Seed Res. 53 (1): 31-35, June, 2025
DOI: https://doi.org/10.56093/SR.v53i1.5

Protein Composition of Cotton Seeds in Relation to Agro-Climatic Variation

V SANTHY*, POOJA VERMA, DEBASHIS PAUL, PRATIK DHAKADE, ANJALI CHOUDHARY AND VN WAGHMARE

ICAR-Central Institute for Cotton Research, Nagpur, Maharashtra-441108, India
*santhy100@gmail.com

(Received April 2025; Revised May 2025; Accepted May 2025)

ABSTRACT: Cotton, globally grown as a fibre crop, is also popular as oil and protein source, which is obtained from cotton seed. An investigation on seed protein content was done in popular cotton cultivars released by ICAR – Central Institute for Cotton Research. Twelve cultivars belonging to three cultivated species (*G. hirsutum*, *G. arboreum* and *G. barbadense*), seeds of which were produced in different locations (Central, North, and South cotton growing regions), were subjected to determination of tris-soluble and seed storage protein fractions. The laboratory investigation was conducted at one location after collecting seeds of same genotypes grown in other locations in the same season for comparing their protein content. Concentration of tris-soluble protein and sequentially extracted storage protein fractions were determined by Bradford assay. Significant variation was observed for tris-soluble and storage protein fractions among genotypes in each location and showed varied relations with seed index in *G. arboreum* and *G. barbadense*. Among the storage fractions, globulin fraction was highest followed by albumins in the studied cultivars. Though the values of each protein component differed widely in different location, the genotypes maintained their rank for both tris-soluble protein content as well as seed storage fraction content irrespective of the location where the seeds were produced.

Keywords: Cotton, seed protein, tris-soluble protein, storage fraction, globulin, albumin

INTRODUCTION

Cotton crop is a popular source of natural fiber which develops from the outer epidermis of cotton seed and constitutes 1/3rd of the total harvest. The significant oil and protein content (21% and 23% by weight respectively) in the seed makes cotton, the fifth largest oil crop in the world and the second most important potential source of plant proteins [1]. Genetic improvement of cotton seed storage protein with successful removal of the gossypol sequestered in the pigment glands is necessary if it is to be used as a source of protein for non-ruminant animals or humans.

Because of continual efforts of breeding cotton for increase in fiber quality and yield, several seed properties were unattended and not improved. The immense potential of cotton breeding for oil and high protein content was discussed [2] and association mapping for their loci which can assist in future breeding was performed [3]. The pivotal role of seed storage proteins as well as several enzymatic proteins in enhancing seed quality concerning dormancy, vigor, and viability, is also well documented [4]. Increase in total soluble seed protein as a result of seed polymer coating with nanoparticles has been reported earlier [5]. A high heritability with additive gene action has been identified for cotton seed protein content and direct selection for its improvement was suggested [6]. Earlier, characterization of genotypes for seed protein in cotton was performed for genetic diversity [7]. Seed protein content in crops including cotton is reported to be highly variable depending on the growth conditions, nutrient supply and genotype [8]. The information on content of seed protein present in presently available Indian cultivars is meagre and hence the study was performed with the following objectives (1) To determine the variability for tris-soluble seed protein and seed storage fraction contents among released cotton cultivars belonging to different species and (2) To evaluate the change in content of tris-soluble and storage protein fractions in seeds on same genotypes produced under different climatic and management condition.

MATERIAL AND METHODS

Twelve cultivars (12) belonging to three cultivated species of cotton suitable for growing at north, central and south zone were utilized for the study as given in Table 1. Fresh seeds of those cultivars suitable for North zone as well

Table 1. Details of the genotypes used in the experimentation and zone of seed production

S. No.	Name of the genotype	Species	Zone/es in which grown to produce seed
1.	CISA 310	*G. arboreum*	North and Central
2.	CISA 614	*G. arboreum*	North and Central
3.	CSH 3075	*G. hirsutum*	North and Central
4.	CSH 3129	*G. hirsutum*	North and Central
5.	MCU 5 VT	*G. hirsutum*	South and Central
6.	Surabhi	*G. hirsutum*	South and Central
7.	Suraj	*G. hirsutum*	South and Central
8.	Supriya	*G. hirsutum*	South and Central
9.	Sumangala	*G. hirsutum*	South and Central
10.	Anjali	*G. hirsutum*	South and Central
11.	LRA 5166	*G. hirsutum*	South and Central
12.	Suvin	*G. barbadense*	South and Central

