REVIEW ARTICLE

Biochemical Systematic: From Amino Acids to Proteomics

ABD EL-ZAHER MOHAMMED ABASERY MUSTAFA[1,2*]
[1]*Department of Botany & Microbiology, College of Science, King Saud University, Riyadh, Saudi Arabia.*
[2]*Botany Department, Faculty of Science, Tanta University, Tanta, Egypt*
**Correspondence: amus@ksu.edu.sa, Mobile No. +966 548549976*

ABSTRACT

Biochemical markers use proteins variability to study the plants genetic diversity and to evaluate the phylogenetic relationships of accessions, populations, subspecies, species of cultivated and wild plants. The biochemical techniques are rapid, accurate and dependable. They have been revealed to be useful in: ensuring the genetic relations between plant taxa, evaluation of the genetic diversity, phylogenetic relationships, germplasm geographical origin, new cultivar identification and description and saving time in the tests of distinctness, uniformity, stability of a candidate cultivar. Moreover, they are utilized in new variety registration and applications of plant variety rights. The current mini-review will focus on consideration of biochemical analyses in plant systematic from the early era up to now, giving more attention to the used techniques.

Keywords: SDS-PAGE, Peptide fingerprint, Serological methods, western blotting, Proteomics

INTRODUCTION

Plant taxonomy was first applied depending on morphological traits. The instability of the morphological characters and its susceptibility to environmental fluctuations resulted in species complexes and lack of clarity in taxonomic relationships, genetic diversity and phenetic classification of plants. For example, the leaf morphology and floral structure of *Heuchera* species (Order: Rosales, Family: Saxifragaceae) showed variation that was attributed to geographic aspects. Trying to overcome such shortage, Taxonomists started to utilize traits other than morphological traits to study taxonomic relationships of plant taxa. The other characters were derived from the micro- and macro-molecules in the plant tissue. The main objective of this review is the consideration of protein molecules in plant systematic.

Systematic Aspects of the Structure of Plant Proteins: Since the enzymes which catalyze the chemical transformations of plant cells are proteins, it can be safely inferred that an enormous number of different plant proteins exist. They show in a pronounced way the general property of plant proteins, namely a relatively low content of Sulphur amino acids compared with animal proteins. Most plant proteins function as enzymes, others act as structural components, storage compounds and possibly as regulatory compounds. Legume proteins are classified in several ways depending on the purpose of the classification, which could be methodological, chemical, functional, or locational[1-3]. Methodological classifications include those based on solubility, electrophoretic mobility, chromatographic or immunological properties. Chemical classifications are based on chemical properties and usually separate proteins into simple and conjugated proteins, the latter being subdivided into glycol-proteins, lipo-proteins, metallo-proteins, flavoproteins and nucleoproteins; examples from each of these groups occur in all plant species. A classification based on function and usually separate proteins into enzymes, structural and storage. In this classification enzymes constitute by far the largest category. Locational classifications refer either to location within the plant or location within the cell. Herein, we will adopt the methodological classification to address the implications of plant proteins on systematics, since it is well suited to the logical way of thinking of the human mind, and it is well publicized among whom working in this area of research. Furthermore, we indented to follow the development in using protein molecules in plant systematics from the early days of this branch of science up to this era.

Systematic Considerations of Proteins Amino Acid Composition: The amino acids makeup of the seed proteins of a number of species (*Glycine max*, *Vigna radiata*, *Phaseolus mungo*, *Cicer arietinum* and *Lens esculenta*) belonging to tribes Glycineae, Phaseoleae and Vicieae, were quantified in reference to standard amino acid composition[4-6]. The tribe Phaseoleae tends to be higher in lysine and lower in arginine than species of the Vicieae. *Glycine max* more nearly resembles members of the Vicieae than the Phaseoleae in its lysine, histidine and phenylalanine content, although the arginine content tends to be more similar to that of the Phaseoleae. Within the tribe Phaseoleae, there is a close similarity between the species of *Phaseolus* and those of *Vigna* and these differ from the analysis of *Dolichos lablab,* particularly with respect to the content of arginine, lysine, leucine and methionine.

