academicJournals

Vol. 12(17), pp. 2157-2163, 24 April, 2013
DOI: 10.5897/AJB12.2725
ISSN 1684–5315 © 2013 Academic Journals
http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Genetic diversity, taxonomy and legumins implications of seed storage protein profiling in Fabaceae

Conrad A. Omonhinmin* and Oluwaseun O. Ogunbodede

Applied Biology and Biotechnology Unit, Department of Biological Sciences, College of Science and Technology, Covenant University, Ota, Ogun State, Nigeria.

Accepted 26 March, 2013

Proteomic evidences can be pivotal to the discovery of new plant proteins and plant relationships, due to the diversity of form it can reveal. Seed storage protein profiles of 20 Fabaceae species: 4 grain-legumes and 16 non-pulses; of 16 genera and 10 tribes were analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) to estimate protein content diversity and the possible genetic relatedness. 28.3% similarity and 71.7% proteomic polymorphism was scored for the species. The high variability expressed by the lot reflects the genetic diversity amongst Fabaceae population. Dendrogram based on the proteomic data clustered the species into four groups. Aside two species, *Albizia lebbeck* and *Albizia zygia* belonging to the tribe *Ingeae* and those of the tribe *Caesalpinieae*, the other species clustered with several other non-traditional cohorts resulting in a rearrangement that showed least semblance with phylogenetic relationships based on traditional morphology taxonomic delimitation. The similarity in profiles can be preliminarily forensic for proteins of importance whether for nutritional, industrial or for improvement of existing crops or for entirely new plants as crops. The protein mix, and the resultant relationship based on seed storage proteins instigates a review of erstwhile taxonomic, agricultural and research perspectives for the Fabaceae.

Key words: Fabaceae, seed protein, polymorphism, genetic diversity, taxonomy, single protein.

INTRODUCTION

Legumes vary from annual and perennial herbs to shrubs, trees, vines/lianas, and even a few aquatics; in size from some of the smallest plants of deserts and arctic or alpine regions to the tallest of rain forest trees as well as constitute conspicuous, and often dominant component of most of the vegetation types distributed throughout temperate and tropical regions of the world (Rundel, 1989). Legumes are particularly diverse in tropical forests and temperate shrub lands with a seasonally dry or arid climate. This preference for semi-arid to arid habitats is related to a nitrogen-demanding metabolism (Sprent and McKey, 1994; Sprent, 2001).

Taxonomically, the family Fabaceae is traditionally divided into three subfamilies, the Caesalpinioideae, Mimosoideae and Papilionoideae; a recognition that is based mainly on floral characteristics with 39 tribes and some 670 genera recognized (Polhill and Raven, 1981; Polhill, 1994). However, recent update of the tribal and generic re-evaluation of the classification of Fabaceae, have resulted from more than 10 years of intensive molecular phylogenetic studies; recognizes 36 tribes, 727 genera and 19,327 species (Lewis et al., 2005).

*Corresponding author. E-mail: conrad.omonhinmin@covenantuniversity.edu.ng or aomoconrad@gmail.com. Tel: +2348056512546.

Table 1. Fabaceae taxa used in seed-protein electrophoresis analysis.

S/N	Sub family	Species	Voucher number	Status	Geographical distribution/sources
1	*Mimosoideae*	*Acacia auriculiformis*	CU0025	Cultivated	South west, Nigeria
2		*Adenanthera pavonina*	CU0026/13	Cultivated	BGBM, Germany*
3		*Acacia scorpioides*	CU0081/15	Cultivated/Wild	BGBM, Germany*
4		*Pentaclethra macrophylla*	CU0056	Wild/Cultivated	Southern Nigeria
5		*Albizia lebbeck*	CU0030	Wild	Southern Nigeria
6		*Albizia zygia*	CU0031	Wild	Southern Nigeria
7	*Caesalpinioideae*	*Bauhinia tomentosa*	CU0027	Cultivated	Southern Nigeria
8		*Bauhinia purpurea*	CU0028/16	Cultivated	Southern Nigeria.
9		*Delonix regia*	CU0032	Cultivated/Wild	Southern Nigeria
10		*Dialium guineense*	CU0033	Wild	Southern Nigeria
11		*Cassia fistula*	CU0018	Cultivated	South West
12		*Cassia javanica*	CU0019/08	Cultivated/Wild	BGBM, Germany*
13		*Caesalpinia pulcherrima*	CU0020	Cultivated	Southern Nigeria
14		*Afzelia africana*	CU0079	Wild	Southern Nigeria
15		*Peltophorum pterocarpum*	CU0055	Cultivated/Wild	Southern Nigeria
16		*Senna semia*	CU0064	Wild	South West, South South, Nigeria
17	*Papilionoideae /Faboideae*	*Phaselous vulgaris*	CU0077	Cultivated	Southern Nigeria
18		*Arachis hypogaea*	CU0080	Cultivated	Southern Nigeria
19		*Vigna unguiculata*	CU0076	Cultivated	Southern Nigeria
20		*Amorpha fruiticosa*	CU0078/108	Cultivated	BGBM, Germany*

