GENETIC DIVERGENCE ANALYSIS ON SOME BREAD WHEAT GENOTYPES GROWN IN ETHIOPIA

*Daniel Hailegiorgis, Mebrahtom Mesfin and S. K. Gangwar
College of Agriculture, Aksum University, Aksum, Tigray, Ethiopia.
*Corresponding author email: danid981@yahoo.com

ABSTRACT

An investigation was carried out with 49 bread wheat genotypes to assess the genetic diversity for yield and yield related traits. The Unweighted Pair Group Method using Arithmetic Averages (UPGMA) technique was used for the analysis. The genotypes were evaluated for 17 Characters and showed wide variability for the components studied. The cluster analysis grouped the 49 bread wheat genotypes into 22 different clusters. This indicates the presence of wide diversity among the tested genotypes. From cluster mean values, genotypes in cluster 9 and 13 deserve consideration for their direct use as parents in hybridization programs to develop high yielding wheat varieties. The genotypes in cluster 3 and 16 may be used for improvement of protein and gluten contents, early maturity-ness and other desirable characters other than grain yield. The result of the principal components analysis revealed that nine principal components (PC1 to PC9) accounted nearly 80 % of the total variation. It was also noted that differentiation of genotypes into different clusters was because of the small contribution of few character rather than the cumulative effect of a number of characters. The information obtained from this study can be used to plan crosses and maximize the use of genetic diversity and expression of heterosis.

KEYWORDS: Hetrosis, hybridization, genotypes, genetic diversity.

INTRODUCTION

Wheat (*Triticum spp.*) is the second major food crop of the world in its importance next to rice. In Ethiopia the crop ranks third in terms of total production next to teff and maize. It is largely grown in the highlands of the country and constitutes roughly 20-30 % of the annual cereal production and plays an appreciable role of supplying the production with carbohydrates, proteins and minerals (Schulthess, *et al.*, 2000).

For a successful breeding program, the presence of genetic diversity and variability play a vital role. Genetic diversity is essential to meet the diversified goals of plant breeding such as breeding for increasing yield, wider adaptation, desirable quality, pest and disease resistance. Genetic divergence analysis estimates the extent of diversity existed among selected genotypes (Mondal, 2003). Precise information on the nature and degree of genetic diversity helps the plant breeder in choosing the diverse parents for purposeful hybridization (Samsuddin, 1985). Since published works of bread wheat on genetic diversity is scanty, the aim of this experiment was to identify genetically divergent bread wheat parents with desirable traits for hybridization particularly for yield and biochemical traits like grain protein and gluten contents.

MATERIALS AND METHODS

Forty-nine bread wheat genotypes/lines which are comprised of different crosses (hybrids), nationally released and introduced varieties were used in the experiment (Table-1). The experiment was conducted in Adet Agricultural Research Center (AARC) during 2007/8 summer season in simple lattice design in six rows of 20 cm spacing and 2.5 m-row length. Sowing was done by hand drilling at a seed rate of 150 kg/ha. The experimental

plots were prepared by tractor ploughing and harrowing. The spacing between plots and replication were 0.2 m and 1.5 m, respectively. DAP and Urea fertilizers were applied at the rate of 100 kg/ha and 150 kg/ha, respectively. All the cultural practices were performed as recommended. The plant data during the cropping season and after harvesting were noted. Observations recorded on plot basis include days to heading, days to maturity, grain filling period, 1000 seed weight, biological yield per plot, seed yield per plot, harvest index per plot, total protein, sedimentation test, starch and gluten contents. The seed yield per plot and biological yield per plot was measured by taking a net plot of 2.5 m * 1.2 m or 3 m^2 and this was used to determine harvest index. All other characters were recorded on single plant basis by randomly taking 5 plants from each experimental plot. The statistical package NCSS (Number Cruncher Statistical System) (Jerry, 2001) was used for genetic divergence calculation, principal component analysis and cluster mean analysis. Euclidean genetic matrices for 1225 pair wise comparisons of 49 genotypes were made using Unweighted Pair Group Method using Arithmetic means (UPGMA).

