Augmenting protein biofortification in maize using teosinte (Zea mays ssp. parviglumis)

S Varalakshmi
Govind Ballabh Pant University of Agriculture and Technology
Narendra Kumar Singh (narendraksingh2@gmail.com)
Govind Ballabh Pant University of Agriculture and Technology
Navneet Pareek
Govind Ballabh Pant University of Agriculture and Technology
V Senthilkumar
Govind Ballabh Pant University of Agriculture and Technology

Research Article

Keywords: Maize, Teosinte, Protein, Biofortification, Variability

Posted Date: July 11th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1753627/v2

Abstract

Maize underwent domestication from the wild species Teosinte (*Zea mays ssp. parviglumis*) in the Balsas River valley of Mexico. While domestication and selective breeding of maize lost major alleles for kernel size and composition, Teosintes have a huge repository of novel and diverse alleles for protein content; have doubled protein content and higher levels of methionine-rich zeins than cultivated maize. To address the issue of protein in maize, we developed and investigated $BC_1F_{2:3}$ population derived from crossing between maize inbred line CML-451 and teosinte (*Zea mays ssp. parviglumis*). The data on ear traits and protein content were recorded on 126 $BC_1F_{2:3}$ individual cobs and parental lines. The protein content of among the $BC_1F_{2:3}$ lines had highly significant variation which varied from 9.53–16.49% for MP51 and MP6 lines, respectively. The results stipulated that protein content of teosinte introgressed maize lines is improved by 41.42% over CML-451. The introgressed population also exhibited significant variation for all ear traits investigated. Principal component analysis and cluster analysis have applied for estimation of genetic diversity among introgressed lines. Proteins content, ear length (cm), ear width (cm), kernel rows/ear and test weight (g) are all significant contributors to maximum diversity among teosinte introgressed maize lines. Cluster analysis highlighted the diversity among lines based on protein content and test weight and grouped them into six clusters. Cluster II lines were characterized by higher protein content and test weight values and were considered for wider use in breeding programmes. The results of the investigation highlight the substantial increase in the protein content of the teosinte derived maize lines and therefore indicates critical role of Zea *mays ssp. parviglumis* in protein biofortification of maize kernels.

1. Introduction

Maize (*Zea mays L. subspecies mays*) has played a major role in shaping the food and poultry industry (Dei et al.2017) and important source of dietary energy and nutrients for African countries and consumed as staple food in various forms (Ekpa et al. 2019). However, intrinsic nutritional deficits of maize regarding protein quality and quantity, effect of anti-nutritional factors and processing nutritional loss causes malnutrition (Ranum et al.2014). From decades, malnutrition remains a widespread problem despite breakthroughs in agriculture research and innovations (Neeraja et al. 2017). Upon consuming maize as dietary source, major malnutrition form faced by people is protein-energy malnutrition (PEM) which mainly influence children stunted growth and underweight which leads to serious diseases like Kwashiorkor and Marasmus (Nyakurwa et al. 2017; Zimmerman et al. 2018 and Kiprop 2020). Apart from food supplementation, maize biofortification for improved both quantity and quality of protein is considered to be a major cost-effective agenda in maize breeders' premises across the continents to address malnutrition (Bouis et al. 2017; Hossain et al. 2019).

Domestication process created substantial morphological differences among cultivated crops from their wild ancestors (Hufford et al. 2012) which changed civilizations to thrive on crops, cultivation practices, and also crop evolution (Flint-Garcia 2013). Likewise, maize has undergone a domestication process during the last ~ 10000 years in the Balsas River Valley of southern Mexico from its wild progenitor Teosinte (*Zea mays ssp. parviglumis*) (Matsuoka et al. 2002). However, domesticated maize exhibited significant morphological changes in plant architecture, inflorescence, and ear and grain characters (Doebley et al. 1995) and also possess remarkably low variation in kernel size and composition compared to wild relatives (Flint-Garcia et al.2009). Teosintes have doubled protein content and have higher levels of methionine-rich -zeins than cultivated maize (Swarup et al., 1995). In the last decade, geneticists and maize breeders have made remarkable efforts to enhance the protein quality and quantity of maize, one such experiment was an intensive selection for magnifying kernel oil and protein concentration at the University of Illinois initiated in 1896 by C.G. Hopkins and developed Illinois high protein line having 26.2% protein content (Hopkins 1899) but this method required long duration and laborious. Another contribution to enhancing protein quality is the unexpected finding of the *opaque-2* mutation and its modified version, Quality Protein Maize (QPM) (Vasal et al. 1980). The mechanism behind this recessive mutant is altered expression of endosperm-specific transcription factor belonging to the bZIP family which is essential for transactivation of zein genes due to which α-zeins expression is reduced and non-zein proteins production was enhanced with increased lysine content (Schmidt et al. 1992; Neto et al. 1995). Even though the quality of protein improved but the percentage of overall protein content remains unchanged (Holding and Larkins, 2009; Wu et al.2014). Still there is gap in science for enhancing total protein content of maize; hence a maize breeder has to search for other alternatives to speed up the breeding programme for elevating both the quantity and quality parameters of maize shortly by integrating diverse alleles. Since it was reported that maize wild relatives have wide variability for protein content (~ 30%) and also improved zein profiles (Wu et al. 2014), re-domestication of wild alleles may prove to be potential alternative for diversification as well as revival of the lost alleles for protein content and hasten crop improvement(Flint-Garcia et al. 2009). Karn et al. (2017) also highlighted the potentiality of *Zea mays ssp. parviglumis* for kernel composition traits using near isogenic lines derived from *parviglumis.*

Previous research has shown that teosinte can be successfully used as a pollen parent for developing teosinte introgressed lines (Singh et al. 2017) and also contributed significantly to the elevation of protein content of maize background lines (Perini and Magoja 1988). Introgression from teosintes into a maize background adds a significant impact on improving inbred lines for protein content and nutritional value. The *Zea diploperennis* introgressed population also exhibited higher mean protein content than the parental lines (Perini et al.1991). Improvement of protein content without a yield penalty has been a major research topic in maize breeding. Ear length, ear width, number of kernel rows per ear, number of kernels per ear, and test weight of grains are the traits that influence grain yield are solely responsible (Sahoo et al. 2021). Thus, based on the little information available, we hypothesized that wild progenitor of maize *Zea mays ssp. parviglumis*, may be a potential source for domestication of genes and alleles for protein biofortification in maize. As such an elite inbred line CML-451 was investigated in cross combination with *Zea mays ssp. parviglumis* for protein, ear and kernel traits. The characterization of introgressed population for various ear morphological characters gives an idea regarding how alien chromatin introgression changes the recipient genome which ultimately visible on phenological characters (Wang et al. 2012; Gonzalez et al. 2019). Multivariate technique principal component analysis (PCA) is widely used to reduce the dimensionality of data and important tool for diversity analysis based on variance of traits (Mohammadi et al.2003). Clustering of genotypes also helps breeders for selection of subsets of population for specific breeding purposes (Rincon et al 1996). At the same time trait association studies also important which aid in selection of superior lines having combination of desirable traits together via connotative way by selecting secondary traits for improvement of complex traits (Meseka et al. 2013). The present investigation therefore planned to develop and analyse teosinte (*Zea mays ssp. parviglumis)*

derived maize lines for protein content, ear and kernel traits. The investigation also seeks to validate the prospects of teosinte in protein biofortification of maize.

2. Materials And Methods

2.1. Plant and seed materials

The Plant and seed materials for present investigation were generated at N.E. Borlaug Crop Research Centre, G.B. Pant University of agriculture and technology, Pantnagar, Uttarakhand. To create teosinte introgressed lines; Teosinte (*Zea mays ssp. parviglumis*) was used as pollen parent which contains 19.67% and an elite maize inbred line CML-451 was used as a seed parent having ~ 9% protein content. The maize line was crossed with teosinte to create F1s, with one backcross with said maize inbred line as a recurrent parent BC_1F_1s was generated. Eventually two generations of selfing were carried out to produce 126 $BC_1F_{2:3}$ population. The individual cobs were harvested, dried and kernels were used for flour preparation for further crude protein analysis.

2.2. Evaluation of ear traits

The teosinte introgressed 126 BC_1F_2 lines were sown in *rainy* season of 2021. Each line was sown 3m row with 60×20 cm planting distance at N.E. Borlaug Crop Research Centre, G.B. Pant University of Agriculture and Technology, Pantnagar, Uttarakhand (India) for evaluation of ear-related traits. These plants were self pollinated to obtain 126 $BC_1F_{2:3}$ individual lines. The data on ear length, ear width, number of kernel rows per ear, number of kernels per row and test weight of grains were recorded on $BC_1F_{2:3}$ individual cobs.

2.3. Protein analysis

The analysis of total protein content of maize kernels was done by using standard Kjeldahl method developed in 1883 by Johann Kjeldahl on automatic system Kjeldahl apparatus (Make: KDI040) placed in Central Analytical Laboratory, Department of Soil Science, G.B. Pant University of Agriculture and Technology, Pantnagar, Uttarakhand. Grain sample of maize is digested with a strong acid so that it releases nitrogen which can be determined by a suitable titration technique. Total protein content of parents and 126 $BC_1F_{2:3}$ individual lines were analyzed and amount of protein present is then calculated from the nitrogen concentration of the grains. The 126 individual lines and parental lines kernels were carefully shelled from their cobs, dried and kernels of each line were ground to fine power using motorized grinder machine and finally stored for further chemical analysis for nitrogen. With three replications, the dried flour is weighed at around 200mg and placed to digestion flasks. The catalyst is delivered in a flask that is included in the digesting mixture and consists of a 10:1 ratio of K_2SO_4 and $CuSO_4$. Then 5ml of concentrated H_2SO_4 is added to the mixture to aid in the conversion of nitrogen to ammonium sulphate. After that, the flasks were placed in a digestive system and heated at 420°C for 90 minutes, with an additional 20 minutes for clearing; after that, the heating was turned off to allow the contents of the flasks to cool before being utilized for distillation. Digestion converts any nitrogen in the grains into ammonia, and other organic matter to CO_2 and H_2O. Ammonia gas is not liberated in an acid solution because the ammonia is in the form of the ammonium ion (NH_4^+) which binds to the sulfate ion (SO_4^{2-}) and thus remains in solution. The alkaline distillation was carried out in an automatic KDI040 Kjeldahl distillation system with 40 percent NaOH, and the flask was carefully heated for 6 minutes to liberate ammonia from the digested mixture. After the digestion has been completed the digestion flask is connected to a *receiving flask* by a tube. The solution in the digestion flask is then made alkaline by addition of sodium hydroxide, which converts the ammonium sulfate into ammonia gas. The ammonia gas that is formed is liberated from the solution and moves out of the digestion flask and trapped into the receiving flask containing 4 percent excess boric acid with a mixed indictor, ensuring complete ammonia release in the receiving flask. The low pH of the solution in the receiving flask converts the ammonia gas into the ammonium ion, and simultaneously converts the boric acid to the borate ion. The nitrogen content is then estimated by titration of the ammonium borate formed with standard 0.1N H_2SO_4 using a suitable indicator to determine the end-point to color changed from green to red. The titer values are recorded for calculation of nitrogen content. A blank sample is usually run at the same time as the material being analyzed to take into account any residual nitrogen which may be in the reagents used to carry out the analysis. Once the nitrogen content has been determined it is converted to a protein content using the appropriate conversion factor. The nitrogen value was multiplied with a factor (F) of 6.25 to get estimate of crude protein content of samples (Singh et al.1999).