* Boanisscher Garten, Botanischer Museum, Berlin Germany.

These genomic or proteomic studies are considered more reliable than morphological or cytological evidences, particularly, sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE) of total seed proteins widely applied in resolving systematic relationships and for inter and intra specific studies (Karihaloo et al., 2002; Yousaf et al., 2008).

Seed proteins are physiologically stable, easy to handle and they operate at the level of gene product where the environment has very little influence (Javaid et al., 2004; Iqbal et al., 2005). These proteins are expressed form of genome and can be used as biomarkers and were properly traced between specimens, they can lead to the identification of proteins of industrial, medicinal or nutritional prospective applications.

Exploring plants' genetic potential is becoming a frontline study, because the information it provides on depleted gene pool of cultivated plants and the genetic erosion that have accompanied human developmental tendencies, are integral to making informed decisions about plant protection, conservation and improvement programs.

The objective of the present study was to evaluate the genetic diversity, taxonomical relationships, and possible exposition of important proteins among 20 Fabaceae species, using SDS-PAGE analysis of seed storage proteins.

MATERIALS AND METHODS

Samples and protein extraction

Seed samples of 20 species of the plant family Fabaceae were received from field surveys collections of eight states in southern Nigeria (8.5°N – 6.45°N; 3.38°E – 7.5°E); and seed lot from Boanisscher Garten, Botanischer Museum, Berlin Germany (Table 1). Seed whole protein extraction was carried out with Norgen All-in-One purification kit® Norgen Biotek Corporation. The flow-through proteins content were stored at -200°C until separation analysis. Gel was fixed with a 500 mL of USP-grade 95% (v/v) ethanol in water and stained in 0.1% (w/v) Coomassie blue R350, 20% (v/v) methanol and 10% (v/v) acetic acid and afterward resultant banding were capture with a digital transilluminator.

Analysis

The numbers of monomorphic and polymorphic protein bands were scored for each sample based on staining intensity as: low, medium and high, which is an index of the subunit constituents of the protein bands. A similarity matrix based on Jaccard's similarity coefficient was generated from protein bands scored as 0 (absent) and 1 (present) and followed by a distance matrix and analysed using