Systematic considerations of peptide fingerprint data of plant seed proteins: Carboxymethylated protein is hydrolyzed into peptides of various sizes when incubated with proteolytic enzymes. These peptides may be partially separated by chromatographic methods and then separated further on paper by first-dimensional chromatography combined with electrophoretic separation in the second dimension. The separated peptides can be located using chromogenic reagents and hence a "map" of the distribution of the peptides obtained. This map will be characteristic for the original protein.

The globulin fraction of seed proteins from five genera belonging to the tribe Vicieae were exposed to tryptic digestion. Comparing the resulting peptide fingerprints revealed close similarity, while they were different from *Phaseolus vulgaris* peptide pattern, and two identical peptide patterns for two species of *Canavalia*. Proteomic analysis was also utilized to study the taxonomic relationships of some members of the tribe Abreae to members of the tribe Vicieae and the tribe Phaseoleae.

Comparisons were made among fingerprints patterns of the globulin fractions of both *Abrus precatorius* (Abreae), five genera of the tribe Vicieae and two genera of the tribe Phaseoleae (*Phaseolus vulgaris* and *Canavalia* spp.). While the results revealed conspicuous differences between the fingerprints pattern of the globulin fraction of the members of each two of the three tribes, greater similarity was exhibited between *Abrus precatorius* (Abreae) and members of the tribe Vicieae rather than that between *Abrus precatorius* and members of the tribe Phaseoleae. This might agree with the conclusion that *Abrus precatorius* is related to but not a member of the Vicieae, based on the electrophoretic properties of globulin fractions[7,8]. They were also able to arrange the genera of this tribe Vicieae as follows: *Vicia* and *Lathyrus* had almost identical patterns; *Lens* had a slightly different pattern, *Pisum* was slightly more different and *Cicer* was the most different from *Vicia*, *Lathyrus*, *Lens* and *Pisum*.

Moreover, the fingerprint patterns of tryptic digests of the vicilin and legumin globulin proteins prepared from the seeds of three Vicieae species, *Pisum sativum*, *Vicia faba* and *Cicer arietinum* were compared. This comparison revealed close similarity among significant parts of amino acid sequences of both vicilin and legumin in *Pisum sativum* seed proteins. This study even revealed that both proteins consist of a number of subunits. The homology extent between vicilin and legumin proteins of *Pisum sativum* was also exhibited in both *Vicia faba* and *Cicer arietinum* but not in any other one species[9].

Systematic Considerations of Electrophoretic Techniques: This is one of the most widely used techniques for protein data comparisons. It utilizes the presence of ionizable groups on the surface of protein molecules; because of these groups, proteins will migrate in solution of suitable pH when subjected to an electrical field. Separation and comparison of different proteins can be made therefore in terms of mobility under standard conditions. The separation is carried out usually on a supporting medium; paper and various kinds of gels have been used. However, proteins tend to streak on paper, and the more powerful methods of electrophoresis on gels and particularly acrylamide gels, have proved more useful. On gels, there is, in addition to separation by electrical properties, a molecular sieving effect, which increases the resolution ability of the technique. On polyacrylamide gels, the sieving effect is more pronounced than on other gels such as agar or starch; hence a higher efficiency of discrimination is obtained.

The banding profiles derived from bulk seed samples of accessions of *Lathyrus sativus*[10], *Lathyrus inconspicuous*[11], *Vicia faba*[12], *Amaranthus*[13], Flax[14], *Trifolium* spp.[15], *Gossypium barbadense*[16], *Phaseolus* spp.[17] and Lens[18] gave rise to distinctive seed protein banding patterns.