%Protein = Factor (F) X %N.

2.4. Statistical analysis

The mean protein values for each line were analyzed statistically by ANOVA using the R software. The descriptive statistics on ear traits were analyzed by one sample T test to check significant differences among lines and subjected to Principal component analysis, correlation and cluster analysis for genetic variability studies using STAR (Statistical Tool for Agricultural Research) software (Gulles et al.2014).

3. Results

3.1 Protein content of parents and Teosinte derived $BC_1F_{2:3}$ population

The protein content of parental lines and introgressed populations is presented in Table 2. The protein content of the pollen parent, Teosinte (*Zea mays ssp. parviglumis)*, was higher (19.67%) than that of the maize inbred line CML-451 (9.02%) which was used as a seed parent. One way Analysis of variance was performed with a null hypothesis of no significant difference between genotypes for mean protein content. The results of the investigation however yielded a highly significant difference among genotypes for protein content. ANOVA therefore indicates that there are adequate variability for protein content in the materials chosen for the analysis (Table 1). The protein content of introgressed lines varies substantially, ranging from **9.53–16.49%** for MP51 and MP6 lines, respectively. Out of 126 lines, four lines (MP6, MP20, MP36, and MP97) had protein content in the range of **16.07–16.49**%.whereas two lines namely MP56

and MP99 had protein content of 15.05 and 15.07.% respectively. Twenty three lines (MP5, MP11, MP12, MP17, MP18, MP24, MP27, MP28, MP31, MP35, MP46, MP58, MP63, MP79, MP80, MP81, MP104, MP107, MP111, MP113, MP119, MP121 and MP124) of the 126 varied in protein content from 14.02- to 14.85% The lines possessed protein content in the range of 13.01 .to 13.95.% were twenty five in number (MP16,MP23,MP29, MP30, MP39, MP40, MP47, MP49, MP52, MP55, MP59, MP65, MP77,MP78, MP82, MP83, MP84, MP86, MP88, MP89, MP94, MP101, MP109, MP118, MP125) whereas another set of thirty seven lines (MP2,MP3, MP4, MP7, MP14, MP19, MP22,MP26, MP33, MP41, MP48, MP50, MP53, MP54, MP61,MP60, MP62, MP64, MP66, MP70, MP71, MP74, MP87, MP92, MP93,MP96, MP98, MP102, MP103, MP105, MP110, MP112, MP115, MP116, MP117, MP123 and MP126) had protein content in the range of 12.00-to12.95%. Sixteen derived lines namely MP13,MP15,MP32,MP42, MP43, MP44, MP45, MP57, MP68, MP72, MP73, MP85, MP90, MP100, MP114 and MP120 possessed protein content of 11.22-to 11.96%, fourteen lines (MP1, MP9, MP10, MP25, MP34, MP37, MP38, MP67, MP69, MP76, MP91, MP95, MP106 and MP108) had protein content from 10.02 to 10.94%. Five lines namely MP8, MP21, MP51, MP75 and MP122 showed protein content of 9.53.to 9.98%. (Table: 2). All the $BC_1F_{2:3}$possessed protein content higher than the seed parent CML-451 but lower than the pollen parent teosinte. The increase in protein content in derived lines over CML-451 were varied from minimum of 9.53% in MP51to a maximum of 16.49% in MP6 Comparison of protein content of 126 lines with protein content of teosinte indicates that the derived possessed minimum of 9.53% protein to maximum of 16.49. % protein content of the pollen parent teosinte. The data across the 126 lines indicated an average increase in protein content of 41.42% over the seed parent CML-451 whereas in comparison to pollen parent, the derived lines showed only up to 16.49% protein content of pollen parent teosinte.

Table 1
ANOVA for protein content of parents and Teosinte derived $BC_1F_{2:3}$ population

Source of Variation	DF	Sum of Squares	Mean Squares	F-Calculated	Pr(> F)
Genotypes	127	1,013.06	7.98	484.61	< 2.2e-16 ***
Error	256	4.21	0.02		
Total	383	1,017.27			

Table 2
Mean protein contents with kernel shape of parents and teosinte derived $BC_1F_{2:3}$ population

Genotypes	Protein content %	Kernel shape	Genotypes	Protein content %	Kernel shape	Genotypes	Protein content %	Kernel shape	Genotypes	Protein content %	Kernel shape
Teosinte	19.67^{a}	P	MP26	12.95^{CDEF}	R	MP53	12.26^{PQR}	F	MP80	14.19^{lmn}	R
CML-451	9.02^{-}	F	MP27	14.48^{ij}	R	MP54	12.76^{FGHIJ}	R	MP81	14.02^{nop}	R
MP1	10.48^{678}	R	MP28	14.16^{mn}	R	MP55	13.17^{yzAB}	P	MP82	13.44^{stuv}	F
MP2	12.17^{QRST}	R	MP29	13.40^{tuvw}	F	MP56	15.05^{de}	F	MP83	13.95^{op}	P
MP3	12.62^{JKLM}	R	MP30	13.23^{vwxyzA}	F	MP57	11.92^{VWX}	F	MP84	13.18^{xyzAB}	P
MP4	12.07^{RSTUV}	R	MP31	14.26^{klm}	R	MP58	14.75^{fg}	R	MP85	11.91^{VWX}	R
MP5	14.72^{fgh}	R	MP32	11.57^{Z}	F	MP59	13.05^{ABCD}	R	MP86	13.36^{tuvwxy}	F
MP6	16.49$_{b}$	R	MP33	12.88^{DEFGHI}	F	MP60	12.55^{KLMN}	F	MP87	12.04^{STUVW}	R
MP7	12.69^{IJKL}	P	MP34	10.67^{56}	F	MP61	12.89^{DEFGHI}	P	MP88	13.45^{stu}	F
MP8	9.94 .	R	MP35	14.04^{nop}	R	MP62	12.71^{HIJKL}	F	MP89	13.38^{tuvwx}	R
MP9	10.78^{45}	F	MP36	16.07^{c}	R	MP63	14.27^{klm}	P	MP90	11.48^{Z1}	R
MP10	10.66^{56}	F	MP37	10.94^{4}	F	MP64	12.52^{LMN}	F	MP91	10.37^{89}	R
MP11	14.15^{mno}	R	MP38	10.33^{89}	R	MP65	13.61^{rs}	R	MP92	12.00^{TUVWX}	R
MP12	14.83^{f}	P	MP39	13.13^{zABC}	R	MP66	12.62^{JKLM}	F	MP93	12.25^{PQRS}	R
MP13	11.80^{XY}	R	MP40	13.46^{stu}	P	MP67	10.26^{9}	R	MP94	13.87^{pq}	R
MP14	12.71^{HIJKL}	R	MP41	12.91^{DEFGH}	F	MP68	11.22^{3}	R	MP95	10.90^{4}	F
MP15	11.28^{123}	R	MP42	11.62^{YZ}	R	MP69	10.02 .	F	MP96	12.07^{RSTUV}	R
MP16	13.26^{uvwxyz}	R	MP43	11.96^{UVWX}	R	MP70	12.06^{RSTUV}	F	MP97	16.28^{b}	P
MP17	14.05^{nop}	R	MP44	11.27^{23}	R	MP71	12.76^{FGHIJ}	F	MP98	12.65^{JKLM}	R
MP18	14.85^{ef}	R	MP45	11.27^{23}	R	MP72	11.46^{Z12}	R	MP99	15.07^{d}	P
MP19	12.47^{MNO}	F	MP46	14.56^{ghij}	R	MP73	11.30^{123}	R	MP100	11.84^{WX}	R
MP20	16.34^{b}	R	MP47	13.69^{qr}	F	MP74	12.92^{CDEFG}	F	MP101	13.30^{tuvwxyz}	R
MP21	9.98 .	R	MP48	12.54^{KLMN}	P	MP75	9.89 .	F	MP102	12.45^{MNOP}	R
MP22	12.75^{FGHIJ}	R	MP49	13.67^{qr}	P	MP76	10.24^{9}	F	MP103	12.39^{NOP}	P
MP23	13.31^{tuvwxyz}	R	MP50	12.29^{OPQ}	P	MP77	13.31^{tuvwxyz}	R	MP104	14.05^{nop}	R
MP24	14.38^{jkl}	R	MP51	9.53^{+}	R	MP78	13.20^{wxyzAB}	R	MP105	12.86^{DEFGHI}	R
MP25	10.30^{89}	F	MP52	13.49^{rst}	P	MP79	14.57^{ghij}	P	MP106	10.62^{567}	F

Different shapes- P = pointed; F = flat; R = round

Treatments with the same letter are not significantly different.