as Central zone were sown at ICAR- Central Institute for Cotton Research, Regional Station, Sirsa and ICAR-Central Institute for Cotton Research, Nagpur, respectively. Those suitable for South and Central zone were grown at ICAR- Central Institute for Cotton Research, Regional Station, Coimbatore and ICAR-Central Institute for Cotton Research, Nagpur, respectively during the same year 2021. At the end of the season, the seeds from the cotton harvested at each location were collected at ICAR- Central Institute for Cotton Research, Nagpur to compare the seed protein and storage protein fraction content for same cultivars grown under different locations. The seeds received from each location were acid delinted and hundred seed weight was recorded. The seed meal obtained after seed coat removal was crushed and subjected to defatting. The crushed, defatted seed meal of different varieties was our study material.

Estimation of Seed Protein

The seed meal crushed with 0.1M tris-HCl solution with pinch of PVP was used for extraction of tris-soluble seed protein content as per the protocol [9]. The soluble protein content of samples resulting from extraction was estimated (μg/ml) using the Bradford assay with bovine serum albumin (BSA) as the protein standard [10]. The concentration was later calculated in the units of mg/g seed. Each fraction of total seed storage proteins such as albumins (using de-ionized water), globulins (using 1M NaCl), prolamines (using 80% ethanol) and glutelins (using 0.01 M NaOH) were extracted by the method [11]. The extracted fraction was again subjected to Bradford assay to determine the concentration. The observations were made on three different seed sample of each variety collected from each location.

Statistical analysis

The experiments were repeated with triplicate samples and statistical computations were performed. The data on seed weight and protein content were checked for normality and analyzed using SAS statistical Software (SAS Institute Inc., Cary, NC). Mean differences of various parameters were tested using appropriate F-test. Means of the treatments were separated based on the Tukey's Honestly Significant Difference test with α=0.05.

RESULT AND DISCUSSION

The variation for seed protein and storage content by Bradford method was investigated in twelve cultivars belonging to three cultivated species which were simultaneously grown under different agro-climatic condition. The influence of environment including seed quality, seed priming, weather in which seed was produced etc. on seed protein content has been reported earlier [12]. Significant variation was detected among genotypes for total seed protein content as well as the storage fraction content such as albumins, globulins, prolamins and glutelin for genotypes in each location (Table 2 and Table 4). The seed index of all genotypes measured as weight of hundred acid delinted seeds also varied significantly. Variation in seed protein composition within the genus of Gossypium has been reported [13]. Bradford assay was used to study genotypic variation for seed protein and four storage fraction content in seeds of diversly grown *G. arboreum* and *G. hirsutum* cultivars which was used later for determining genetic diversity [14].

Tris-soluble and storage protein fractions for seeds produced at North and Central zone

The results on tris soluble seed protein content as well as the seed storage fractions for each of the four

Table 2. Interaction of cotton genotypes and location of production for seed protein content (Central zone and North zone)

Genotype	Tris Protein (mg/g)	Globulin (mg/g)	Albumin (mg/g)	Prolamin (mg/g)	Glutelin (mg/g)	Seed Index (g)
Central zone (Nagpur)						
CISA-310 (A)	23.8^{b}	50.9^{ab}	20.9^{a}	7.8^{ab}	3.6^{bc}	4.72^{de}
CISA-614 (A)	29.1^{ab}	54.6^{ab}	24.7^{a}	7.6^{ab}	8.1^{a}	4.67^{e}
CSH-3075 (H)	30.7^{ab}	36.1^{ab}	21.2^{b}	8.8^{ab}	1.6^{c}	6.45^{c}
CSH-3129 (H)	34.4^{a}	40.6^{ab}	19.8^{b}	10.6^{a}	4.6^{bc}	6.90^{bc}
North zone (Sirsa)						
CISA-310 (A)	30.5^{ab}	63.3^{a}	26.2^{a}	7.5^{ab}	3.3^{c}	5.87^{cd}
CISA-614 (A)	32.1^{a}	61.9^{a}	27.2^{a}	6.8^{b}	6.8^{ab}	6.03^{c}
CSH-3075 (H)	30.5^{ab}	25.1^{b}	22.6^{a}	9.4^{ab}	2.6^{c}	8.00^{ab}
CSH-3129 (H)	33.1^{a}	50.2^{ab}	23.4^{a}	2.1^{c}	3.4^{bc}	8.35^{a}

