. PROTEOMICS–Clinical Applications. 2015; 9(3-4):277-288.
4. Armengaud J. Next-generation proteomics faces new challenges in environmental biotechnology. Current Opinion in Biotechnology. 2016; 38:174-182.
5. Banerjee S, S Mazumdar. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. International Journal of Analytical Chemistry. 2012, 40.
6. Bi X, Wang Y, Zhao F, Zhang Y, Rao L, Liao X *et al.* Inactivation of Escherichia coli O157: H7 by high pressure carbon dioxide combined with nisin in physiological saline, phosphate-buffered saline and carrot juice. Food Control. 2014; 41:139-146.
7. Fenn JB, M Mann *et al.* Electrospray ionization–principles and practice. Mass Spectrometry Reviews. 1990; 9(1):37-70.
8. Förster F, Han BG, Beck M. Visual proteomics. In Methods in enzymology. Academic Press. 2010; 483:215-243.
9. Gillet LC, Leitner A *et al.* Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing. Annual Review of Analytical Chemistry. 2016; 9(1):449-472.
10. Hotson Andrew, Renee Chosed, Hongjun Shu, Kim Orth, Mary Beth Mudgett. Xanthomonas type III effector XopD targets SUMO-conjugated proteins in planta. Molecular Microbiology. 2003; 50:377-389.
11. Kumar M, Verma S *et al.* Genomic and proteomic analysis of lignin degrading and polyhydroxyalkanoates accumulating β-proteobacterium *Pandoraea* sp. ISTKB. Biotechnology for Biofuels. 2018; 11(1):154. doi.org/10.1186/s13068-018-1148-2.
12. Mann M. Innovations: Functional and quantitative proteomics using SILAC. Nature reviews Molecular Cell Biology. 2006; 7(12):952.
13. Matallana-Surget S, Joux F, Wattiez R, Lebaron P. Proteome analysis of the UVB-resistant marine bacterium Photobacterium angustum S14. PLoS One. 2012; 7(8):42299.
14. Nickell S, Kofler C, Leis AP, Baumeister W. A visual approach to proteomics. Nature reviews Molecular Cell Biology. 2006; 7(3):225.
15. Okanishi H, Kim K, Masui R, Kuramitsu S. Lysine propionylation is a prevalent post-translational modification in Thermus thermophilus. Molecular & Cellular Proteomics, 2014, 113.
16. Oikonomou CM, Jensen GJ. Cellular Electron Cryotomography: Toward Structural Biology in Situ. Annual Review of Biochemistry. 2017; 86(1):873-896.
17. Ong SE, Mann M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nature protocols. 2006; 1(6):2650.
18. Ong SE, Mann M. Mass spectrometry–based proteomics turns quantitative. Nature chemical biology. 2005; 1(5):252.
19. Schubert OT, Röst HL *et al.* Quantitative proteomics: challenges and opportunities in basic and applied research. Nature Protocols. 2017; 12:1289. doi.org/10.1038/nprot.2017.040
20. Schulze WX, Gleixner G *et al.* A proteomic fingerprint of dissolved organic carbon and of soil particles. *Oecologia.* 2005; 142(3):335-343.
21. Vasileva D, Suzuki-Minakuchi C, Kosono S, Yoshida M, Okada K, Nojiri H. Proteome and acylome analyses of the functional interaction network between the Carbazole-Degradative plasmid pCAR1 and host *Pseudomonas putida* KT2440. Environmental Microbiology Reports. 2018; 10(3):299-309.
22. Wiese S, Reidegeld KA, Meyer HE, Warscheid B. Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research. Proteomics. 2007; 7(3):340-350.
23. Wilmes P, A Heintz-Buschart *et al.* A decade of Metaproteomics: Where we stand and what the future holds. Proteomics. 2015; 15(20):3409-3417.
24. Xu J, Zhang L, Hou J, Wang X, Liu H, Zheng D *et al.* iTRAQ-based quantitative proteomic analysis of the global response to 17β-estradiol in estrogen-degradation strain *Pseudomonas putida* SJTE-1. Scientific reports. 2017; 7:41682.
25. Zimaro T, Thomas L, Marondedze C, Garavaglia BS, Gehring C, Ottado J. *et al.* Insights into *Xanthomonas axonopodis pv. citri* biofilm through proteomics. BMC Microbiology. 2013; 13(1):186.