3.2. Ear traits of parents and Teosinte derived $BC_1F_{2:3}$ population

The one sample t test was performed to analyze individual line data without replication because data was recorded on F2:3 single cobs and found significant differences (p = 0.001) among genotypes for all traits (Table:3). The *Zea may ssp. parviglumis* had ear characteristics of 5-6cm ear length, 0.51cm ear width, two kernel rows ear, contains 5–8 kernels per ear and test weight of 3.64g (naked kernels) and 7.04g (with seed coat) (Fig. 1,Table 2). Maize inbred line CML-451 had ear length of 14.85 cm, ear diameter of 3.41 cm, 14 kernel rows, 30 kernels per row and test weight of 28.56g. The data on individual ears of parents

and introgressed populations are illustrated in Table 3. The introgressed population exhibited wide variation for all ear traits. Ear length ranged from 6.8 to 23.5cm for MP74 and MP115 lines, respectively. Ear width varied from 2.0cm for MP7 and MP12 lines to 3.5cm for MP123 line. Number of kernel rows per ear also varied significantly and ranged from 8 (MP12, MP21, MP22, MP24, MP29, MP52, MP56, MP60 and MP68) to 16 (MP108 and MP118) lines respectively. Number of kernels per row varied from 4 (MP17, MP20 and MP21) to 30 (MP93) representing wide variation for number of kernels per row among teosinte derived maize lines. The results also indicated genetic variation for kernel weight; teosinte kernels are small enclosed within a hard seed coat and maize kernels are naked and bold, while teosinte derived maize lines exhibited differences for test weight ranging from 12.36g (MP60) to 30.58 g (MP111). Apart from ear characteristics, introgressed lines were observed for kernel shape. Teosinte seeds are black in colour, with brown pointed kernels, whereas maize kernels are bold, flat, and bright yellow in colour (Fig. 2). The teosinte derived $BC1F_{2:3}$ population had a variety of kernel forms and observed round, flat and pointed shaped kernels (Table 2). Because of the introgression of genetic material from teosinte into the maize background, an introgressed offspring's kernels shape has been modified. Among 126 teosinte derived lines, 73 lines had round shaped kernels, 33 lines had flat shaped kernels and remaining 21 lines showed pointed kernel shape. The simple linear regression analysis was done by plotting the protein content of each line against the kernel shape of the respective lines. The results revealed a non-significant correlation between protein content and kernel shape, indicating that kernel shape is not an effective trait for the selection of higher protein-content lines (Fig:3). However, when lines were grouped into three based on the shape of kernels, the round-shaped kernel group has a mean protein content of 12.73%, the flat-shaped kernel group has a 12.07% mean protein content, but the pointed kernel shape group has the highest mean protein content of 13.71 compared to other groups. Still, one notable observation was that all pointed kernel-shaped lines contained more than 12% protein content, and MP97 introgressed line displayed a pointed kernel shape with a protein content of 16.28%. Hence, the data disclosed that the shape of kernels was not enough to select the lines having higher protein due to a lack of clear distinction of introgressed lines into maximum and minimum protein content lines based on kernel shape.

Table 3
Ear traits of parents and Teosinte derived $BC_1F_{2:3}$ population (parents, MP1-MP62)

Genotypes	Ear Length (cm)	Ear Width (cm)	Kernel Rows/Ear	No. of Kernels/Row	Test Weight (g)	Genotypes	Ear Length (cm)	Ear Width (cm)	Kernel Rows/Ear	No. of Kernels/Row	Test Weight (g)
Teosinte	5.5	0.509	2	4	3.64	**MP31**	9.5	2.004	12	16	21.42
CML-451	14.85	3.409	14	32	28.56	**MP32**	13	2.307	10	17	23.62
MP1	14.3	2.501	10	10	26.26	**MP33**	12	2.208	10	15	23.57
MP2	9.4	2.502	10	6	23.05	**MP34**	14.5	3.005	12	19	16.13
MP3	12.5	2.507	10	10	22	**MP35**	8.7	2.109	10	20	18.42
MP4	16.2	3.008	14	17	26.63	**MP36**	9	2.509	12	11	18.4
MP5	9.7	2.506	12	10	14.94	**MP37**	10.5	2.509	14	15	18.7
MP6	10.8	2.502	12	12	21.23	**MP38**	12.6	2.601	14	20	12.72
MP7	12.5	2.002	10	12	20.97	**MP39**	12.7	2.708	14	21	16.09
MP8	11.8	2.702	10	5	22.6	**MP40**	13	2.205	10	12	17.212
MP9	11.7	2.203	12	10	20.96	**MP41**	14.9	3.206	12	22	26.6
MP10	9.7	2.805	12	7	24.04	**MP42**	10.4	2.906	14	16	16.94
MP11	10.4	2.504	14	6	18.23	**MP43**	15	3.003	12	24	16.56
MP12	8.5	2.002	8	7	20.85	**MP44**	13.7	2.908	14	24	16.2
MP13	10.4	2.005	10	9	18.37	**MP45**	9.7	2.401	10	15	14.42
MP14	8	2.203	10	8	21.97	**MP46**	12.5	2.301	12	14	27.34
MP15	12.6	2.604	10	18	24.84	**MP47**	10.5	2.602	12	18	15.4
MP16	12.7	2.503	12	17	17.81	**MP48**	11.7	3.207	12	17	24.04
MP17	13	3.003	14	4	18.83	**MP49**	14.3	2.406	10	8	21.798
MP18	8	2.005	12	5	13.36	**MP50**	9.7	3.002	14	10	20.96
MP19	9.8	2.005	12	6	28	**MP51**	15.5	2.907	12	24	24.62
MP20	12.8	2.104	10	4	16.75	**MP52**	12.6	2.408	8	16	22.5
MP21	12	2.102	8	4	21.86	**MP53**	9.4	3.105	14	18	18.5
MP22	13	2.204	8	12	21.44	**MP54**	13.5	2.706	12	16	18.24
MP23	14.2	2.506	10	14	23.7	**MP55**	13.7	2.803	10	11	28.022
MP24	13.6	2.401	8	15	22.09	**MP56**	14.2	2.504	8	14	21.5
MP25	12.7	2.205	12	13	27	**MP57**	10	3.007	14	24	13.05
MP26	8.5	2.408	14	8	18.2	**MP58**	11.6	2.809	12	17	15.12
MP27	9.8	2.009	10	12	21.275	**MP59**	12.8	2.708	12	12	23.54
MP28	11.5	2.608	10	8	15.87	**MP60**	8.5	2.308	8	11	12.36
MP29	10	2.303	8	15	26.53	**MP61**	14	2.601	10	23	17.86
MP30	8.5	2.205	10	12	23.82	**MP62**	11.8	2.604	12	15	20.82
MP63	11.6	2.804	12	15	22.76	**MP97**	16.2	3.107	12	21	26.604
MP64	9.7	2.609	10	14	19.58	**MP98**	10.8	2.502	12	12	16.88
MP65	14.8	2.505	10	11	20.348	**MP99**	16.5	3.103	12	15	24.775
MP66	14	2.608	10	17	19.76	**MP100**	12	2.809	14	19	16.03
MP67	11.7	2.503	12	19	20.76	**MP101**	10.7	2.502	12	8	19.62
MP68	18.5	2.409	8	5	29.5	**MP102**	11.3	2.703	10	20	22.3
MP69	10.8	2.808	12	15	16.976	**MP103**	12.6	2.801	12	14	18.38
MP70	12	3.001	10	19	21.8	**MP104**	11.8	2.407	10	11	18.7

Genotypes	Ear Length (cm)	Ear Width (cm)	Kernel Rows/Ear	No. of Kernels/Row	Test Weight (g)	Genotypes	Ear Length (cm)	Ear Width (cm)	Kernel Rows/Ear	No. of Kernels/Row	Test Weight (g)
MP71	10.8	3.004	12	21	17.342	**MP105**	16.4	2.709	12	20	21.02
MP72	16.5	3.003	14	24	14.28	**MP106**	11.8	2.704	12	20	21.016
MP73	11.9	2.902	12	15	25.24	**MP107**	13.4	2.402	12	9	13.54
MP74	6.8	2.309	12	17	14.76	**MP108**	10.6	2.708	16	21	10.56
MP75	12.7	2.503	12	17	12.81	**MP109**	15	2.504	10	21	27.91
MP76	9.7	2.808	14	17	15.14	**MP110**	12.7	2.309	12	13	15.42
MP77	16.8	2.709	10	17	22.04	**MP111**	14	2.804	10	10	30.58
MP78	13	2.602	10	11	20.4	**MP112**	10	3.306	14	21	24.54
MP79	9.4	2.609	12	12	19.08	**MP113**	12.3	2.401	10	16	18.044
MP80	9.5	2.209	12	6	22.725	**MP114**	15	2.804	12	21	19.26
MP81	9.5	2.506	12	13	15	**MP115**	23.5	3.105	10	8	20.29
MP82	8.5	2.007	10	11	16.76	**MP116**	10.5	2.909	14	18	20.234
MP83	14.5	2.905	14	15	20.72	**MP117**	10.7	2.603	12	16	20.32
MP84	12.5	2.702	12	13	20.828	**MP118**	15.5	3.207	16	20	26.12
MP85	16.2	3.008	12	17	19.87	**MP119**	10.8	2.607	10	13	20.19
MP86	10	2.506	10	15	21.98	**MP120**	12.8	2.505	10	16	14.174
MP87	9.5	2.706	12	14	14.48	**MP121**	15.8	2.301	12	18	15.83
MP88	12.8	2.904	12	22	22.12	**MP122**	14.5	3.208	14	25	19.06
MP89	16.8	2.807	14	19	18.67	**MP123**	9.5	3.501	12	17	28.28
MP90	11.6	2.508	10	14	19.112	**MP124**	7.8	2.605	10	9	20.36
MP91	15	2.805	10	28	17.66	**MP125**	14.5	3.109	14	18	18.344
MP92	11.7	2.602	12	16	20.354	**MP126**	9.7	2.601	12	10	15.712
MP93	16.7	2.801	10	30	22.68	**Mean**	12.15	2.6	11.39	14.91	19.90
MP94	12.3	2.507	10	22	15.37	**StdDev**	2.63	0.38	1.95	5.67	4.28
MP95	13	2.803	12	26	15.44	**SE_Mean**	0.2326	0.0334	0.1726	0.5015	0.3785
MP96	11	2.805	12	11	18.942	**t Value**	52.23	78.03	65.98	29.74	52.58

Principal component analysis:

The variability amongst the Teosinte derived BC1F2:3 lines was analyzed using principal component analysis (PCA). Out of the six PCs, the first three explained 75.61 percent of the total variation in the data (Table 4). The perusal of the Table 4 indicated that the first, second, third, fourth, fifth, and sixth PCs accounted for 41.81, 20.47, 13.33, 12.11, 8.27, and 4.01 percent of the total variation, respectively. The PC1 accounts for 41.81 percent of total variability with positive eigen values for Ear Length (0.3634), Ear Width (0.5562), No of Kernel Rows/Ear (0.4253), No. of Kernels per Row (0.4418) and test weight (0.1853) while eigen value for protein content on first axis was negative (-0.3896). Ear Width was the character which accounted for maximum variability on PC 1. The PC2 accounted for 20.47% of total variance of the data The major characters that made a significant contribution to the second component were protein content (0.1232), Ear Length (0.4833), Ear Width(0.0098) and test weight (0.7349) which possessed positive Eigen values whereas No. of. Kernels/Row (-0.1900), No of Kernel Rows/Ear (-0.4183) had negative Eigen values on second principal component. The PC3 had 13.33% of the total variability observed in population and traits positively contributed to this axis were Ear Width (0.1545), No of Kernel Rows/Ear (0.4106) and test weight (0.4501) while protein content (-0.2224), Ear Length (-0.4946) and No. of. Kernels/Row (-0.5576) showed negative Eigen values on PC3. However, test weight trait, provided more variability to PC2 and PC3, with values of 0.7349 and 0.4501, respectively.