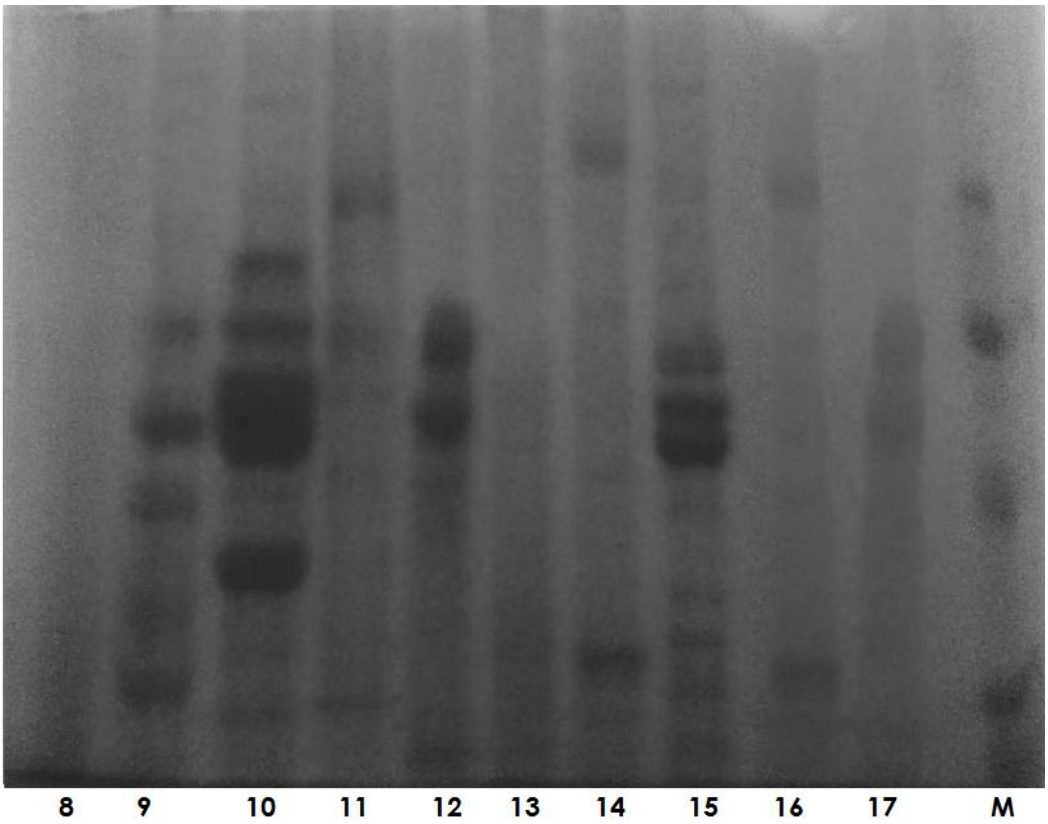

Figure 1. The banding pattern obtained by SDS-PAGE with 12% separating gel from some of the Fabaceae species of interest. Lane M = protein standard; lazne 8 = *C. pulcherrima*; lane 9 = *D. regia*; lane 10 = *D. guineense*; lane 11 = *P. pterocarpum*; lane 12= *P. vulgaris*; lane 13 = *S. semia*; lane 14 = *P. macrophylla;* lane 15 = *A. hypogeal*; lane 16 = *A. lebbeck*; lane 17 =*A. zygia.*

SPSS 15.0 for Windows. A hierarchical cluster was generated from the similarity matrix and compared with a previous similarity cluster generated from morphological data according to Polhill (1994).

RESULTS

The SDS gel electrophoresis of reproducible storage proteins for the 20 Fabaceae species resulted in bands that ranged from molecular weights of 14 to >100 kDa. Eighty nine (89) bands were detected (Figure 1).

The total seed storage proteins segregated into six distinct groups based on molecular weight of the proteins, ranging from 14 to 24, 25 to 30, 31 to 40, 41 to 62, 63to 100 and >100 kDa proteins. Each group of protein phenotypes consist of several subunits, from which specific proteins can be identified (Kottapalli et al., 2008). Higher polymorphism was recorded for proteins ranging from 31 to 62 kDa, accounting for 43.5% of the total protein bands. The least polymorphism and numbers of protein bands (9.6%) were recorded for the lower molecular weight (<24 kDa) proteins (Table 2). This indicates that the Fabaceae may predominantly express medium molecular weight genome products.

Protein profile similarity index

Analysis of the Jaccard similarity coefficient and distance for the protein profile resulted in a mean similarity of 28.3% and thus 71.7% (approximately 72%) dissimilarity. The pairwise analysis of the species (20) against the protein groups (6) recorded the least (5%) similarity and the highest (50%) similarity. Likewise, a high mean Jaccard distance of 0.668 was recorded for the taxa proteins profiled, an indication of the degree of dissimilarity between the taxa.

Cluster analysis

A dendogram of the protein polymorphism generated resulted in a phylogenetic tree construct that highlighted the clumping of the taxa into related groups based on the protein polymorphism data. Using average linkage distance, the hierarchical cluster was grouped into four distinct clusters (Figure 2) defining the phylogenetic relationship at intra-subsectional for the taxa studied.

The dendogram at 40% cluster distance revealed four clusters. The first cluster (Group A) with eight members:

Table 2. Characterization of six protein-banding patterns (groups) generated from 20 Fabaceae species.