Systematic considerations of serological methods: In the early days of using proteins as a source of systematic information, serology was the only technique used to derive these characters. Serological work on plants started with using the precipitin reaction in which protein extracts from a given plant are injected into a rabbit, causing the formation of antibodies in its blood serum. The serum containing antibodies (antiserum) is then mixed with a suspension of the protein to be tested (the antigen) and the antibody and the antigen react forming a precipitate; hence the term precipitin reaction. The strength of the reaction is regarded as a measure of the protein similarity of the samples and therefore to some extent of the plants being compared.

The advent of Ouchterlony double immuno-diffusion, makes the determination of immunological identity possible between the species. In this technique, a well is cut in an agar gel, then surrounded with six wells in a circle. Antigen in the central hole and antibody in the surrounded six holes, are allowed to diffuse towards each other in the gel. By this way, we can test the sera of four samples, and two positive control sera in each gel.

Both legumins and vicilins purified from the legume plants belonging to the tribe Vicieae; *Vicia faba, Cicer arientinum, Pisum sativum* and *Lens esculentum,* were allowed to diffuse against anti-legumin and anti-vicilin antisera separately in an Ouchterlony double immunodiffusion test. All the tested proteins gave a precipitin arc[19], indicating a degree of similarity between both legumin and vicilin purified from Pisum *sativum* and those of the other members of Vicieae. Their results revealed some variation in the degrees of similarity exhibited between the legumins or vicilins of the different studied plants. While the legumins of *Vicia faba* and *Lens esculentum* showed a total identity, a partial identity was exhibited for legumin of *Cicer arientinum.* Also, vicilin from *Pisum sativum* exhibited a partial identity with those from *Vicia faba*, *Lens esculentum* and *Cicer arientinum*. However, the degree of this partial identity was higher in case of both *Vicia faba* and *Lens esculentum* than that of *Cicer arientinum*, the reaction of which, produced a larger spure. The data obtained from such comparisons can provide considerable insight into taxonomic relationships.

Another recent development in serology was the use of "Western Blotting". In this technique, electro-blotting is applied to transfer the SDS-PAGE separated proteins to a nitrocellulose paper, which is immersed in a solution of specific antibodies, visualized by peroxidase-coupled antibodies, then stained with 4-chloro-1-napthol[20]. This technique opens the door to utilize the subunit composition of specific proteins as a biochemical marker in studying taxonomic relationships. Taxonomic relationships among some members of the tribe Vicieae were actually studied by utilizing "Western Blotting" of their legumin seed proteins[19]. This study revealed a very strong reaction between *Pisum* legumin antiserum and all the legumin subunits of both *Vicia faba* and *Lens esculentum* under both reducing and non-reducing conditions. *Cicer* exhibited some variation, where its low molecular weight legumin subunit showed no reaction with *Pisum* legumin antiserum under non-reducing conditions. Moreover, some bands of its legumin basic (β) subunits couldn't react with *Pisum* legumin antiserum under reducing conditions. These results agreed with those obtained by using comparative double diffusion technique, leading to a convincing interpretation for the variation revealed between the reactions of both partial identity with *Cicer* and total identity with the rest of the members of the tribe Vicieae. Such kind of studies lead also to the recommendation of classifying of *Cicer* as a member of the tribe Vicieae rather than neither other tribes nor a separate tribe Cicerideae[19].

Potential Usefulness of Isoenzymes/ Allozymes In Systematics: Isoenzymes/allozymes analyses have been widely applied to measure levels of variation within and among populations, accessions and species. For example, they were used in assessment of genetic relationships within *Brassica rapa* subspecies, *Medicago* species and *Lathyrus* species[21-22] and to study the genetic diversity in cultivated *Amaranthus* species and their wild relatives[23].

Potential Value of Proteomic Technique in Systematics: In proteomics, specific properties of a large number of expressed polypeptides are analyzed. These properties include identity, quantity, activity and molecular interactions. This technique has proven to be a successful tool in assessing the genetic proximity and phylogenetic relationships among the different plant taxa at the genus level and inter- and intra-specific levels[24].