Table 4
Eigen value, contribution of variability and eigen vectors for the principal component axis of introgressed lines

Parameters	PC1	PC2	PC3	PC4	PC5	PC6
Eigen Values	2.5087	1.2285	0.7996	0.7264	0.4962	0.2407
Proportion of Variance	41.81	20.47	13.33	12.11	8.27	4.01
Cumulative Proportion	41.81	62.29	75.61	87.72	95.99	100
Protein .content	-0.3896	0.1232	-0.2224	0.8651	-0.1864	-0.0205
Ear Length (cm)	0.3634	0.4833	-0.4946	0.0899	0.5882	-0.1890
Ear Width (cm)	0.5562	0.0098	0.1545	0.3016	-0.0242	0.7584
Kernel Rows/Ear	0.4253	-0.4183	0.4106	0.3859	0.1946	-0.5374
No. of Kernels/Row	0.4418	-0.1900	-0.5576	-0.0605	-0.6422	-0.2044
Test Weight (g)	0.1752	0.7349	0.4501	-0.0042	-0.4105	-0.2412

Table 5
Correlation matrix among the ear and kernel traits of introgressed lines

	Protein content (%)	Ear Length (cm)	Ear Width (cm)	No of Kernel Rows/Ear	No. of Kernels/Ro w
Ear Length (cm)	-0.1911*				
Ear Width (cm)	-0.3816**	0.4299**			
No of Kernel Rows/Ear	-0.3249**	0.0834	0.6232**		
No. of. Kernels/Row	-0.3391**	0.3285**	0.5024**	0.3335**	
Test Weight (g))	-0.1036	0.3089**	0.2689**	-0.0526	-0.0351
Significance levels ** = P < 0.01, * = P < 0.05					

3.4. Cluster analysis

One hundred twenty-six Teosinte derived maize lines along with parents were subjected to cluster analysis using Ward's method agglomerative clustering, applying squared Euclidean distance as the distance measure, and were grouped into 6 clusters based on protein content and test weight of kernels (Table 6). The cluster I was consisted of Teosinte (*Zea mays ssp. parviglumis*) which had protein content of 19.67% and test weight of 3.645g. Cluster II grouped 4 lines (MP46, MP97, MP99 and MP111) in which mean of protein content was 15.13% and kernels test weight of 27.32g. Clusters III had 28 lines (MP6, MP11, MP12, MP17, MP20, MP24, MP27, MP31, MP35, MP36, MP49, MP52, MP56, MP63, MP65, MP77, MP78, MP79, MP80, MP83, MP84, MP86, MP88, MP101, MP104, MP113, MP119 andMP124,) having 14.25% and 20.47 g cluster means for protein content and test weight, respectively. Cluster IV possessed maximum number of lines (46) with mean protein content of 12% and a test weight of 23.51g. The seed parent, CML-451, was also grouped in this cluster. Thirteen lines grouped in Cluster V (MP34, MP37, MP38, MP44, MP69, MP75, MP76, MP90, MP91, MP95, MP108, MP114 and MP122) has mean protein content and test weight of 10.60% and16.14g respectively. Cluster VI contained 36 lines (MP5, MP13, MP16, MP18, MP26, MP28, MP39, MP40, MP42, MP43, MP45, MP47, MP53, MP54, MP57, MP58, MP60, MP61, MP71, MP72, MP74, MP81, MP82, MP87, MP89, MP94, MP96, MP98, MP100, MP103, MP107, MP110, MP120, MP121, MP125 and MP126) in which cluster mean for test weight was 16.12g protein content of 12.92%. The dendrogram was constructed by the Ward method based on Euclidean distance coefficient matrices, categorizing parents and 126 teosinte introgressed lines into 6 major clusters based upon the Euclidean distance of protein content and test weight of kernels combined together (Fig: 4). The maximum diversity was observed between ClusterI (Teosinte) and the rest of the clusters with a Euclidean genetic distance of more than 8. This classification re-affirmed the greater genetic variability among teosinte-derived maize lines for test weight and protein content of kernels. Teosinte falls under one cluster which might be due to having entirely different test weight and protein content than maize-derived lines. Within each cluster, more similar lines were grouped and genetically dissimilar lines were grouped between clusters. Forming larger number of clusters represents the wider diversity for traits considered for classification of lines and data represented in the Table 6.The scatter plot matrix (Fig: 5) was also created to know the association between protein content and test weight of kernels and observed an interrelation of protein content and test weight of kernels but in a negative direction.

Table 6
Cluster mean for total protein content and Test weight of parents and Teosinte derived $BC_1F_{2:3}$ population

Cluster Name	Name of Genotypes	Number of Genotypes	Cluster mean for protein content %	Cluster mean for Test weight(g)
Cluster1	Teosinte (*Zea mays ssp. parviglumis*)	1	19.67	3.64
Cluster2	MP46 MP97 MP99 MP111	4	15.13	27.32
Cluster3	MP6 MP11 MP12 MP17 MP20 MP24 MP27 MP31 MP35 MP36 MP49 MP52 MP56 MP63 MP65 MP77 MP78 MP79 MP80 MP83 MP84 MP86 MP88 MP101 MP104 MP113 MP119 MP124	28	14.25	20.47
Cluster4	CML-451 MP1 MP2 MP3 MP4 MP7 MP8 MP9 MP10 MP14 MP15 MP19 MP21 MP22 MP23 MP25 MP29 MP30 MP32 MP33 MP41 MP48 MP50 MP51 MP55 MP59 MP62 MP64 MP66 MP67 MP68 MP70 MP73 MP85 MP92 MP93 MP102 MP105 MP106 MP109 MP112 MP115 MP116 MP117 MP118 MP123	46	12.00	23.51
Cluster5	MP34 MP37 MP38 MP44 MP69 MP75 MP76 MP90 MP91 MP95 MP108 MP114 MP122	13	10.60	16.14
Cluster6	MP5 MP13 MP16 MP18 MP26 MP28 MP39 MP40 MP42 MP43 MP45 MP47 MP53 MP54 MP57 MP58 MP60 MP61 MP71 MP72 MP74 MP81 MP82 MP87 MP89 MP94 MP96 MP98 MP100 MP103 MP107 MP110 MP120 MP121 MP125 MP126	36	12.92	16.12

Correlation analysis: A correlation study indicated a different degree of association among the ear traits and protein content of teosinte derived lines and parents. Protein content exhibited significant negative correlation coefficients of -0.1911, -0.3861, -0.3249, and − 0.3391 and non-significant correlation coefficient of -0.1036 with ear length, ear width, number of kernel rows per ear, number of kernels per row and test weight of kernels, respectively (Table: 5) Ear length was significantly and positively correlated with ear width, number of kernel/rows per row and test weight of kernels. Similarly, ear width was significantly and positively correlated with number of kernel rows per ear, number of kernels per row and test weight. Number of kernel rows per ear exerted significant positive association with number of kernels per row while non-significant negative with test weight. Number of kernels per row implied negative correlation with test weight.

4. Discussion

Genetic variability has greatly reduced in cultivated maize due to domestication bottlenecks and both natural and artificial selections (Doebley et al. 2004) which causes a severe effect on availability allelic divergence in primary gene pool for agriculturally important traits such as biotic and abiotic stresses as well as nutritional parameters (Yamasaki et al.2007; Maazou et al. 2021). By and large genetic diversity is low in the case of protein content in cultivated maize compared to their wild relatives (Flint-Garcia et al.2013; Zhang et al.2017). Apart from plant architecture, the major trait responsible for domestication syndrome is kernel size and composition, which is important in terms of human and animal nutrition. Maize wild progenitor teosintes have a huge repository of novel and diverse alleles for many parameters including protein content (Paulis and Wall 1977). Allelic introgression from teosintes for diversification and enhancement of protein content in maize is seems to be very valid proposition considering the limited genetic variation for protein in cultivated maize and in primary gene pool. Several researchers created introgressed teosinte populations using different species of teosinte such as *Zea perennis* (Perini and Magoja 1988), *Zea diploperennis* (Perini et al.1991), *Zea. mays ssp. mexicana* (Magoja and Nivio, 1982) and *Zea mays ssp. parviglumis* (Doebley et al.1994); these populations were utilized for the study of kernel composition and quality parameters and reported higher average protein content than maize parent. In the present study, we explored *Zea mays ssp. parviglumis* for diversification and enhancement of protein content of maize inbred line CML-451 and realized diversification in derived lines with wide range of protein content and elevated up to ~ 16%. Four $BC_1F_{2:3}$ lines namely MP6, MP20, MP36, and MP97 of the 126 analysed had protein content of around 16% which are considered as transgressive segregants which are having improved quantity of protein content and also exhibits higher test weight of kernels compared to rest of the lines derived from teosinte similar results are akin the Wang et al. (2008), who noted significant improvement in protein content over the maize parent, with about 26.6% increased protein content among *Z. mays ssp. mexicana* introgressed population. Perini and Magoja (1988) used *Zea perennis* and developed derived population in which they noted 50% higher protein content over the protein content of maize. Recently, Karn et al. (2017) identified 12 alleles from teosinte and three QTLs for protein content explaining 23% of the total variation and also highlighted major QTL on chromosome 3 acts as major driving force for loss of alleles for protein content during domestication for about 60%and starch content was increased to 34%. They observed increased protein content in population derived by alien introgression approach. Further, they opined that using wild relative nutritional quality of maize can be improved successfully. The recent advancement in maize protein science was the cloning of the higher total protein content QTL present in "Zea mays ssp Parviglumis", i.e., teosinte high protein 9 (Thp9), which is found on chromosome 9, encodes the enzyme asparagine synthetase 4, the expression of which was found to be higher in teosinte than in cultivated maize germplasm (Wu et al. 2022). In addition one of the important observations was about kernel shape. Teosinte-derived lines show variation for the trait and 126 lines are categorized into round, flat, and pointed shaped kernels compared to flat shaped kernels of maize parent. Similar results were observed by Wang et al. (2008), who observed triangle-shaped kernels in wild derived lines but did not mention protein content of triangle-shaped lines and their interrelation with kernel shape. Likewise, Liu et al. (2016b) also studied kernel size and shape characters with their effects on kernel weight and discovered allelic effects but no information regarding the protein content

of introgressed lines. Even though we observed non significant linear regression between kernel protein content and shape, about > 12% of protein content was found in all pointed-shaped kernels (Table: 1). The kernel shape was modified from flat to pointed and round in progeny lines derived from teosinte due to genetic material introgression into the maize genome, which led to the inclination of the maize kernel shape towards the kernel shape of teosinte (Fig: 2). Therefore, the results shows improved protein content of teosinte-derived lines which clearly indicates the potential of *prviglumis*-teosinte and can be utilized in pre-breeding programmes for the development and diversification of maize lines for kernel protein content.

