Potentials of Proteomics in Plant Improvement

E.E. Okoli¹, M.J. Nworji¹, and C. M. Okoronkwo²

¹Department of Crop Science and Horticulture, Chukwuemeka Odumegwu Ojukwu University, P.M.B. 6059 Awka, Anambra State, Nigeria.
²Department of Agronomy, College of Crop and Soil Sciences, Michael Okpara University of Agriculture, Umudike. P.M.B. 7267, Abia State, Nigeria.

*correspondence: Dr. Evans Ebuka Okoli; ee.okoli@coou.edu.ng

ABSTRACT
The large-scale study of proteins is referred to as proteomics and is often regarded as an emerging technology when compared to genomics or transcriptomics, that is, as not having reached the same level of maturity. While the successful implementation of proteomics workflows and technology still requires significant levels of expertise and specialization, great strides have been made to make the technology more powerful, streamlined and accessible. Many outstanding improvements in the large scale study of proteins and many other but related improvements in plant biotechnology techniques offer various new ways to encourage the usage by plant breeders for crop improvement. A combinatorial approach of accelerated gene discovery through genomics, proteomics, and other important branches of biotechnology, as an applied approach, is proving to be an effective way to speed up the crop improvement programs of the world. This review present and discuss the potentials of proteomics as an effective biotechnological tool for crop improvement programs, especially how it helps in the investigation of both biotic and abiotic stress tolerance mechanisms in plants and this to a very large extent will lead to sustainable agriculture.

Keywords: proteomics; breeding; plant biotechnology; plant improvement; biotechnological tool; stress tolerance mechanism.

INTRODUCTION
World’s population is on a steady increase and almost half of this population estimate lives in a state of hunger. In an effort to eradicate that ugly spot of hunger from the face of the earth, we need to significantly increase the production and supply of food by integrating different elements and strengthening the plant breeding tools for crop improvements (Beddington et al, 2012). A major hurdle for crop improvement programs faced by the plant breeders is a limited gene pool of domesticated crop species.

The identification of potential useful genes across the animal and plant kingdom that could play key roles toward the improvement of important crop traits, generally derived from research in molecular biology including genomics and proteomics, is a crucial step. Such newly discovered genes, when placed into a desired crop species and then utilized for breeding programs, could be a boon to human society.

Proteins are highly complex substance that is present in all living organisms; they are polymers of amino acids and they play important role in metabolic activities. Primary structure of protein is determined by the sequence of specific amino acids, encoded by the mRNA, which directs the proper folding of the polypeptide chain into the secondary structure. One type of secondary structure is the alpha helix, a region of the polypeptide that folds into a corkscrew shape. Beta strands are linear structures of polypeptides, bonding together to form a flat beta sheet. Turns and coils interact chemically with each other to form the unique three dimensional shapes of the proper three dimensional structures that create the final protein. Many proteins, however, have several different polypeptide subunits that make the final active protein. For these proteins, the interactions between the different subunits form the quaternary structure.

One of the most promising developments to come from the study of human genes and proteins has been the identification of potential new drugs for the treatments of disease. This relies on genome and proteome information to identify proteins associated with a disease. The term “proteomics” was first coined in 1995 and was defined as the large-scale characterization of the entire protein complement of a cell line, tissue, or organism (Wasinger et al, 1995).

Proteomics is the large-scale study of proteins particularly their composition, structures, functions, and interactions of the proteins directing the activities of cell (Anderson and Anderson 1998). The main theme of interest is that proteomics gives a much better understanding of an organism than genomics.

Genomics can give a rough estimation of expression of a protein. Most of the proteins function in collaboration with other proteins, and the main goal of proteomics are to identify which proteins interact.

Available Online at www.ijscia.com | Volume 2 | Issue 1 | Jan-Feb 2021
After genomics, proteomics is often considered as the advanced step in the study of biological systems.

It is much more complicated than genomics, mostly because while an organism’s genome is more or less constant, the total protein expression profile always changes with time, micro and macro environmental conditions. Mass spectrometry (MS) has been widely used in forensic science in the identification of compounds, particularly illicit drugs. MS is a technique that allows the detection of compounds by separating ions by their unique mass (mass-to-charge ratios) using a mass spectrometer.