Genetic proximity was evaluated between the two species *Quercus petraea* and *Quercus robur* by utilizing proteomic data[25]. They compared the 2D-PAGE protein pattern of 23 oaks of the two species from six European countries representing the natural geographic distribution of white oaks in Europe. 530 polypeptide spots were scored, 101 spots of them were polymorphic. In this study, low level of genetic differentiation was revealed between the two species as indicated by the very close interspecific and intraspecific distances; the interspecific dissimilarity was 0.36, while the intraspecific dissimilarities were 0.35 and 0.33 for *Q*. *petraea* and *Q*. *robur* respectively. Proteomics were also utilized to study genetic distance between *Arabidopsis* and other five species from the family Brassicaceae[26]. Other studies discriminated three wheat varieties by quantitative and qualitative analysis of their 2D-PAGE protein patterns[27-29]. Moreover, the proteome approach was used to study interspecific relationships[30]. The phloem proteomic patterns of both resistant Manchurian ash and susceptible black, green, and white ash were compared using the Difference Gel Electrophoresis (DIGE). The differences among these proteomic patterns revealed a close association with the phylogenetic relationships among the four species studied.

Data availability: The paper contains all the relevant data

Funding and support: There is no funding or support.

Competing interests: There are no competing interests for the author (Author declaration).

REFERENCES

1. Nelson DL, Cox MM. Principles of biochemistry, 4th Ed. W.H. Freeman and Company. 2006 Nov; Ch3:75-115.
2. Kessel A, Ben-Tal N. Introduction to proteins: structure, function, and motion. CRC Press. Sec. Edition. 2018 March.
3. Tseng YY, Li W-H. Classification of protein functional surfaces using structural characteristics. Proceedings of the National Academy of Sciences. 2012 Jan; 109:1170-1175.
4. Velusamy K, Mohan VR. Nutritional and anti-nutritional assessment of underutilized legume D. lablab var. vulgaris L. Bangladesh Journal of Scientific and Industrial Research. 2013 Jul; 48:119-130.
5. Kilonzi SM, Anselimo O, Makokha AO, Kenji GM. Physical characteristics, proximate composition and anti-nutritional factors in grains of lablab bean (*Lablab purpureus*) genotypes from Kenya. Journal of Applied Biosciences. 2017 Jun; 114: 11289-11298
6. Thakur S, Shrivastava SK, Shrivastava M. Amino acid profile of some new varieties of leguminous seeds. IIOABJ. 2017 Aug; 8:48–53.
7. Jackson P, Milton JM, Boulter D. Fingerprint patterns of the globulin fraction obtained from seeds of various species of the Fabaceae. New Phytol. 1967 Jan; 66:47-56.
8. Boulter D, Thurman DA, Derbyshire E. A disc electrophoretic study of globulin proteins of legume seeds with reference to their systematics. New Phytol. 1967 Jan; 66:27-36.
9. Jackson P, Boulter D, Thurman DA. A comparison of some properties of vicilin and legumin isolated from seeds of *Pisum sativum*, *Vicia fab* a and *Cicer arietinum*. New Phytol. 1969 Jan; 68: 25-33.
10. Sammour RH, Mustafa A, badr S, Tahr W. Genetic variations in accessions of *Lathyrus sativus* L. Acta Bot. Croat. 2007 Jan; 66, 1–13.
11. Radwan S, El-Koly A, Sammour RH. Genetic variation among accessions of *Lathyrus inconspicuous* (L.) as revealed by SDS Polyacrylamide Gel Electrophoresis. Acta Agriculturae Slovenica. 2013 March; 101:21-30.
12. Sammour RH. An electrophoretic analysis of the variation in the seed storage proteins of *Vicia faba* L. Cultivars. Fedds Repertorium. 1992 Jan; 103:555-557.
13. Sammour RH, Hamoud MA, Abd Alla SA. Electrophoretic variations in *Amaranthus*. Bot. Bull. Acad. Sin. 1993 Jan; 34:37-42.
14. Sammour RH. Flax seed proteins: Comparison by various PAGE-techniques in slabs. Journal of Agronomy and Crop Science. 1988 Apr; 160:271-276.
15. George NM, Ghareeb A, Fawziand NM, Saad S. Electrophoretic pattern of seed proteins in *Trifolium* L. and its taxonomic implications. Bangladesh Journal of Plant Taxonomy. 2013 Jun; 20: 19–26. doi: 10.3329/bjpt.v20i1.15461.
16. Sammour RH. Using SDS-PAGE in Identification of *Gossypium barbadense* L, cultivars. Egypt. J. Bot. 1990 33:169- 174 (Published on ResearchGate on June, 1994).
17. Sammour RH, Radwan SA, El-Koly A. Genetic diversity in *Phaseolus* spp. as revealed by SDS-PAGE markers. Plant Genetic Resources Newsletter. 2007 Sep; 151:69-75.
18. Sammour RH. Species relationships in genus *Lens* as indicated by electrophoresis, -a reappraisal. Feddes Repertorium. 1994 Jan; 105:283-286.
19. Sammour RH. Molecular reassessment of relationships within Vicieae using electrophoretic and immunochemical techniques. African Crop Science Journal. 2005 Sep; 13:27-39.
20. Towbin H, StaehelinT, Gorden J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets; procedures and some applications. Proceedings of the National Academy of Sciences USA. 1979 Sep; 76:4350-4354.
21. Karam MA, Morsi YS, Sammour RH, Refaat MA. Assessment of genetic relationships within *Brassica rapa* subspecies based on polymorphism. Int J Curr Microbiol Appl Sci. 2014 March; 3:1-10.
22. Karam MA, Sammour RH, Ahmed MF, Ashour FM, El-Sadek LM. Relationships among some *Medicago* species as revealed by isozyme electrophoresis. J. Union Arab Biol. 1999 Apr; 9: 269-279.
23. Safaa R, Salwa B, Mohammed M, Sammour RH. Isozyme analysis of the genetic diversity in cultivated and wild species of *Amaranthus*. Acta Botanica Hungarica. 2012 Sep; 54:391-408.
24. Ishtiaq M, He Q., Wang Y, Xiao PG, Cheng YY. Biosystematics and plant proteomics: role of proteomics in plant phylogenetic analysis. Pakistan Journal of Biological Sciences. 2007 Oct; 10: 3487-3496.
25. Barreneche T, Bahrman N, Kremer A. Two-dimensional gel electrophoresis confirms the low level of genetic