Domestication of crops from their wild ancestors has created numerous morphological and physiological variations which made them adaptable to various agricultural cropping systems. Domesticated from teosinte (*Zea mays ssp. parviglumis*) by single domestication event leads to two major modifications in maize is plant architecture and female inflorescence (Matsuoka et al., 2002). Since grain is an economical part of maize, major emphasis is given to female inflorescence characters, which is an important morphological difference between maize and *Zea mays ssp. parviglumis.* Thus, the striking difference between maize and its wild relatives is the female inflorescence, i.e. ear, which leads to various conclusions regarding maize origin and the diversity of species (Iltis 2000; Orr et al.2002). Teosinte ear has few seeds that are trapezoidal in shape and covered in stony seed coats (Brown et al. 2011). The ear and kernel traits are so significant because of their direct contribution to yield and have been domesticated for many years (Liu et al., 2016b). Introgressive hybridization is considered to be useful for morphological characterization of introgressed lines for grasping knowledge regarding the behavior of ear and kernel traits, and desired effects can be reintroduced into cultivated maize germplasm from teosinte (Wang et al.2012). There is a whopping allelic effect observed among teosinte derived near isogenic lines ranging from − 2.33 to 0.24 and identified four major QTLs on chromosomes 1, 2, 4, and 5 for Kernel Row Number (Liu et al. 2016b). Artificial selection and domestication undoubtedly increased the kernel size of maize compared to the wild progenitor due to small effects of individual loci and identified QTLs for kernel weight using F2 maize-teosinte populations (Doebley et al.1994; Briggs et al.2007). Liu et al. (2016b) reported eight QTLs for the weight of kernels and proved how wild relatives cherish beneficial allelic forms for kernel characteristics and their composition. In this study, a wide range of diversity was observed among parental and introgressed lines for all ear traits. Our observations are consistent with those of Kumar et al. (2019), who characterized a BC1F4 teosinte derived population for various ear traits and noted wide diversity for all ear traits ranging from 8.00 cm to 13.50 cm for ear length; 1.50 cm to 3.00 cm for ear diameter; 10 to 13 for the number of kernel rows/ear; 10 to 27 for the number of kernel rows per row; and 12.1 to 17.8 g for test weight among introgressed lines. Diversification of ear associated parameters is the key features we noted while analyzing 126 $BC_1F_{2:3}$ lines. Maazou et al. (2021) while investigating the SNP-based molecular characterization on ear and kernel traits realized the potential of *Zea mays ssp. Mexicana* in broadening the genetic base using introgressive breeding approach. Alien gene introgression from teosinte has successfully employed with substantial modification and improvement in kernel traits (kernel composition, kernel weight, kernel row number, kernel area, and kernel length) by Liu et al. (2016a, b); Karn et al. (2017).

Estimating genetic diversity and aligning the population diversity in simplified and meaningful forms make it easier for selection and further advancement of genetic material based on certain morphological characters (Sokolov and Guzhva 1997). Principal component analysis and cluster analysis have proved to be powerful statistical tools for analyzing diversity and also reducing the dimensionality of multivariate data (Mohammadi and Prasanna 2003). This categorizes the lines based on the variability of parameters, ultimately helping in the selection of diverse lines for breeding purposes (Shrestha, 2016). To investigate maize diversity, Goodman first time used multivariate analysis in 1967. Smith et al. (1984) distinguished different teosinte species using a PCA plot for taxonomic classification, and a scatter plot displays genetic distances and similarity based on dot aggregation among individual genotypes. In the present study, we observed substantial diversification of introgressed lines. Doebley et al. (1991) also noted morphological differences between maize and teosinte, as well as variability for kernel traits among maize-teosinte F2 populations. In this investigation, Principal component analysis and cluster analysis grouped 126 wild derived lines into six principal components and clusters respectively. Numerous reports available on studies of genetic diversity using principal component analysis; Pinheiro de Carvalho et al.(2008) analyzed diversity among 43 open pollinated varieties based on 41 morphological characters grown in different ecological conditions helps in preserving their identity and exploitation of varietal adaptability. Sánchez González et al. (2018) have also applied PCA to study ecological distribution of different teosinte species based on several descriptors. Maazou et al. (2021) employed principal component analysis to display genetic correlations between parents and mexicana introgressed germplasm for detection of alleles by SNP genotyping. In our study, traits which are investigated for variability in $BC_1F_{2:3}$ population are protein content, ear length (cm), ear width (cm), kernel rows/ear and test weight (g). The results are consistent with those of earlier reports of Amegbor et al. (2022). Cluster I having Teosinte is entirely different from remaining 5 clusters which is represented in dendrogram (Fig: 5). Cluster II contained those lines which possessed higher values for both protein content and test weight and can be a potential source for protein and test weight in line development programme. Clusters III had lines with improved protein content of 14.25% but have lower test weight of kernels of about 20.47g. Members of Cluster IV had mean protein content of 12% which is relatively low when compared with other member of the $BC_1F_{2:3}$ lines yet it is quite higher than the protein content in maize. However, test weight of this cluster is 23.51g which is lower than the maize parental line parent CML-451. Cluster V lines showed lower value for both traits i.e. protein content was 10.60% and 16.14g of test weight. Cluster VI is characterized of 12.92% of protein and test weight of 16.0 g. The protein content in lines decreased dramatically as the test weight of kernels increased. Therefore Selection of greater protein-containing lines without lowering test weight, as well as lines with higher values for both features, are given priority in breeding programmes (Xu et al.2021). Recurrent selection methods will be advisable for the improvement of test weight of kernels (Wiersma et al. 2001). Adu et al. (2019) also applied cluster analysis for classification of 96 inbred lines based on structure of population by using SNP markers. The substantial genetic variability among wild-derived populations (*Zea mays* ssp. *parviglumis*) were noted for morphological traits based on cluster analysis and further reported significant diversity among teosinte derived populations compared to maize germplasm with the help of SSR markers (Adhikari et al. 2021).

Correlation studies among variables help breeders with an indirect selection of superior lines for further breeding work (Amegbor et al.2022). Quantitative characters are highly influenced by climatic conditions due to their polygenic and low heritability nature (Qi et al. 2010), especially in maize kernel composition trait are governed by polygenes and highly influenced by environmental conditions (Wilson et al. 2004). Hence character associations are useful for the selection of superior lines based on related traits (Amini et al. 2013). Correlation analysis in the present experiment between protein content and ear traits exhibited different degrees of associations, which might aid in the development of effective selection criteria. We found a negative correlation between protein content and ear traits. Partially similar results were observed by Jio et al. (2021), who reported a negative correlation between protein quality and kernel and

yield traits. Additionally, protein content has a non-significant negative relationship with the test weight, validating the work of Langyan et al. (2021). In pulses, there is a negative correlation between protein content and seed size (Saxena et al. 1984), but in maize, a similar trend is observed in the current study, correlating with the results of Langyan et al. (2021). Kernels of teosinte are very small but have higher protein content. While undergoing domestication, probably some allelic variants for seed size and protein content were lost/modified, leading altered grain size and nutritional/biochemical composition with protein content of nearly half of its ancestor, and higher starch content of the kernel (Paulis and Wall 1977). Similar observation over the years by scientists proved the hypothesis that selection for bold seeds (higher test weight) leads to a substantial decrease in the protein content of kernels in different crops (Sadras et al. 2007; Garcia et al 2009).

Conclusion And Future Prospects

In this study, we determined to investigate whether teosinte could be exploited in development and diversification of maize germplasm for one of the important but limiting trait protein content in maize. Surprisingly, we found that the BC1F2:3 population had higher protein content than the maize parent CML-451, demonstrating the potential of wild progenitor *parviglumis*-teosinte in bio-fortification of maize for improved protein content. We also looked at ear and kernel characteristics, and found a large range of variance for all of the traits studied. Wide diversity was observed among teosinte-derived lines based on the results of principal component analysis and cluster analysis. Proteins content, ear length (cm), ear width (cm), kernel rows/ear and test weight (g) are traits responsible for greater genetic variability. In addition we also highlighted effect of teosinte genetic introgression changes the kernel shape from flat shape to pointed or triangular shaped kernels and their association with protein content. Correlation studies pore limelight on domestication syndrome's effects on nutritional aspects and give the inter-relation among ear and kernel traits and protein content and found strong negative association between protein content and the rest of kernel and ear traits. Recurrent selection may be opted to break the negative relationship for developing lines with higher protein content and higher test weight of kernels. The major emphasis in the study given on increment of total protein quantity using *Zea mays ssp. parviglumis*; quality parameters like Lysine and Tryptophan content can be analyzed to see if these traits exist in wild maize species. Furthermore, the Zea genus contains other species with a wide range of variability in total protein content and protein quality, which can be recommended for improving maize nutritional traits by incorporating those species into large-scale hybridization programmes and speeding up the bio-fortification of maize for protein content.

Abbreviations

Cm: centimeter

g: grams

MP: Maize-*parvaglumis*

PCA: principal component analysis

PEM: protein-energy malnutrition

QPM: Quality Protein Maize

Declarations

Ethics Approval and Consent to participate: Investigation is line with the research ethics and also complied with the ethics of the journal. All authors are given consent to participate.

Consent for publication: All authors are given consent for publication.