* Means are compared by Tukey's HSD test. Levels not connected by same letters within a column are statistically different ($p<0.05$); A: *arboreum* genotype, H: *hirsutum* genotype

Table 3. Total soluble and storage protein content with respect to location of Production (Central and North)

Location	Tris Protein (mg/g)	Globulin (mg/g)	Albumin (mg/g)	Prolamin (mg/g)	Glutelin (mg/g)	Seed Index (g)
Central (Nagpur)	29.5^{a}	45.5^{a}	21.7^{b}	8.7^{a}	4.5^{a}	5.68^{b}
North (Sirsa)	31.6^{a}	50.2^{a}	24.9^{a}	6.5^{b}	4.0^{a}	7.06^{a}

*Means are compared by Tukey's HSD test. Levels not connected by same letters within a column are statistically different ($p<0.05$)

genotypes grown and seed produced in north and central zone is provided in Table 2. The seed protein content was highest for the *hirsutum* cultivar, CSH-3129 followed by CSH 3075 respectively at both Sirsa of north zone as well as Nagpur of Central zone. Remaining two *G. arboreum* cultivars showed lower values for seed protein content and also seed index compared to the *G. hirsutum* cultivars at both the locations. The finding is totally in agreement with report that diploid genomes generally have low values for seed index as well as seed protein [15].

Among the seed storage protein fraction, water soluble albumins and salt soluble globulins were higher in all the cultivars irrespective of species indicating prominence of these fractions in cotton seed. The classification of cotton seed proteins as water-soluble albumins, alkali-soluble glutelins, salt-soluble globulins and alcohol-soluble prolamins was proposed [16]. The prolamin and glutelin fractions were significantly low in all the cultivars. The higher presence of globulins and albumins were earlier reported for cotton seed [14]. Seeds produced within same location also showed significant difference between cultivars for both albumins and globulins as mentioned in the beginning of the results. Both the major storage fractions such as albumin and globulin were highest in two *arboreum* cultivars viz. CISA 310 and CISA 614 compared to *hirsutum* at both the locations *viz.* Sirsa and Nagpur inspite of their lower seed index. Among the two major fractions, globulin storage fraction showed higher values than albumin fraction in seeds of all cultivars produced at both the locations and is in total agreement with earlier studies in cotton seed [14, 17].

The comparison for traits between both locations revealed significant difference for values of tris-soluble protein content, albumin storage fraction content and seed index for seeds produced at north and central zone (Table 3). The values for seed protein content across locations are bound to differ under different field management and rate of nitrogen application [18]. However, non-significant location x genotype interaction determined [19] suggest that genotypes maintain their rank for seed protein content across environments. Thus, selection for high seed protein content cultivars might be effective for broad range of environments.

Tris-soluble and storage protein fractions for seeds produced at South and Central zone

The results on tris soluble seed protein content as well as the seed storage fractions for each of the eight genotypes grown and seed produced in South and

Table 4. Seed Protein content and seed Index for cotton genotypes produced (Central and South Zone)