differentiation between *Quercus robur* L. and *Quercus petraea* (Matt.). Liebl. For. Genet. 1996 Jul; 3: 89-92.
26. Marques B, Chane-Favre SL, Zivy M, Thiellement H. Comparative proteomics to establish genetic relationships in the Brassicaceae family. Proteomics 2001 Nov; 1: 1457-1462.
27. Jacobsen S, Nesic L, Petersen M, Sondergaard I. Classification of wheat varieties: Use of two-dimensional gel electrophoresis for varieties that cannot be classified by matrix assisted laser desorption/ionization-time of flight-mass spectrometry and an artificial neural network. Electrophoresis. 2001 May; 22: 1242-1245.
28. Zivy M, Thiellement H, de Vienne D, Hofmann JP. Study on nuclear and cytoplasmic genome expression in wheat by two-dimensional gel electrophoresis. Theor. Applied Genet. 1983 Jul; 66:1-7.
29. Zivy M, Thiellement H, de Vienne D, Hofmann JP. Study on nuclear and cytoplasmic genome expression in wheat by two-dimensional gel electrophoresis. Theor. Applied Genet. 1984 Jul; 68: 335-345.
30. Whitehill JGA, Popova-Butler A, Green-Church KB, Koch JL, Herms DA and Bonello P. Interspecific Proteomic Comparisons Reveal Ash Phloem Genes Potentially Involved in Constitutive Resistance to the Emerald Ash Borer. PLoS ONE 2011 Sep; 6(9): e24863. doi:10.1371/journal.pone.0024863