Competing Interests: Authors do not have any financial or non-financial interests that are directly or indirectly related to the work submitted for publication.

Funding: No funding is available for this particular investigation and also preparation of manuscript.

Author's Contribution: VS, NKS: Concept and Planning; VS, NP: Protein analysis; VS, NKS: Wrote main manuscript, prepared tables and figures, VS, NKS, NP, SV: Reviewed the manuscript; all authors have read and approved the final manuscript.

Acknowledgements: the work was supported by the G.B. Pant University of Agriculture and Technology, Pantnagar, Uttarakhand (India). We are thankful to Department of Soil Science in the same university for providing laboratory facilities for analyzing protein content of genotypes.

Data Availability: All raw data files generated during study were provided in supplementary files.

The wild species utilized in the current study is locally adapted Teosinte species of maize; we recommend that authors comply with the IUCN Policy Statement on Research Involving Species at Risk of Extinction and the Convention on the Trade in Endangered Species of Wild Fauna and Flora. We also ensure that authors followed all procedures in accordance to the guidelines.

References

1. Adhikari, S., Joshi, A., Kumar, A., & Singh, N. K. (2021). Diversification of maize (Zea mays L.) through teosinte (Zea mays subsp. parviglumis Iltis & Doebley) allelic. Genetic Resources and Crop Evolution, *68*(7), 2983–2995. doi.org/10.1007/s10722-021-01170-z

2. Amegbor, I. K., van Biljon, A., Shargie, N., Tarekegne, A., & Labuschagne, M. T. (2022). Heritability and Associations among Grain Yield and Quality Traits in Quality Protein Maize (QPM) and Non-QPM Hybrids. *Plants*, *11*(6), 713. doi.org/10.3390/plants11060713
3. Amini, Z., Khodambashi, M., & Houshmand, S. (2013). Correlation and path coefficient analysis of seed yield related traits in maize. *International Journal of Agriculture and Crop Sciences (IJACS)*, *5*(19), 2217–2220. URL: http://ijagcs.com/.../2217-2220.pdf
4. Bouis, H. E., & Saltzman, A. (2017). Improving nutrition through biofortification: a review of evidence from HarvestPlus, 2003 through 2016. Global Food Security, *12*, 49–58. https://doi.org/10.1016/j.gfs.2017.01.009
5. Briggs, W. H., McMullen, M. D., Gaut, B. S., & Doebley, J. (2007). Linkage mapping of domestication loci in a large maize–teosinte backcross resource. Genetics, *177*(3), 1915–1928. doi.org/10.1534/genetics.107.076497
6. Brown, P. J., Upadyayula, N., Mahone, G. S., Tian, F., Bradbury, P. J., Myles, S., & Rocheford, T. R. (2011). Distinct genetic architectures for male and female inflorescence traits of maize. PLoS genetics, *7*(11), e1002383. doi.org/10.1371/journal.pgen.1002383
7. Dei, H. K. (2017). Assessment of maize (Zea mays) as feed resource for poultry. Poultry science, *1*, 1–32. http://dx.doi.org/10.5572/65363
8. Doebley J, Stec A, Gustus C (1995) Teosinte branched1 and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics 141:333–346. https://doi.org/10.1093/genetics/141.1.333
9. Doebley J. The genetics of maize evolution. Annu Rev Genet. 2004;38:37–59. DOI: 10.1146/annurev.genet.38.072902.092425
10. Doebley, J., & Stec, A. (1991). Genetic analysis of the morphological differences between maize and teosinte. Genetics, *129*(1), 285–295. https://doi.org/10.1093/genetics/129.1.285
11. Doebley, J., Bacigalupo, A., & Stec, A. (1994). Inheritance of kernel weight in two maize-teosinte hybrid populations: implications for crop evolution. Journal of Heredity, *85*(3), 191195.https://doi.org/10.1093/oxfordjournals.jhered.a111434
12. Ekpa, O., Palacios-Rojas, N., Kruseman, G., Fogliano, V., & Linnemann, A. R. (2019). Sub-Saharan African maize-based foods-processing practices, challenges and opportunities. Food Reviews International, *35*(7), 609–639. https://doi.org/10.1080/87559129.2019.1588290
13. Flint-Garcia, S. A. (2013). Genetics and consequences of crop domestication. Journal of Agricultural and Food Chemistry, *61*(35), 8267–8276. https://doi.org/10.1021/jf305511d
14. Flint-Garcia, S. A., Bodnar, A. L., & Scott, M. P. (2009). Wide variability in kernel composition, seed characteristics, and zein profiles among diverse maize inbreds, landraces, and teosinte. Theoretical and Applied Genetics, *119*(6), 1129–1142. DOI 10.1007/s00122-009-1115-1
15. Gonzalez-Segovia, E., Pérez-Limon, S., Cíntora-Martínez, G. C., Guerrero-Zavala, A., Janzen, G. M., Hufford, M. B., … Sawers, R. J. (2019). Characterization of introgression from the teosinte Zea mays ssp. mexicana to Mexican highland maize. PeerJ, *7*, e6815. https://doi.org/10.7717/peerj.6815
16. Goodman, M. M. (1968). The Races of Maize: II. Use of Multivariate Analysis of Variance to Measure Morphological Similarity 1. *Crop Science*, *8*(6), 693–698. https://doi.org/10.2135/cropsci1968.0011183X000800060016x
17. Holding, D. R., & Larkins, B. A. (2009). Zein storage proteins. In *Molecular genetic approaches to maize improvement* (pp. 269–286). Springer, Berlin, Heidelberg. doi: 10.1007/978-3-540-68922-5_19
18. Hopkins, C. G. (1899). Improvement in the chemical composition of the corn kernel. Journal of the American Chemical Society, *21*(11), 1039–1057. https://doi.org/10.1021/ja02061a012
19. Hossain, F., Sarika, K., Muthusamy, V., Zunjare, R. U., & Gupta, H. S. (2019). Quality protein maize for nutritional security. In *Quality breeding in field crops* (pp. 217–237). Springer, Cham.DOI: 10.1007/978-3-030-04609-5_11
20. Hufford, M. B., Xu, X., Van Heerwaarden, J., Pyhäjärvi, T., Chia, J. M., Cartwright, R. A., … Ross-Ibarra, J. (2012). Comparative population genomics of maize domestication and improvement. Nature genetics, *44*(7), 808–811.
21. Iltis, H. H. (2000). Homeotic sexual translocations and the origin of maize (Zea mays, Poaceae): A new look at an old problem. Economic Botany, *54*(1), 7–42.
22. Jilo, T. (2021). Nutritional benefit and development of quality protein maize (QPM) in Ethiopia. Cereal Research Communications, 1–14.
23. Karn, A., Gillman, J. D., & Flint-Garcia, S. A. (2017). Genetic analysis of teosinte alleles for kernel composition traits in maize. *G3: Genes, Genomes, Genetics*, *7*(4), 1157–1164. https://doi.org/10.1534/g3.117.039529
24. Kiprop, J. C. (2020). Malnutrition: Current challenges and future perspectives. http://dx.doi.org/10.5772/intechopen.92007
25. Kumar, A., Singh, N. K., Adhikari, S., & Joshi, A. (2019). Morphological and molecular characterization of teosinte derived maize population. Indian Journal of Genetics and Plant Breeding, *79*(04), 670–677. https://doi.org/10.31742/IJGPB.79.4.4
26. Langyan, S., Dar, Z. A., Chaudhary, D. P., Shekhar, J. C., Herlambang, S., El Enshasy, H., … Rakshit, S. (2021). Analysis of nutritional quality attributes and their inter-relationship in maize inbred lines for sustainable livelihood. *Sustainability*, *13*(11), 6137. https://doi.org/10.3390/su13116137
27. Liu, Z., Cook, J., Melia-Hancock, S., Guill, K., Bottoms, C., Garcia, A., & Flint-Garcia, S. A. (2016a). Expanding maize genetic resources with predomestication alleles: maize–teosinte introgression populations. *The plant genome*, *9*(1), plantgenome2015-07. https://doi.org/10.3835/plantgenome2015.07.0053
28. Liu, Z., Garcia, A., McMullen, M. D., & Flint-Garcia, S. A. (2016b). Genetic analysis of kernel traits in maize-teosinte introgression populations. *G3: Genes, Genomes, Genetics*, *6*(8), 2523–2530. https://doi.org/10.1534/g3.116.030155
29. Maazou, A. R. S., Qiu, J., Ai, L., Arabzai, M. G., Qian, M., Gul, H., & Liu, Z. (2021). Preliminary Phenotypic and SNP-Based Molecular Characterization of Maize (Zea mays L.)-Mexicana (Zea mays SSP. Mexicana) Introgression Lines under Inbred Background of 48 – 2. *American Journal of Plant Sciences*, *12*(7), 1073–1089. DOI: 10.4236/ajps.2021.127075
30. Magoja, J.L., Nivio, A.A., 1982. Specific relationship of storage proteins in maize and its wild relatives. Maize Genetics Cooperation Newsletter 56, 106e107.