Well lane	Taxa	Protein groups by molecular weight (kDa)						
		14-24	25-30	31-40	41-62	63-100	>100	Number of band
1	*A. auriculiformis*	Abs	Low	Abs	Low	Abs	Low	3
2	*A. pavonina*	Abs	Low	Low	Low	Abs	Low	4
3	*B. tomentosa*	Abs	Low	Abs	Abs	Abs	Abs	1
4	*B. purpurea*	Low	Low	Med	Abs	Low	Abs	5
5	*A. scorpioides*	Abs	Low	Med	Abs	Low	Abs	4
6	*C. fistula*	Abs	Abs	Med	Abs	Low	Med	5
7	*C. javanica*	Low	Abs	Low	Med	Abs	Low	5
8	*C. pulcherrima*	Low	Abs	Abs	Abs	Abs	Abs	1
9	*D. regia*	Abs	Low	Low	High	Low	Abs	6
10	*D. guineense*	Abs	Low	Med	High	Low	Low	9
11	*P. pterocarpum*	Abs	Low	Abs	Med	Med	Abs	5
12	*P. vulgaris*	Low	Abs	Abs	High	Abs	Abs	4
13	*S. semia*	Abs	Abs	Low	Abs	Abs	Abs	1
14	*P. macrophylla*	Abs	Abs	Low	Low	Low	Low	4
15	*A. hypogaea*	Low	Med	Med	High	Med	Med	11
16	*A. lebbeck*	Abs	Abs	Low	Med	Low	Abs	4
17	*A. zygia*	Abs	Abs	Low	Med	Abs	Abs	3
18	*V. unguiculata*	Low	Abs	Med	Abs	Low	Abs	4
19	*A. africana*	Abs	Abs	Med	Abs	Low	Med	5
20	*A. fruiticosa*	Low	Abs	Low	Med	Abs	Low	5

Abs = Absence; Med = medium.

Delonix regia, *Amorpha fruiticosa*, *Albizia lebbeck*, *Albizia zygia*, *Pentaclethra macrophylla*, *Phaselous vulgaris*, *Cassia javanica* and*Peltophorum pterocarpum* from the 3 sub families linking 6 tribes: *Cassieae, Caesalpinieae, Ingeae, Phaseoleae* and *Amorphheae.* The second cluster (Group B) with five members: *Caesalpinia pulcherrima*, *Senna semia*, *Bauhinia tomentosa*, *Acacia auriculiformis* and *Adenanthera pavonina* from 2 sub-families; Caesalpinoideae and Mimosoideae, linking five tribes: *Cassieae, Caesalpinieae, Circideae, Acacieae* and *Mimoseae.* The third cluster (Group C) with five members: *Cassia fistula, Afzelia africana, Bauhinia purpurea, Vigna unguiculata* and *Acacia scorpioides* from the three sub families and five tribes: *Cassieae, Circideae, Detarieae, Acacieae* and *Phaseoleae.* The forth cluster (Group D) with two members: *Dialum guineense, Arachis hypogaea* from two subfamilies and two tribes; *Cassieae* and *Aeschynomeneae*.

DISCUSSION

Genetic diversity

Protein homology (28.3%) was observed across the taxa studied, which represent the measure of similarity and thus the degree of inter-specific closeness amongst the species. 71.7% polymorphism was recorded for the seed lot. These results confirm as expected, some degree of closeness between the taxa as well as reveals the level of variations that can exist among species members of the same family (Carmona et al., 2010).

The taxa studied presented three categories of high, moderate and low molecular weight polypeptides. Recording 89 bands and a similarity of 28.3 and 71.7% protein polymorphism, the protocol applied could be useful for species/cultivar identification and protein markers generation for the family, particularly with majority of the protein homology within the 30 to 40 and 41 to 62 kDa molecular weight protein range (Figure 1 and Table 2). SDS-PAGE studies recording protein similarity and differences across different species, accessions have been highlighted by Valizadeh (2001) for grain legumes in Iran, Ghafoor and Ahmad (2005) for *Vigna mungo*, Ishtiaq et al. (2010) for Ranunculaceae and De Britto et al. (2012) for Apocynaceae. These studies including the present one, establishes that profiling seed storage protein regarded as independent of environmental fluctuations using SDS-PAGE is a reliable tool for investigating intra and inter-specific variations in plants (Hames, 1990; Carmona et al., 2010; Tchiagam et al., 2011).

The degree of polymorphism recorded shows the diversity of DNA products (proteins) within the Fabaceae and offers potentials for research, industrial and economic applications of such proteins when deciphered.