RESULT AND DISCUSSION
Genetic Divergence analysis
The unweighted pair group method using arithmetic averages (UPGMA) resulted in classifying the 49 bread wheat genotypes into 22 distinct clusters (Table 1). This indicates the presence of wide diversity among the tested genotypes.
Estimation of genetic distance and cluster analysis
The results of the analysis of variance (ANOVA) revealed the presence of significant difference among the tested genotypes for all the characters studied and this justified

the need to estimate genetic distance values for the genotypes. The estimate of Euclidean genetic distance value ranged from 0.47 (between ETBW5324 and ETBW5307) to 2.63 (between ETBW5320 and CRAKER/ST825) (Table 2). The frequency distribution of genetic distance value for all 1225 pairs of comparisons indicates that 95 % of the pair comparisons had values between 0.51 and 2.00. Only two pairs accounting 0.16 % of the total had genetic distance value smaller than 0.5 and four pair comparisons which are only 0.4 % of the total had values greater than 2.5. HAR2093/HAR604 and CRAKER/ST825 had the greatest genetic distance from all other genotypes included in this study. Therefore, the results of the genetic distance has shown that there is a room for the genetic improvement of bread wheat varieties and the information generated can be used to plan wide crosses, to exploit genetic diversity and maximize the expression of hetrosis.

Principal components analysis

Principal component analysis (PCA) reflects the importance of the largest contributor to the total variation at each axis of differentiation (Sharma, 1998). The eigenvalues are often used to determine how many factors to retain. The sum of the eigenvalues is usually equal to the number of variables. Therefore, in this analysis the first factor retains the information contained in 3.181 of the original variables. The coefficients defining the first nine principal components of these data are given in Table 3. The coefficients are scaled, so that they present correlations between observed variables and derived components.

Nine principal components, PC1 to PC9, which are extracted from the original data and having latent roots greater than one, accounting nearly 80 % of the total variation. Suggesting these principal component scores might be used to summarize the original 17 variables in any further analysis of the data. Out of the total principal components retained, PC1, PC3, PC8 and PC 4 with values of 18.71 %, 9.68 %, 9.22 % and 8.15 % respectively contributed more to the total variation.

According to Chahal and Gosal (2002) characters with largest absolute value closer to unity within the first principal component influence the clustering more than those with lower absolute value closer to zero. Therefore, in the present study, differentiation of the genotypes into different clusters was because of relatively high contribution of few characters rather than small contribution from each character. Accordingly, the first principal component had high positive component loading from protein content, sedimentation volume and wet gluten content; and high negative loading from grain yield, biomass yield and starch content. The positive and negative loading shows the presence of positive and negative correlation trends between the components and the variables. Therefore, the above mentioned characters which load high positively or negatively contributed more to the diversity and they were the ones that most differentiated the clusters.

Hence, the major contributing characters for the diversity in the third principal component (PC3) were days to maturity, harvest index and days to heading; grain filling period and days to heading in principal component four (PC4); biological yield, days to heading, grain yield and starch content in principal component eight (PC8).

Usually it is customary to choose one variable from these identified groups. Hence, for the first group protein content is best choice, which had the largest loading from component ones, days to maturity for the third, grain filling period for the fourth group, biological yield for the eighth group.

Cluster mean analysis

The mean value of the characters in each cluster is presented in Table 4. The characteristic feature of each cluster is discussed here below.

Cluster 1 consisted of 2 genotypes. This cluster was found to have a characteristic feature of low yielding, moderately high number of seeds per spike and having more number of tillers per plant. The genetic distance value of these varieties was 0.793. Two genotypes which include the standard check HAR 3730 constitute cluster 2. This cluster can be characterized by having high biomass production, low level of protein and wet gluten contents. The genetic distance value of these varieties was 0.806.Cluster 3 had two genotypes with a characteristic feature of very low grain yielding, biomass production and thousand seed weight but having high level of protein, wet gluten and starch contents. These cluster exhibited the greatest genetic distance from all other genotypes included in this study. According to Ghaderi et al., (1984), increasing parental distance implies a great number of contrasting alleles at the desired loci, and to the extent that these loci recombine in the F2 and F3 generation following a cross of distantly related parents, the greater will be the opportunities for the effective selection for yield factors. Thus, crossing of genotypes from these clusters with other clusters may produce higher amount of hetrotic expression in the first filial generations (F1's) and wide range of variability in subsequent segregating (F2) populations. The genetic distance value of these genotypes was 0.812.