31. Matsuoka, Y., Vigouroux, Y., Goodman, M. M., Sanchez, J., Buckler, E., & Doebley, J. (2002). A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences*, *99*(9), 6080–6084. https://doi.org/10.1073/pnas.052125199
32. Meseka, S., Fakorede, M., Ajala, S., Badu-Apraku, B., & Menkir, A. (2013). Introgression of alleles from maize landraces to improve drought tolerance in an adapted germplasm. Journal of Crop Improvement, *27*(1), 96–112. https://doi.org/10.1080/15427528.2012.729259
33. Mohammadi, S. A., & Prasanna, B. M. (2003). Analysis of genetic diversity in crop plants—salient statistical tools and considerations. *Crop science*, *43*(4), 1235–1248.https://doi.org/10.2135/cropsci2003.1235
34. Neeraja, C. N., Babu, V. R., Ram, S., Hossain, F., Hariprasanna, K., Rajpurohit, B. S., & Datta, S. K. (2017). Biofortification in cereals: progress and prospects. Current Science, 1050–1057. https://www.jstor.org/stable/26494165
35. Neto, G. C., Yunes, J. A., da Silva, M. J., Vettore, A. L., Arruda, P., & Leite, A. (1995). The involvement of Opaque 2 on β-prolamin gene regulation in maize and Coix suggests a more general role for this transcriptional activator. Plant molecular biology, *27*(5), 1015–1029. doi: 10.1007/BF00037028
36. Nyakurwa, C. S., Gasura, E., & Mabasa, S. (2017). Potential for quality protein maize for reducing proteinenergy undernutrition in maize dependent Sub-Saharan African countries: A review. *African Crop Science Journal*, *25*(4), 521–537.: http://dx.doi.org/10.4314/acsj.v25i4.9
37. Orr, A. R., Mullen, K., Klaahsen, D., & Sundberg, M. D. (2002). Inflorescence development in a high-altitude annual Mexican teosinte (Poaceae). *American Journal of Botany*, *89*(11), 1730–1740. https://doi.org/10.3732/ajb.89.11.1730
38. Osgood-Zimmerman, A., Millear, A. I., Stubbs, R. W., Shields, C., Pickering, B. V., Earl, L., … Hay, S. I. (2018). Mapping child growth failure in Africa between 2000 and 2015. Nature, *555*(7694), 41–47. doi:10.1038/nature25760
39. Paulis, J. W., & Wall, J. S. (1977). Comparison of the protein compositions of selected corns and their wild relatives, teosinte and Tripsacum. Journal of Agricultural and Food Chemistry, *25*(2), 265–270. https://doi.org/10.1021/jf60210a022
40. Perini, L. A., & Magoja, J. L. (1988). Effect of perennial teosinte introgression in maize on kernel protein content. Maize Genet Coop Newsl, *62*, 80.
41. Perini, L.H., Pischedda, G., Magoja, J.L., 1991. Diploperennial teosinte introgressed population of maize: kernel protein content. Maize Genetics Cooperation Newsletter 65, 4
42. Pinheiro de Carvalho, M. Â. A., Ganança, J. F. T., Abreu, I., Sousa, N. F., dos Santos, T. M. M., Clemente Vieira, M. R., & Motto, M. (2008). Evaluation of the maize (Zea mays L.) diversity on the Archipelago of Madeira. Genetic Resources and Crop Evolution, *55*(2), 221–233. DOI 10.1007/s10722-007-9230-9
43. Qi, X., Li, Z. H., Jiang, L. L., Yu, X. M., Ngezahayo, F., & Liu, B. (2010). Grain-Yield Heterosis in Zea mays L. Shows Positive Correlation with Parental Difference in CHG Methylation. *Crop science*, *50*(6), 2338–2346.https://doi.org/10.2135/cropsci2010.01.0023
44. Ranum, P., Peña-Rosas, J. P., & Garcia-Casal, M. N. (2014). Global maize production, utilization, and consumption. *Annals of the new York academy of sciences*, *1312*(1), 105–112. https://doi.org/10.1111/nyas.12396
45. Rincon, F., Johnson, B., Crossa, J., & Taba, S. (1996). Cluster analysis. An approach to sampling variability in maize [Zea mays] accessions. *Maydica (Italy)*.
46. Sadras, V. O. (2007). Evolutionary aspects of the trade-off between seed size and number in crops. Field Crops Research, *100*(2–3), 125–138. https://doi.org/10.1016/j.fcr.2006.07.004
47. Sahoo, S., Adhikari, S., Joshi, A., & Singh, N. K. (2021). Use of wild progenitor teosinte in maize (Zea mays subsp. mays) improvement: present status and future prospects. Tropical Plant Biology, *14*(2), 156–179.https://doi.org/10.1007/s12042-021-09288-1
48. Sánchez González, J. D. J., Ruiz Corral, J. A., García, G. M., Ojeda, G. R., Larios, L. D. L. C., Holland, J. B.,& García Romero, G. E. (2018). Ecogeography of teosinte. PLoS One, *13*(2), e0192676. https://doi.org/10.1371/journal.pone.0192676
49. Saxena, K. B., Faris, D. G., Singh, U., & Kumar, R. V. (1987). Relationship between seed size and protein content in newly developed high protein lines of pigeonpea. Plant Foods for Human Nutrition, *36*(4), 335–340. Doi:10.1007/bf01892354
50. Schmidt, R. J., Ketudat, M., Aukerman, M. J., & Hoschek, G. (1992). Opaque-2 is a transcriptional activator that recognizes a specific target site in 22-kD zein genes. The Plant Cell, *4*(6), 689–700. https://doi.org/10.1105/tpc.4.6.689
51. Shrestha, J. (2016). Cluster analysis of maize inbred lines. Journal of Nepal Agricultural Research Council, *2*, 33–36. https://doi.org/10.3126/jnarc.v2i0.16119
52. Singh, D., Chhonkar, P. K., & Pande, R. N. (1999). Soil reaction in soil, plant, water analysis method: manual. IARI, ICAR, New Delhi, *1*(4.2), 11–13.
53. Singh, N. K., Kumar, A., Chandra, H., Pal, K., & Verma, S. S. (2017). Enhancement of maize allelic diversity using wild relative teosinte (Zea mays ssp. parviglumis). Indian Journal of Plant Genetic Resources, *30*(3), 253–257. DOI: 10.5958/0976-1926.2017.00032.8
54. Smith, J. S. C., Goodman, M. M., & Stuber, C. W. (1984). Variation within teosinte. III. Numerical analysis of allozyme data. Economic Botany, *38*(1), 97–113.
55. Sokolov, V. M., & Guzhva, D. V. (1997). Use of qualitative traits for genotypic classification of inbred maize lines. Kukuruza I sorgo, *3*, 8–12.
56. Swarup, S., Timmermans, M. C., Chaudhuri, S., and Messing, J. (1995). Determinants of the high-methionine trait in wild and exotic germplasm may have escaped selection during early cultivation of maize. Plant J. 8, 359–368.
57. Vasal, S. K., Villegas, E., Bjarnason, M., Gelaw, B., & Goertz, P. (1980). Genetic modifiers and breeding strategies in developing hard endosperm opaque-2 materials. *Genetic modifiers and breeding strategies in developing hard endosperm opaque-2 materials.*, 37–73.
58. Wang, L., Gu, X., Wang, X., Qu, M., Luan, J., Zhang, X., & Zhang, J. (2012). Introgression of genetic material from Zea mays ssp. mexicana into cultivated maize was facilitated by tissue culture. Pak. J. Bot, *44*(4), 1419–1426.
59. Wang, L., Xu, C., Qu, M., & Zhang, J. (2008). Kernel amino acid composition and protein content of introgression lines from Zea mays ssp. mexicana into cultivated maize. Journal of Cereal Science, *48*(2), 387–393. https://doi.org/10.1016/j.jcs.2007.09.014

60. Wiersma, J. J., Busch, R. H., Fulcher, G. G., & Hareland, G. A. (2001). Recurrent selection for kernel weight in spring wheat. *Crop Science*, *41*(4), 999–1005. https://doi.org/10.2135/cropsci2001.414999x
61. Wilson, L. M., Whitt, S. R., Ibaéñez, A. M., Rocheford, T. R., Goodman, M. M., & Buckler IV, E. S. (2004). Dissection of maize kernel composition and starch production by candidate gene association. The Plant Cell, *16*(10), 2719–2733. https://doi.org/10.1105/tpc.104.025700
62. Wu, Y., & Messing, J. (2014). Proteome balancing of the maize seed for higher nutritional value. Frontiers in plant science, *5*, 240. https://doi.org/10.3389/fpls.2014.00240
63. Wu, Y., Huang, Y., Wang, H., Zhu, Y., Huang, X., Li, S., & Wang, W. (2022). Teosinte high protein 9 enhances the seed protein content and nitrogen utilization efficiency in maize. Research Square (Preprint) DOI: https://doi.org/10.21203/rs.3.rs-1701760/v1
64. Yamasaki, M., Wright, S. I., & McMullen, M. D. (2007). Genomic screening for artificial selection during domestication and improvement in maize. Annals of Botany, *100*(5), 967–973. https://doi.org/10.1093/aob/mcm173
65. Zhang, H., Mittal, N., Leamy, L. J., Barazani, O., & Song, B. H. (2017). Back into the wild—Apply untapped genetic diversity of wild relatives for crop improvement. *Evolutionary Applications*, *10*(1), 5–24. https://doi.org/10.1111/eva.12434

Figures

Figure 1

Kernels of *Zea mays ssp. parviglumis.* a) With seed coat, b) without seed coat

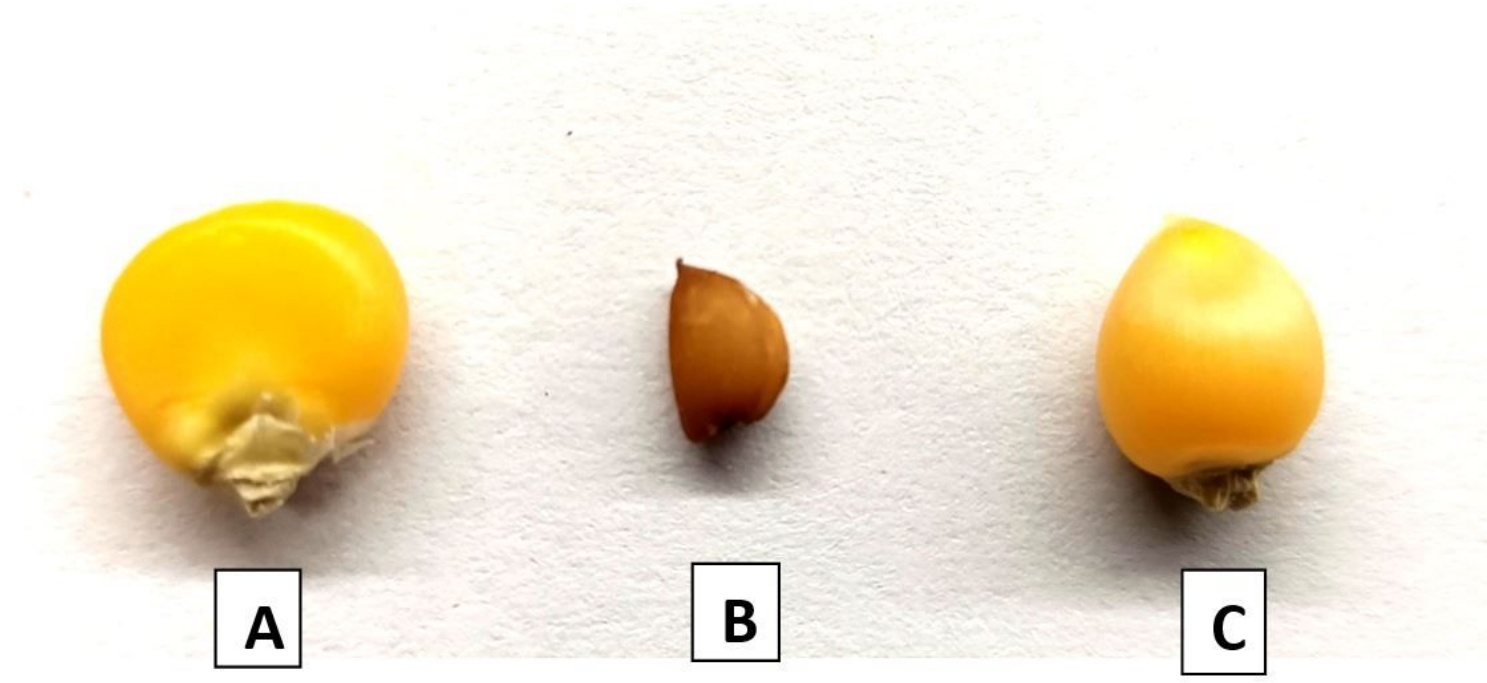

Figure 2

Kernels of two parents and progeny line; A. CML-451; B. *Zea mays ssp. parviglumis*, C. Introgressed progeny line MP97