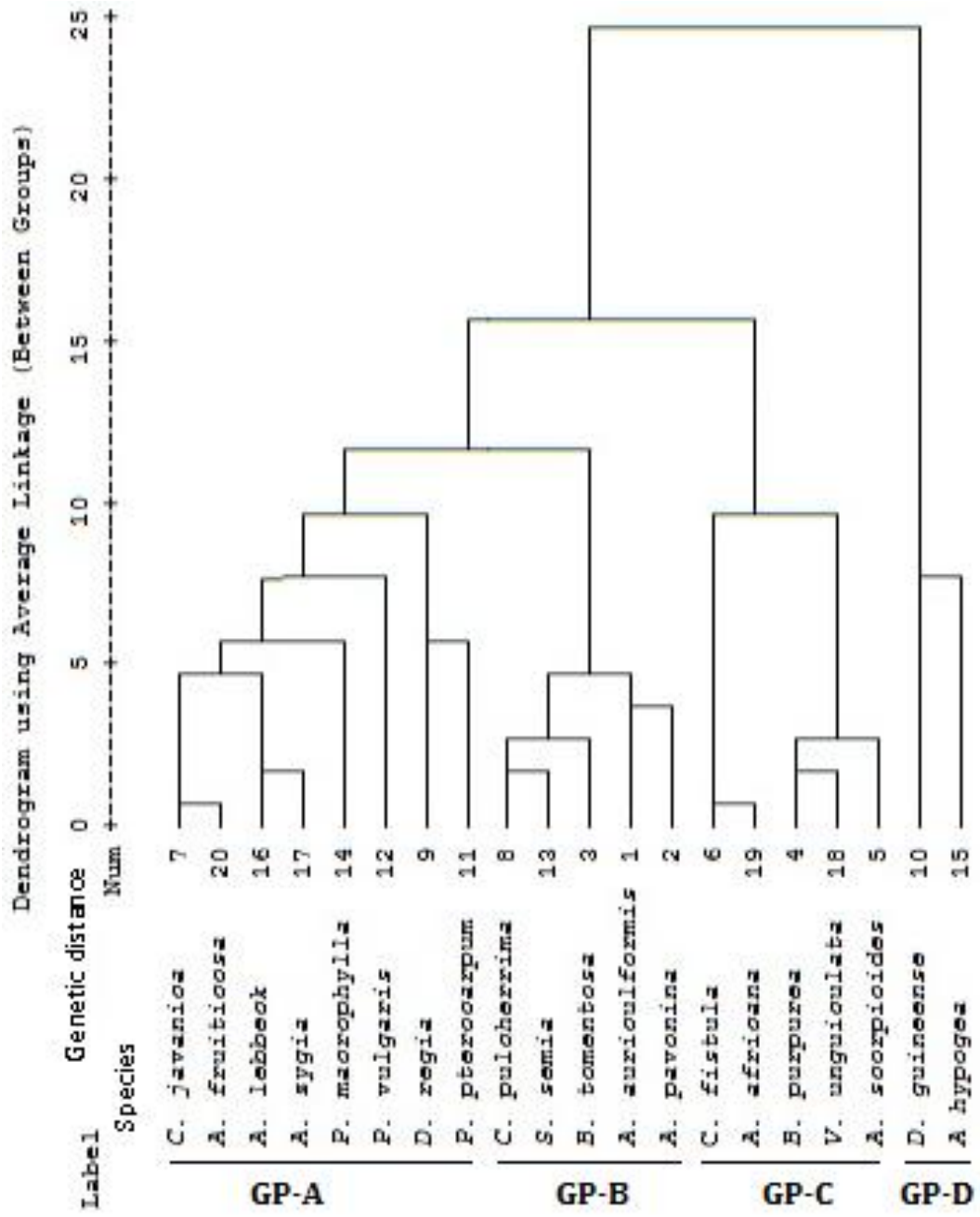

Figure 2. Dendrogram of 20 Fabaceae species based on distances produced from Jaccard's similarity coefficients of seed storage protein; with four clusters: GP-A (8 members); GP-B (5 members); GP-C (5 members); GP-D (2 members).

Taxonomic characterization

Taxonomically, the family Fabaceae have been traditionally divided into three subfamilies; Caesalpinioideae, Mimosoideae and Papilionoideae; and according to last formal classification by Polhill (1994), prior to the advent of molecular evidences, 39 tribes and 670 genera were determined for the family.