Genotype	Total soluble Protein (mg/g)	Globulin (mg/g)	Albumin (mg/g)	Prolamin (mg/g)	Glutelin (mg/g)	Seed Index (g)
Central Zone (Nagpur)						
MCU-5VT (H)	30.8^{abc}	57.2^{abc}	25.8^{a}	5.2^{cde}	1.8^{fg}	7.97^{ef}
Surabhi (H)	28.9^{abc}	35.0^{ab}	24.5^{a}	6.2^{bcd}	1.4^{fg}	8.00^{ef}
Suraj (H)	30.0^{abc}	38.6^{abcd}	15.8^{bc}	7.7^{bcd}	1.3^{fg}	8.61^{cde}
Supriya (H)	26.1^{bcd}	32.2^{cd}	17.4^{abc}	7.3^{bcd}	1.5^{fg}	6.45^{g}
Suvin (B)	28.0^{abcd}	17.6^{d}	14.1^{c}	7.7^{bcd}	4.5^{cde}	9.55^{abc}
Sumangala(H)	29.6^{abc}	23.9^{d}	22.2^{abc}	5.3^{cde}	1.1^{g}	7.16^{fg}
Anjali (H)	30.8^{abc}	31.1^{cd}	25.7^{a}	3.9^{de}	3.4^{def}	8.30^{def}
LRA-5166 (H)	33.2^{ab}	25.2^{d}	24.0^{ab}	1.2^{a}	7.8^{b}	7.17^{fg}
South Zone (Coimbatore)						
MCU-5VT (H)	34.3^{a}	66.8^{a}	28.3^{a}	7.6^{bcd}	1.6^{fg}	9.37^{bcd}
Surabhi (H)	30.1^{abc}	26.9^{d}	21.5^{ab}	10.3^{ab}	6.0^{bc}	8.07^{ef}
Suraj (H)	27.9^{abcd}	22.7^{d}	21.9^{ab}	12.4^{a}	5.3^{cd}	10.10^{ab}
Supriya (H)	24.5^{cde}	26.7^{d}	21.4^{ab}	8.6^{abc}	2.1^{fg}	9.48^{abcd}
Suvin (B)	21.1^{de}	16.7^{d}	11.9^{c}	1.2^{e}	1.9^{fg}	10.72^{a}
Sumangala(H)	27.7^{c}	16.2^{d}	14.4^{bc}	6.4^{bcd}	2.6^{efg}	9.31^{bcd}
Anjali (H)	31.8^{ab}	35.2^{bcd}	24.5^{a}	4.7^{cde}	1.6^{fg}	8.87^{bcde}
LRA-5166 (H)	31.5^{abc}	25.3^{d}	26.3^{a}	10.1^{ab}	11.1^{a}	8.78^{cde}

* Means are compared by Tukey's HSD test. Levels not connected by same letters within a column are statistically different ($p<0.05$); H: *hirsutum* genotype; B*: barbadense* genotype

Table 5. Total soluble and storage protein content with respect to location of Production **(**Central and South)

Location	Total solubleProtein	Albumin	Globulin	Prolamin	Glutelin	Seed Index
Nagpur (Central)	29.7^{a}	24.1^{a}	$36.2.^{a}$	6.9^{a}	2.8^{b}	7.90^{b}
Coimbatore (South)	27.4^{b}	21.2^{a}	29.6^{b}	7.7^{a}	4.0^{a}	9.34^{a}

* Means are compared by Tukey's HSD test. Levels not connected by same letters within a column are statistically different ($p<0.05$)

Central zone is provided in Table 4. The seed index varied significantly across genotypes with in same location when produced in south as well as central zone. The *G. barbadense* genotype, Suvin showed the highest value for seed index followed by *G. hirsutum* varieties, Suraj, and Sumangala respectively. The variability for tris-soluble seed protein content was significant for seeds produced in the south region at Coimbatore. Despite having a high seed index, the *barbadense* variety, Suvin showed lowest values for tris-soluble protein content in seeds produced at both central and south locations (Table 5). This is unlike what was observed in the comparison mentioned earlier between north and central. The finding is supported by the earlier report that there is a varying relationship between seed protein content and seed index in different species [15]. However, with respect to the ranking for seed protein content, seeds of same genotypes viz. MCU 5 VT, Anjali and LRA 5166 showed higher values when produced in both central and south locations. For the seed storage fractions, the albumin and globulin fraction were highest as observed earlier and prolamin, glutelin fractions were very low. In these seeds, not much of difference could be observed between albumin and globulin storage fraction content. It is pertinent to mention that, though the values are different, genotypes showing highest albumin and globulin fraction content was same in both the locations of central and south as observed in the comparison between central and north produced seeds.

CONCLUSION

The genotypes under investigation revealed significant variability for both tris-soluble seed protein as well as seed storage protein fraction on seeds produced at all locations. The difference in values for these traits in same genotypes across locations proved that seed production environment and post-harvest storage environment can influence the seed protein content and its storage fractions in cotton. However, the genotypes maintained their ranking of highest or lowest values, with consistency, for both tris-soluble protein content as well as seed storage fraction content irrespective of the location where the seeds were produced. The cultivar differences and ranking for seed protein content presented in the above study would allow a cotton researcher to select cultivars based upon the trait and/or use in future breeding programme for improved cotton seed composition.