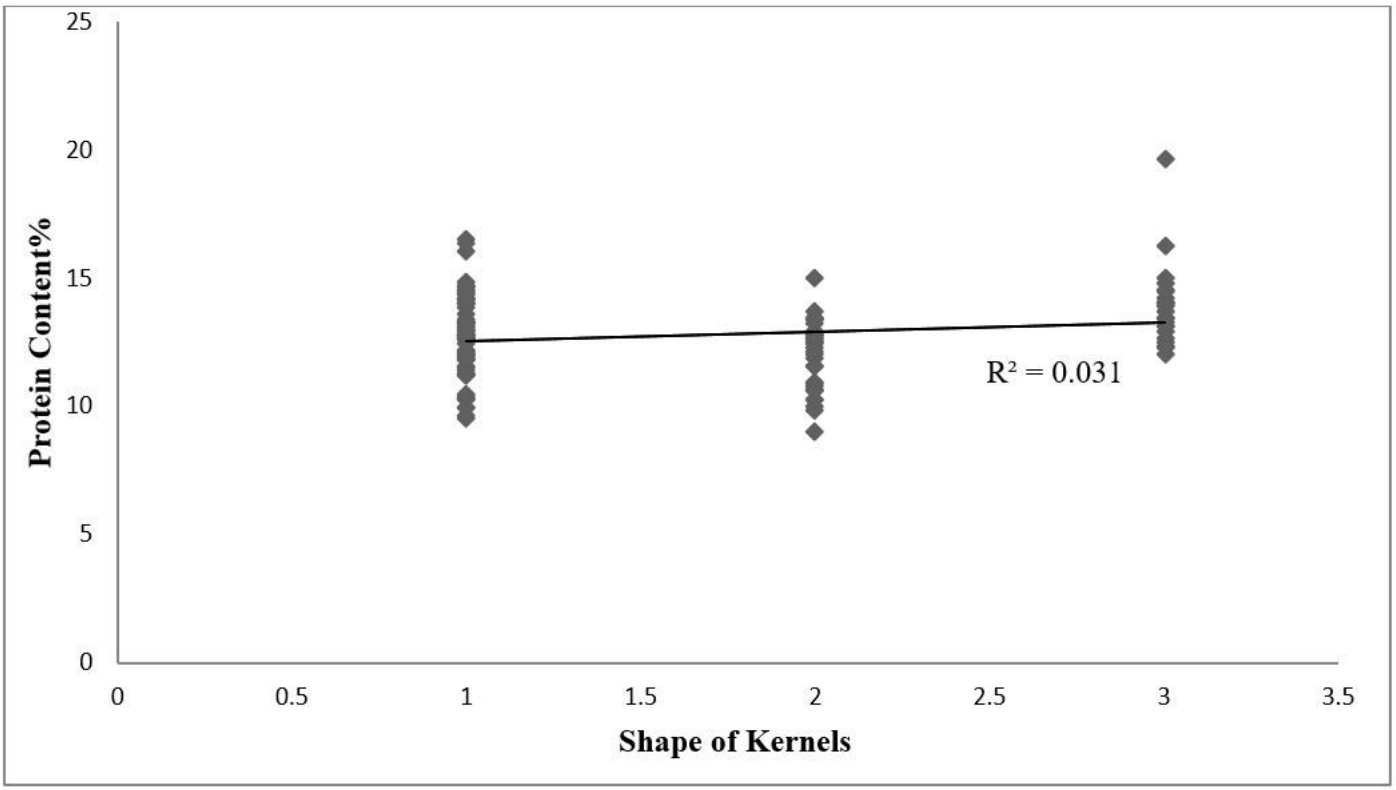

Figure 3

Kernel shape and Protein content of parents and teosinte derived 126 $BC_1F_{2:3}$ population. Graph represents interrelationship between shape and protein content of kernels, kernel shape scores includes: 1= Round, 2= Flat and 3= Pointed kernels.

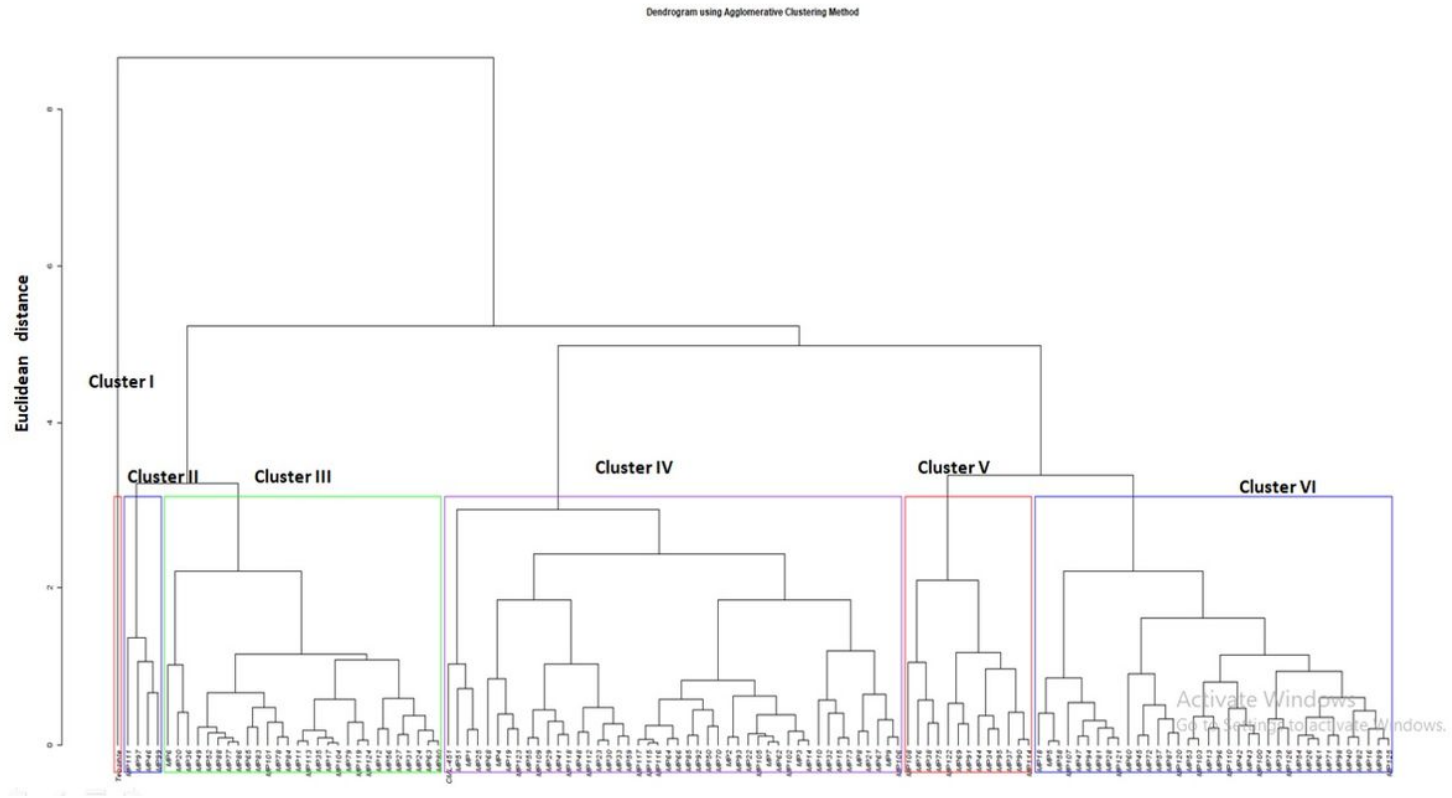

Figure 4

Dendrogram based on cluster agglomerative clustering of parents and teosinte derived 126 $BC_1F_{2:3}$ population based on test weight and protein content of lines.

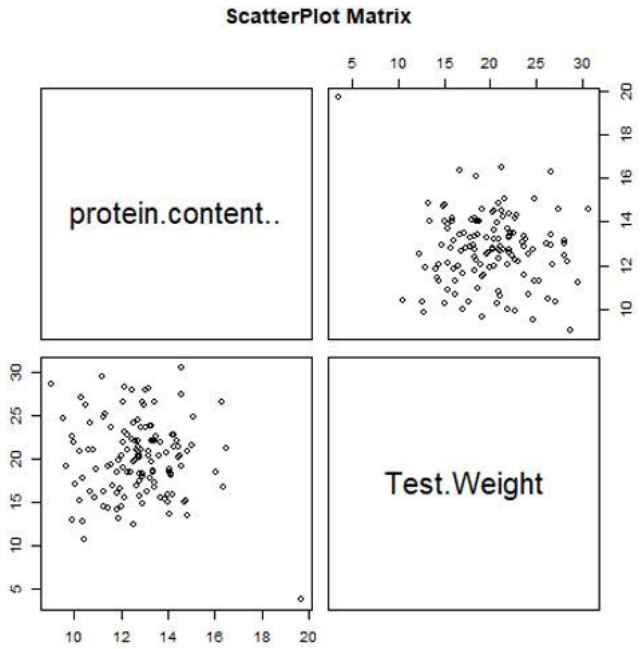

Figure 5

A simple Scatter plot matrix illustrating relationship between protein content and Test weight of Teosinte introgressed lines kernels

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementary.xlsx