The method relies on the fact that every compound has a unique fragmentation pattern (mass spectrum). The sample is ionized; the sample ions are separated based on their differing masses and relative abundance. Anambra State, Nigeria.

**TYPES OF PROTEOMICS**

Proteomics are classified into different groups based on the protein response under stress conditions:

- **Expression proteomics**
  Expression proteomics is used to study the qualitative and quantitative expression of total proteins under two different conditions. Normal cell and diseased cell can be compared to understand the protein that is responsible for the stress or diseased state or the protein that is expressed due to disease. Typically, expression proteomics studies are addressed to the investigation of the expression protein patterns in abnormal cells (Pandey et al., 2000).

- **Structural proteomics**
  Structural proteomics helps to understand three dimensional shapes and structural complexities of functional proteins. Structural prediction of a protein when its amino acid sequence is determined directly by sequencing or from the gene with a method called homology modelling. Structural proteomics can give detailed information about the structure and function of protein complexes present in a specific cellular organelle. It is possible to identify all the proteins present in a complex system such as membranes, ribosomes, and cell organelles and to characterize all the protein interactions that can be possible between these proteins and protein complexes. Different technologies such as X-ray crystallography and Nuclear Magnetic Resonance (NMR) spectroscopy were mainly used for structure determination.

- **Functional proteomics**
  Functional proteomics explains the understanding of protein functions as well as unrevealing molecular mechanisms within the cell which depend on the identification of the interacting protein partners. The association of an unknown protein with partners belonging to a specific protein complex involved in a particular mechanism would in fact, be strongly suggestive of its biological function (Ho et al., 2002).

Furthermore detailed description of the cellular signaling pathways might greatly benefit from the elucidation of protein-protein interactions in-vivo (Lewis et al., 2000).

**TECHNIQUES INVOLVED IN PROTEOMICS**

In proteomic analysis both analytical and bio-informatics tools are used to characterize protein structure and functions. Analytical techniques 2-D gel electrophoresis, Matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF-MS) can be used. In case of bio-informatics numbers of software tools were used.

- **2-D gel electrophoresis**
  In 2-D gel electrophoresis, protein samples are resolved based on charge, in a step called isoelectric focusing, and then based on molecular weight in second step (Mikkelsen and Cortón 2004). The result is an image in thousands of small spots, each representing a protein.

A good 2-D gel can resolve one thousand to two thousand protein spots, which appear after staining, as dots in the gel. 2-D gel electrophoresis technique is mainly used to compare two similar samples to find specific protein differences.

<table>
<thead>
<tr>
<th>2-D Electrophoresis workflow chart</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample preparation</td>
</tr>
<tr>
<td>First Dimensional separation-IEF</td>
</tr>
<tr>
<td>Second Dimensional separation - Based on Molecular Weight</td>
</tr>
<tr>
<td>Detection</td>
</tr>
<tr>
<td>Protein Excision, Digestion and Identification</td>
</tr>
</tbody>
</table>

Prepare the protein at a concentration and in a solution suitable for isoelectro-focusing (IEF). Choose a method that maintains the native charge, solubility, and relative abundance of proteins of interest. Separate proteins according to isoelectric point (pI) by IEF. Select the appropriate IPG strip length and pH gradient for the desired resolution and sample load. Select appropriate sample loading and separation conditions. Separate proteins according to size by Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE). Select the appropriate gel size and composition and separation conditions. Visualize proteins using either a total protein stain or fluorescent protein tags. Select a staining technique that matches sensitivity requirements and available imaging equipment.

Capture digital images of the 2-D patterns using appropriate imaging equipment and software. Then analyse the patterns using 2-D software. Excise protein spots of interest from the gel digest the proteins, and the digests by MS.

- **Mass Spectrum (MS) Analysis**
  Mass spectrometry is an analytical technique that produces spectra of the masses of the atoms or molecules comprising a sample of material. The spectra are used to determine the elemental or isotopic signature of a sample, the masses of particles and of molecules, and to elucidate the chemical structures of molecules, such as peptides and other chemical compounds. Mass spectrometry works by ionizing chemical compounds to generate charged molecules or molecule fragments and measuring their mass to charge ratios (Sparkman, 2000). MALDI-TOF is the most useful technique for protein identification.