In the present study, emerging groups (A to D) in the dendogram from the Jaccard's similarity matrix revealed a divergence from the traditional taxonomic status of the Fabaceae by majority of the species studied. Comparatively, only two (*A. lebbeck* and *A. zygia*) of the 20 species maintained parallel positions in the phylogenetic trees (Figure 3) from the morphological (Polhill, 1994) and the proteomic evidences (Figure 2). In

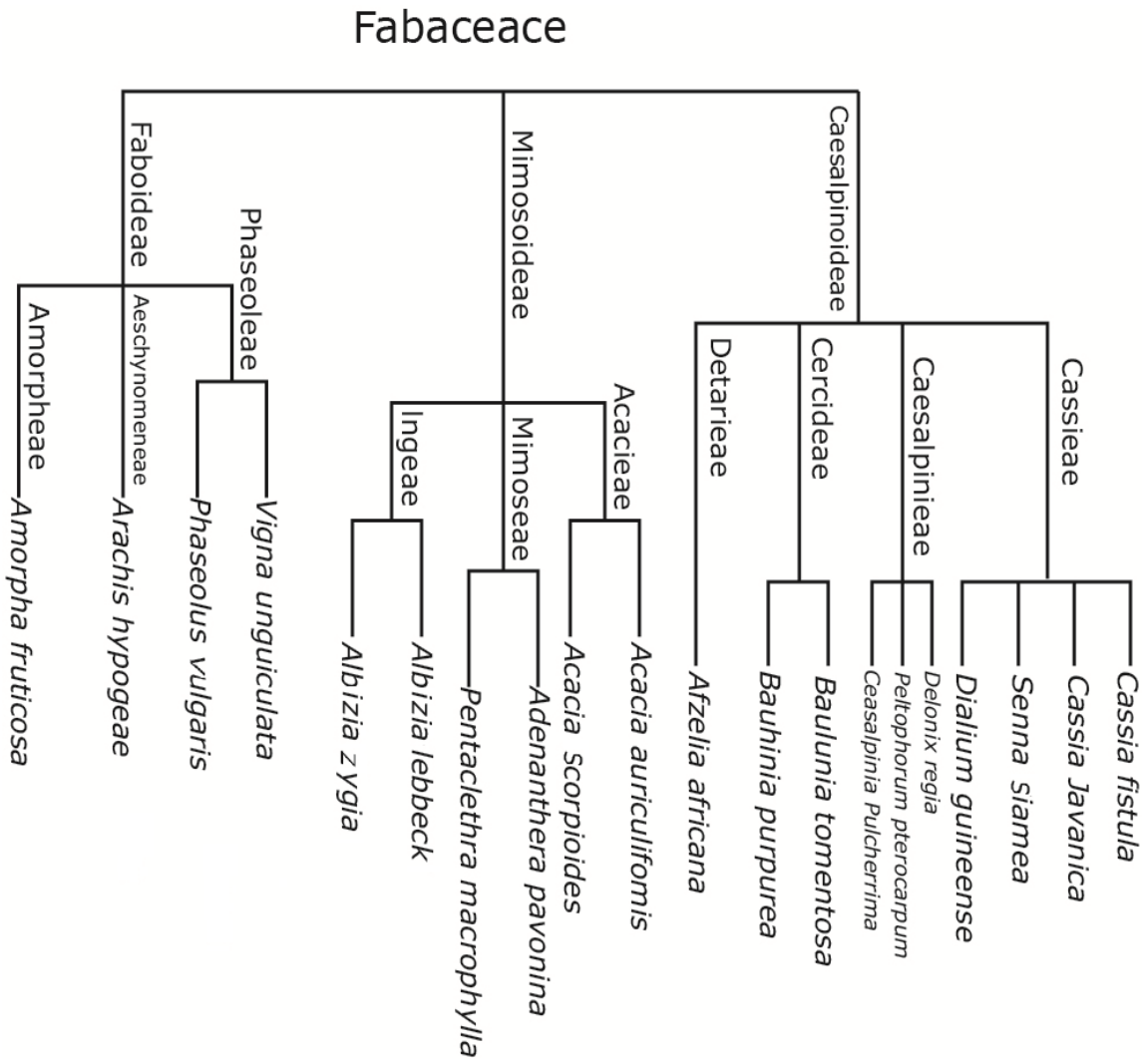

Figure 3. Traditional delimitation of the 20 Fabaceae species into three sub-families and ten tribes based on morphological characteristics according to Polhill (1994).

addition, the species *A. hypogaea* that is traditionally grouped in the sub-family Papilionoideae; increasing segregates from the group, forming an out-group (Group D) with *D. guineense*. In the seed protein profile of 11 grain legumes, Valizadeh (2001), recorded similar (single-member) out-group for *A. hypogaea.*

The species studied shared morphological characteristics that warranted their groupings, such as the tree habit for the tribe *Cassieae;* long cylindrical indehiscent pods for the genus *Cassia* and members of the tribe *Phaseoleae* are grain legumes with edible seeds. However, the present protein profile reflects a considerable re-alignment of the taxa across sub-family and tribal boundaries, which suggest that SDS-PAGE seed storage proteins profiling is a reliable tool for taxonomic studies and the evidences from such studies can lead to the identification of new groups and in some cases the delimitation into a new clade for species like *A. hypogaea*. When added to the growing pool of molecular (genomic and proteomic) evidences, these data will allow for better clarification and taxonomic review of the family, Fabaceae.