Cluster 4 had two genotypes with a characteristics feature of long duration of maturity, relatively low yielding, small number of seeds per spike and small tiller number per plant. The genetic distance value of these varieties was 0.827. Cluster 5 had two genotypes and this cluster could be characterized by having long duration of maturity, low grain and biomass yield. The genetic distance value of these varieties was 0.835. Cluster 6 had two genotypes and this cluster could be characterized by having low grain and biomass yield but with high protein and wet gluten contents. The genetic distance value of these varieties was 0.836. Cluster 7 had two genotypes having a characteristics feature of long duration of heading, low starch content, relatively high grain yield and very high biomass yield. The genetic distance value of these varieties was 0.891. Cluster 8 had two genotypes. This cluster is mainly characterized by having high number of seeds per spike. The genetic distance value of these varieties was 0.926. Three genotypes constitute cluster 9 and this cluster had a characteristics of relatively high grain, biomass yields, high sedimentation volume and long grain filling period. Six genotypes constitute cluster 10 with characteristic feature of long spike length, high spikelets per spike and high sedimentation volume. The range of dissimilarity obtained in this cluster was 0.569

(between ETBW5354 and ETBW5341) to 1.172 (between ETBW5354 and CMSS96YO2555S4-0Y). The largest number of genotypes accounting 26.50 % of the total was included in cluster 11 and this cluster had characteristics of small tiller number per plant, relatively high protein and wet gluten contents and long grain filling period. The range of dissimilarity obtained was 0.503 (between variety 27 and variety 38) to 1.492 (between variety 32 and variety 36). The rest clusters were singletons.

Crosses involving parents belonging to more divergent clusters would be expected to manifest maximum hetrosis and wide variability in genetic architecture (Singh *et al*., 1987). In the present study, cluster 3 was more divergent than the others. However, the chance of getting segregates with high yield level is quite limited when one of the characters has a very low yield level. Cluster 3 had the lowest mean performance in seed yield and other important characters. This indicates that the chance of getting segregates with high yield is limited between crosses of cluster 3 with the other clusters. The selection of parents should also consider the special advantage of each cluster and each genotype within a cluster depending on specific objective of hybridization (Chahal and Gasal, 2002). Thus, crosses involving cluster 9 and 13 with any other cluster except cluster 3, 4, 6 and 22 are suggested to exhibit high hetrosis and could result in segregates with higher seed yield.

At a cutoff 1.00 the dendrogram revealed 22 distinct clusters which are stemmed from two main clusters. Eleven are real clusters and the rest eleven are singletons (Figure 1). Cophenotic value of 0.75 or more are usually recommended for the best fit of the cluster analysis. The present study confirmed a good fit with the genetic distance matrix by having cophenotic correlation coefficient of r = 0.756. The two delta goodness of fit statistics delta (0.5) and delta (1.0) were 0.141 and 0.178, respectively. When comparing to clustering configurations, the configuration with the smallest delta value fits the data better.

CONCLUSION AND RECOMMENDATION
The present investigation provided considerable information useful in genetic improvement of bread wheat. Genotypes grouped into cluster 3 showed maximum inter cluster diversity. From cluster mean values, genotypes in cluster 9 and 13 deserve consideration for their direct use as parents in hybridization programs to develop high yielding wheat varieties. The genotypes in cluster 3 and 16 may be used for improvement of protein and gluten contents, early maturity-ness and other desirable characters other than grain yield. There is significant genetic variability among tested genotypes that indicates the presence of excellent opportunity to bring about improvement through wide hybridization by crossing genotypes in different clusters.