- **MALDI-TOF-MS**
  Matrix Assisted Laser Desorption/Ionisation is a soft ionization technique used in spectrometry, allowing to analysis the biomolecules like DNA, protein, peptides. Biomolecules and synthetic polymers have low volatility and are thermally unstable, which has limited the use of MS as a means of characterization. These problems have been minimized through the development of MALDI-TOF MS, which allows for the mass determination of biomolecules by...
ionization and vaporization without degradation, a Laser beam used to ionize the sample (Wu et al., 1994).

Protein sample have been characterized by HPLC or SDS PAGE by generating peptide maps. These peptide maps have been used as fingerprints of protein or as a tool to know the purity of a known protein in a known sample. Mass spectrometry gives a peptide map when proteins are digested with proteolytic enzymes like trypsin. This peptide map can be used to search a sequence database to find a good match from the existing database.

PROTEOMICS TECHNIQUES OFFER NEW TOOLS FOR PLANT BIOTECHNOLOGY

The knowledge of key proteins that play crucial roles in the proper growth and development of a plant is critical to propel the biotechnological improvement of crop plants. These proteins maintain cellular homeostasis under a given environment by controlling physiological and biochemical pathways.

A search of the published research literature revealed that genomics and proteomics are the two major wheels that keep the discovery of novel genes rolling, which can eventually be placed into the pipeline for crop improvement programs.

Two-dimensional electrophoresis (2-DE) and mass spectroscopy (MS), two of the most widely used proteomics methods, are used to catalogue and identify proteins in different proteome states or environments.

Advances in 2-DE have been extremely helpful in bringing proteomics close to biotechnological programs; however, due to some drawbacks and disadvantages associated with gel-based proteomics, e.g., labour intensiveness, insensitivity to low-copy number proteins, low reproducibility and the inability to characterize complete proteomes, many gel-free proteomic techniques have also become a valuable tool for scientists (Baggeman et al., 2005; Lambert et al., 2005; Scherp et al., 2011; Jayaraman et al., 2012).

POTENTIALS OF PROTEOMICS AS A BIOTECHNOLOGY TOOL IN CROP IMPROVEMENT PROGRAMS

- Molecular Markers are to Assist Plant Breeders

Proteomics offers novel gene (DNA) identifications to plant biologists and breeders. Marker assisted selection (MAS), which is the employment of DNA markers in a plant breeding program, has extensively been used to select desired genes/quantitative trait loci (QTLs) in the development of a comparatively superior breeding line (Collard and Mackill, 2008). Damerval et al. (1994) used an approach that brought proteomic and MAS components together; they identify protein quantity loci (PQL) that explained some of the spot intensity variation, out of the 72 proteins analyzed, 70 PQLs were identified for 42 proteins, 20 of which had more than one PQL. This type of approach is especially useful in breeding programs because, through intensive breeding selection, lines could be available with differing phenotypic degrees that help in drawing correlations between responsive genes and observed stress tolerance phenotypes. This correlation can further be verified by analysing advanced mapping populations such as recombinant inbred lines (RILs), near isogenic lines (NILs), and double haploid lines (Salekdeh and Komatsu, 2007).

Furthermore, the co-segregation of a protein and the QTL (or the trait) can be studied in the two parental lines from which the mapping populations were developed. Finally, the plant breeders should be able to integrate the selected genes in marker-assisted breeding programs to improve the trait under study (Salekdeh and Komatsu, 2007).

The major limitation of this technique is that it works only within the same species because the parents need to be cross-compatible to transfer the superior genes/alleles through this molecular breeding approach. Under such limitations, embryo rescue or genetic engineering, which has no boundaries for gene transfer, could be very useful (Varshney et al., 2011).