Single proteins and leaf protein concentrate applications

Comparative analysis of the study taxa profile highlighted possible legumins of importance. In the present study, *D. guineense*, clustered with *A hypogaea* and earlier profiling and identification of *A hypogaea* proteins (Kottapalli et al., 2008) recorded *Arachi*-specific as well as universal single polypeptides ranging from 18.0 to 96.6 kDa (Javaid et al., 2004). Two of such protein spots

aligned with similar spots in *D. guineense* with molecular weight of 27.8 and 32.14 kDa corresponding to the proteins glycinin, and Gly1 (Catsimpoolas et al., 1971; Kottapalli et al., 2008). These 11S-globulin proteins have been implicated in reduction of cardiovascular, pro-atherogenic factors and incidence of chronic diseases (Duranti, 2006; Fassini et al., 2011). While genomic mapping may be required to mark the precise genes coding for such proteins, quantitative proofs using techniques like protein profiling are fundamental to such molecular studies.

Fabaceae species like *Cicer arietinum, Desmodium distortum, Lens culinaris, Phaseolus calcaratus, Phaseolus mungo*, *Pisum sativum, Psophocarpus tetragonolobus, Vigna radiate* and *Vigna unguiculata*, have shown excellent amino acid contents comparable to *Medicago sativa* (alfalfa) and *Glycine max* meal (OTA, 1983; Zia-Ul-Haq et al., 2012; Khalid et al., 2012). In the present study, *V. unguiculata* clustered with four other species like *C. fistula* and *A. africana.* Similarly, *P. vulgaris* clustered with seven other species such as *P. macrophylla.* These species posses seasonal variability and other characters that qualify them as good candidates for commercial level leaf protein concentrate (LPC) production in tropical Africa, alongside other non-leguminous plants like *Telferia occidentalis* (Agbede et al., 2008*;* Adeyeye et al., 2011; Aletor and Adebayo, 2012).

Storage protein profiling of Fabaceae seeds from the study highlighted a high degree of genetic diversity within the Fabaceae. The level of inter-specific variations may warrants a review of the traditional taxonomic delimitations in the family. The comparative protein banding underscores possible important species for single proteins like glycinin as well as for LPC production.

ACKNOWLEDGEMENTS

We acknowledge Covenant University, Ota and the technical assistance of Dr. Rotimi S. O.

REFERENCES

Adeyeye EI, Omolayo FO (2011). Chemical composition and functional properties of leaf protein concentrates of *Amaranthus hybridus* and *Telfairia occidentalis.* Agric. Biol. J. N. Am. 3:499-511.

Agbede JO, Adegbenro M, Onibi GE, Oboh C, Aletor VA (2008). Nutritive evaluation of *Telfairia occidentalis* leaf protein concentrates in infant foods. Afr. J. Biotechnol. 7(15):2721-2727.

Aletor O, Adebayo AO (2012). Nutritive and Physico-Chemical Characteristics of the Leaves and Leaf Protein Concentrates from Two Edible Vegetables: A Comparative Evaluation. W. J. Dairy Food Sci. 7(1):54-58.

Carmona S, Alvarez JB, Caballero L (2010). Genetic diversity for morphological traits and seed storage proteins in Spanish rivet wheat. Biol. Planta. 54(1):69-75.

Catsimpoolas N, Kenney JA, Meyer EW, Szuhaj BF (1971). Molecular weight and amino acid composition of glycinin subunits. J. Sci. Food Agric. 22(9):448-450.

De Britto AJ, Kumar PBJR, Gracelin DHS (2012). Studies on protein profile of some medicinally important species of Apocyanaceae family using SDS-PAGE. Chem. Bio. Phy. Sci. 2(2):792-796.

Duranti M (2006). Grain legume proteins and nutraceutical properties. Fitoterapia 77:67-82.

Fassini PG, Noda RB, Ferreira ES, Silva MA, Neves VA, Demonte A (2011). Soybean glycinin improves HDL-C and suppresses the effects of rosuvastatin on hypercholesterolemic rats. Lip. Health Dis. 10:165-172.