REFERENCES

1. QING L, D LLEWELLYN, SP SINGH, AND AG GREEN (2012). Cotton seed development: opportunities to add value to a byproduct of fibre production, *Cotton Flowering and Fruiting*, chapter **9**, 134-161.
2. MUTHUSAMY A AND N JAYABALAN (2013). Variations in seed protein content of cotton (*Gossypium hirsutum* L.) mutant lines by in vivo and in vitro mutagenesis. *Journal of Environmental Biology,* **34**: 11-16.
3. GUIZHEN L, H MEI, S WANG, X LI, X ZHU, AND T ZHANG (2015). Association mapping of seed oil and protein contents in upland cotton. *Euphytica*, **205**: 637–645.
4. BHAT RS, RV HUNJE AND SS BHARATH (2024). Fresh seed dormancy in groundnut: Mechanisms, Factors and Implications for Crop Improvement – a Review *Seed Research,* **52**: (1) 1-8.
5. CHOUGALA LD, PARASHIVMURTHY, MS HARISH AND R SIDDARAJU (2024). Seed Polymer Coating with ZnO, SiO2 and TiO2 Nanoparticles: An Innovative Seed Quality Enhancement Technique in Sweet Pepper. *Seed Research*, ***52***(2): 124-129.
6. POOJA, SR PUNDIR, O SANGWAN, M RATHI, S KUMAR, AND DIGAMBER (2020). Study of genetic parameters and clustering on the basis of morpho-bio traits of cotton genotypes *Journal of Cotton Research and Development,* **34**(2): 183-191.
7. BHATTI KH, Y BOLEK, H TEKEREK, AND A BARDAK (2016). Screening of Cotton Genotypes for Protein Content, Oil and Fatty Acid Composition. *Journal of Agriculture Science*, **8**(5): 107-121.
8. BELLALOUI N, RB TURLEY, AND SR STETINA (2021). Cottonseed Protein, Oil, and Minerals in Cotton (*Gossypium hirsutum* L.) Lines Differing in Curly Leaf Morphology. *Plants*, **10**: 525. https://doi.org/10.3390/ plants10030525.
9. RANJAN S, R MATCHA, B MADHURI, AND NR BABU (2012). Comparative evaluation of protein extraction methods from few leguminous seeds. *International Journal of Advanced Biotechnology and Research,* **3(2)** pp 558-563.
10. BRADFORD M (1976). A rapid and sensitive method for the quantitation of microgram quantities.
11. PANDA AM, R MISHRA, A KUMAR, A KERKETTA, AND N SONI (2018). Seed storage protein evaluation of few rice varieties used by tribal people of Chhattisgarh. *Journal of Rice Research,* **11**(2):31-35.
12. VIDYA CJ AND N NETHRA (2024). Seed Priming: Tool towards Sustainable Agriculture. *Seed Research*, **52(1)**: 30-40.
13. WU M, W PEI, T WEDEGAERTNER, J ZHANG, AND J YU (2022). Genetics, Breeding and Genetic Engineering to Improve Cottonseed Oil and Protein: A Review. *Front Plant Sci*. **10(13)**: 864850.
14. SINGH AND KAUR (2019). Comparative studies on seed protein characteristics in eight lines of two Gossypium species *Journal of Cotton Research,* **2**: 6 https://doi.org/10.1186/ s42397-019-0024-3
15. HINZE L, P HORN, N KOTHARI, J DEVER, J FRELICHOWSKI, K CHAPMAN, AND R PERCY (2015). Nondestructive Measurements of Cottonseed Nutritional Trait Diversity in the U.S. National Cotton Germplasm Collection. *Crop Science*, **55**, pp 770-782.
16. SAMMOUR RH, MN EL-SHOURBAGY, AM ABO-SHADY, AND AM ABASARY (1995). Proteins of cottonseed (*Gossypium barbadense*) extraction and characterization by electrophoresis. *Qatar Univ. Sci. J.*, **15(1)**: 77- 82.
17. FLATIE D, T TESFAYE, KM BABU, M GIBRIL, AND F KONG (2023). Extraction and characterization of protein from cottonseed meal. *Environ Waste Management Recycling*, **6**(2): 140.
18. PARKASH V, JL SNIDER, A BRUCE, A ERMANIS, G VIRK, N KAUR, G COLLINS, AND KD CHAPMAN (2023). Effects of cultivar and nitrogen application rate on lint, seed, oil, and protein yields of field-grown cotton. *Crop Science,* **63**(3): 1541-1554.
19. MERT M, Y AKISCAN, AND O GENCER (2004). Inheritance of Oil and Protein Content in Some Cotton Generations. *Asian Journal of Plant Sciences*, **3**: 174-176.