PROTEOMICS HELP THE INVESTIGATIONS OF ABIOTIC AND BIOTIC STRESS TOLERANCE MECHANISMS

As with any living organism, crop plants also have to cope with various biotic and abiotic stress conditions. Contrary to greenhouse nurseries, plants in the field experiences a combination of various biotic and abiotic stresses either concurrently or at different developmental stages throughout the growing season (Tester and Bacic, 2005; Mittler, 2006). A recent estimate suggested that the increased temperatures of the past two decades have caused a loss of approximately 5 billion US Dollars by impacting the yields of major food crops such as wheat, rice, maize, and soybeans (Peng et al., 2004). Temperatures reaching 35°C in the field cause rice and maize to show sterility. Such high heat conditions in the field also lead to flowering and fruiting failure in other crops. Molecular plant physiologists know very well that heat stress increases membrane damage and impairs metabolic functions (Taiz and Zeiger, 2010). A plant breeder needs to activate the proper protection systems in a crop plant to enable the survival of the plant’s cells under such heat stress conditions.

Heat stress tolerance is a complex mechanism and is controlled by multiple genes and proteins involving a number of physiological and biochemical changes in the cell, e.g., adjustments in the membrane structure and function, tissue water content, protein composition, lipids, and primary and secondary metabolites (Huang and Xu, 2008). Global proteomic profiling projects are useful techniques for increasing the knowledge base of plant breeders. For example, a study comparing various wheat cultivars with different heat tolerance capabilities revealed low molecular weight (16–17 kDa) heat shock protein (HSPs) and other metabolic proteins crucial for the heat tolerance phenotype (Majoul et al., 2004).

Proteins from the HSP family and the transcription factors upstream of these HSPs have been found to have crucial roles in providing thermo-tolerance to the crop.

Disarming the function of HSP100 by introducing an antisense construct in tomato plants resulted in their poor survival under heat stress conditions (Yang et al., 2006). However, in another study, transgenic lines over expressing a different HSP protein (HSP70) showed superior thermo-tolerance in soy bean plants (Zhu et al., 2006). Furthermore, protein–protein interaction studies have proved that HSP90 interacts with calmodulin-binding protein (CBP) (Virdi et al., 2009).

Thus, the studies by Zhang et al. (2009) showed that the knock down of calmodulin resulted in reduced thermo-tolerance. Proteins other than HSPs, e.g., CBP in the above study, have been identified in other proteomic studies as differentially expressed proteins during heat stress conditions.

Süle et al. (2004) proposed S-adenosyl methionine synthetase as a molecular marker for screening heat-tolerant germplasms. Even with this information, knowledge on the systemic response of plants during heat stress remains limited because plant perception and response to a single stress is different than to a combination of multiple stresses. There is another major constraint to world agriculture in the form of limited water...
availability for crop irrigation. Recent climate variability from year to year predicts a worsening situation in the future. World climatologists predict that global warming will result in more frequent and severe droughts in the coming years. Drought stress causes a decrease in carbon usage by the photo-synthetic machinery that result in net yield losses on the farm. Physiological experiments have shown that drought conditions inhibit plant photosynthesis within a short time of a limited water supply resulting in a drop in the CO₂ assimilation rates (Ribas-Carbo et al., 2005). To minimize water loss, plants need to close their stomata under water deficient conditions.

The guard cells help the plant in the process of controlling the opening and closing of the stomata. The closure/opening of the stomata is controlled by the plant hormone, abscisic acid (ABA).

In a plant cell, ABA flux concentrations are controlled in response to the availability of water to the plant. ABA has been found to play an indispensable role in the plant response to drought conditions by inducing many transcription factors.

In this direction, the guard cell proteome profiling by Zhao et al. (2008) revealed 336 proteins responsive to water stress conditions, with a further 52 proteins considered to be signalling proteins. Abiotic stresses in general cause a water deficit condition in cells that results in a myriad of complex cellular and physiological responses at the plant cellular and organismal levels. In general, the net photosynthesis rate is reduced either because of stomatal closure or via metabolic impairment (Reddy et al., 2004).

The changes in mitochondrial respiration and the photosynthetic electron transport chain lead to the generation of highly toxic reactive oxygen species (ROS), such as super oxides and peroxides, and cause chemical damage to the DNA and proteins. This damage has serious effects on cellular metabolism (Mittler, 2002).