Ghafoor A, Ahmad Z (2005). Diversity of agronomic traits and total seed protein in black gram *Vigna mungo* (L.) Hepper. Acta Biol. Cra. Ser. Bot. 47(2):69-75.

Hames BD (1990). One-dimensional polyacrylamide gel electrophoresis. In: Hames, B.D. and Rickwood, D., Eds., *Gel Electrophoresis of Proteins*, 2nd Edition, Oxford University Press, New York, 382.

Iqbal SH, Ghafoor A, Ayub N (2005). Relationship between SDS-PAGE markers and *Ascochyta* blight in chickpea. Pak. J. Bot. 37:87-96.

Ishtiaq Ch M, Mumtaz SA, Wang Y, Yiyu C, Mehmood T, Ashraf M (2010). Proteins as Biomarkers for Taxonomic Identification of Traditional Chinese Medicines (TCMs) from Subsection *Rectae* Genus *Clematis* from China. World Appl. Sci. J. (Special Issue of Biotechnology & Genetic Engineering) 8:62-70.

Javaid A, Ghafoor A, Anwar R (2004).Seed storage protein electrophoresis in groundnut for evaluating genetic diversity. Pak. J. Bot. 36:25-29.

Karihaloo JL, Kaur M, Singh S (2002). Seed protein diversity in *Solanum melongena* L., and its wild and weedy relatives. Gent. Res. Crop Eval. 49(6):533-539.

Khalid II, Elhardallou SB, Elkhalifa EA (2012). Composition and Functional Properties of Cowpea (*Vigna ungiculata* L. Walp) Flour and Protein Isolates. Am. J. Food Technol. 7:113-122. doi: 10.3923/ajft.2012.113.122.

Kottapalli KR, Payton P, Rakwal R, Agrawal GK, Shibato J, Burow M, Puppala N (2008). Proteomics analysis of mature seed of four peanut cultivars using two-dimensional gel electrophoresis reveals distinct differential expression of storage, anti-nutritional, and allergenic proteins. Plant Sci. 175:321-329.

Lewis G, Schrire B, MacKinder B, Lock M (2005). Legumes of the world. Royal Botanic Gardens, Kew, UK.

OTA (1983). Plants: The Potentials for Extracting Protein, Medicines, and Other Useful Chemicals–Workshop Proceedings (Washington, D. C.: U.S. Congress, Office of Technology Assessment, OTA-BP-F-23, September 1983).

Polhill RM (1994). Classification of the Leguminosae. pp. xxxv–xlviii in Phytochemical Dictionary of the Leguminosae (FA, Bisby, J, Buckingham, JB, Harborne). Chapman and Hall, New York, NY.

Polhill RM, Raven PH (1981). Advances in legume systematics, parts 1 and 2. Royal Botanic Gardens, Kew, UK.

Rundel PW (1989). Ecological success in relation to plant form and function in the woody legumes. **In**: CH, Stirton, JL, Zarucchi,). Advances in legume biology, Monog. Syst. Bot. Missou. Bot. Gard. 29:377-398.

Sprent JI, McKey D (1994). Advances in legume systematics, Part 5, The nitrogen factor. Royal Botanic Gardens, Kew, UK.

Sprent JI (2001). Nodulation in legumes. Royal Botanic Gardens, Kew, UK.

Tchiagam JN, Bell JM, Nassourou AM, Njintang NY, Youmbi E (2011). Genetic analysis of seed proteins contents in cowpea (*Vigna unguiculata* L. Walp.). Afr. J. Biotechnol. 10(16):3077-3086.

Valizadeh M (2001). Seed Storage Protein Profile of Grain Legumes Grown in Iran, Using SDS-PAGE. J. Agric. Sci. Technol. 3:29.37-292.

Yousaf Z, Masood S, Shinwari ZK, Khan MA, Rabani A (2008). Evaluation of taxonomic status of medicinal Species of the genus *Hyoscyamous, Withania, Atropa* and *Datura* based on poly acrylamide gel electrophoresis. Pak. J. Bot. 40(6):229.39-2297.

Zia-ul-haq M, Shahid SA, Ahmad S, Qayum M, Rasool N (2012). Mineral contents and antioxidant potential of selected legumes of Pakistan. J. Med. Plant Res. 6(32):4735-4740. doi: 10.5897/JMPR12.255.