During evolution, plants have developed several strategies to address ROS, e.g., avoidance by anatomical adaptation, photosynthesis suppression and photo-system and antenna protein complex modulations. Several metabolites, such as ascorbate and glutathione, and enzymes, such as peroxidases and super oxide dismutases, help to scavenge the ROS (Mittler, 2006).

Another plant strategy to address drought conditions is to maintain the turgor pressure of plant cells by the over production of osmolytes, such as proline, glycine betaine, and trehalose.

These metabolites provide secondary protective effects to proteins against misfolding (Hare et al., 1998). Moreover, dehydration responsive proteins, such as dehydrins and HSPs, are over produced to protect the intracellular metabolic machinery (Wang et al., 2003).

In short, with such a wealth of knowledge, drought-tolerant plants can be generated by the modification of these mechanisms, e.g., ABA signalling can be adjusted for the better survival of a crop plant under such stress conditions. The level of sphingosine-1-phosphate, a messenger molecule, is controlled by ABA through the sphingosine kinase protein. In another study using a sphingosine-1-phosphate lyase mutant, the accumulation of sphingosine-1-phosphate decreased the fresh weight drop in the CO₂ within a short time of a limited water supply resulting in a drop in the CO₂ assimilation rates (Ribas-Carbo et al., 2005).

To minimize water loss, plants need to close their stomata under water deficient conditions. A wealth of knowledge, drought stress conditions, with a further 52 proteins considered to be signalling proteins. Abiotic stresses in general cause a water deficit condition in cells that results in a myriad of complex cellular and physiological responses at the plant cellular and organismal levels. In general, the net photosynthesis rate is reduced either because of stomatal closure or via metabolic impairment (Reddy et al., 2004).

The changes in mitochondrial respiration and the photosynthetic electron transport chain lead to the generation of highly toxic reactive oxygen species (ROS), such as super oxides and peroxides, and cause chemical damage to the DNA and proteins. This damage has serious effects on cellular metabolism (Mittler, 2002).

During evolution, plants have developed several strategies to address ROS, e.g., avoidance by anatomical adaptation, photosynthesis suppression and photo-system and antenna protein complex modulations. Several metabolites, such as ascorbate and glutathione, and enzymes, such as peroxidases and super oxide dismutases, help to scavenge the ROS (Mittler, 2006).

Another plant strategy to address drought conditions is to maintain the turgor pressure of plant cells by the over production of osmolytes, such as proline, glycine betaine, and trehalose.

These metabolites provide secondary protective effects to proteins against misfolding (Hare et al., 1998). Moreover, dehydration responsive proteins, such as dehydrins and HSPs, are over produced to protect the intracellular metabolic machinery (Wang et al., 2003).

In short, with such a wealth of knowledge, drought-tolerant plants can be generated by the modification of these mechanisms, e.g., ABA signalling can be adjusted for the better survival of a crop plant under such stress conditions. The level of sphingosine-1-phosphate, a messenger molecule, is controlled by ABA through the sphingosine kinase protein. In another study using a sphingosine-1-phosphate lyase mutant, the accumulation of sphingosine-1-phosphate decreased the fresh weight loss of plants under drought stress conditions by controlling water loss from the stomata (Nishikawa et al., 2008). Hajheidari et al. (2005) report the predominance of proteins that are related to ROS management and protein stability after investigating the proteomic profiling of field-grown plants under drought stress conditions.

All of the studies described above, and many that are not included here are decent examples that prove that proteomics is highly capable of discovering novel genes/proteins that could be potential candidates for further studies through biotechnological approaches, with time, the data sets for crop proteomics will strengthen further and that researchers will be able to see examples in which such proteomic-based knowledge is used directly for the improvement of the stress tolerance of a crop plant (Agrawal et al., 2012).

CONCLUSION

During the recent past, world agriculture has come under more climatic variability along with less arable land availability per person, which compounds the stress situation on producer groups. In the present scenario, pressure is building upon the plant breeders and plant biologists to come up with "smart crop varieties" that are better suited genotypes with the ability to withstand a wider range of climatic variability to tackle the food insecurities of future generations along with maintaining/exceeding quality parameters.

Proteomics is also used to know plant-insect interactions that help identify candidate genes involved in the defensive response of plants to herbivore and this aspect can be exploited by plant breeders.

REFERENCES