Further studies on the influence of environment and agronomic practices on the genetic potential of the varieties in different wheat environment is necessary. This is helpful to stratify the environments based on quality and yield suitability. Generally, the development of bread wheat varieties possessing higher grain yield, reasonably higher protein content, stronger gluten and reasonably higher HLW is important.
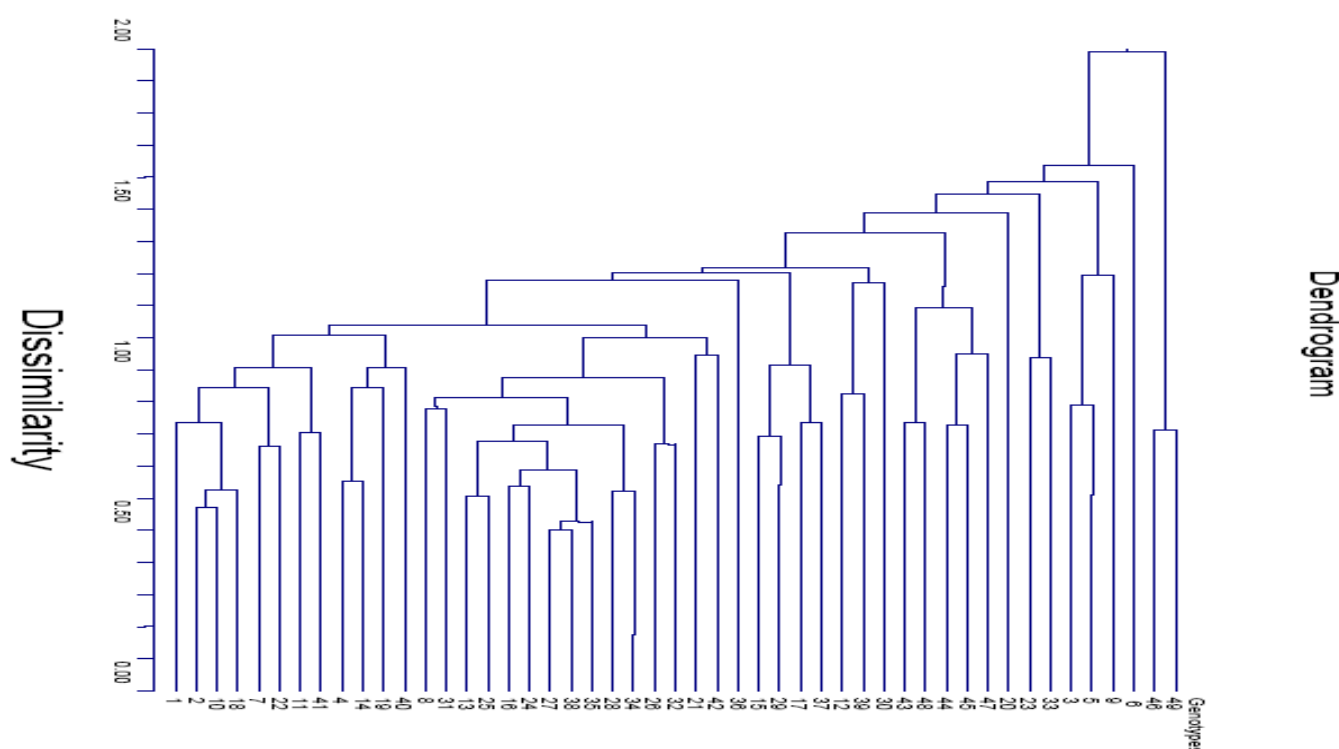

Figure 1. Dendrogram depicting genetic relationships among 49 bread wheat genotypes

REFFERENCES
Chahal, G.S., and Gosal, S.S. (2002) Principles and Procedures of Plant Breeding: Biotechnology and Conventional Approaches. Narosa Publishing House, New Delhi.

Ghaderi, A., Adams, M.W.and Nassib, A.M. (1984) Relationship between genetic distance and heterosis for yield and morphological traits in dry edible bean and faba bean. *Crop Sci. 24*: 37-42.

Jerry H.L. (2001) Number Cruncher Statistical System (NCSS). Statistical Software. Kaysville, Utah.

Mondal, M.A. (2003) Improvement of potato (Solanum tuberosum L.) through hybridization and in vitro culture technique. PhD Thesis. Rajshahi University, Rajshahi, Bangladesh.

Samsuddin, A.K. (1985) Genetic diversity in relation to heterosis and combining analysis in spring wheat. Theor. Appl. Genet., 70: 306-308.

Schulthess, U.B., Feil, J. and Jutzi, S.C. (2000) Yield independent variation in grain nitrogen and phosphorus concentration among Ethiopian wheat. Agron. J. 89(3): 497-506

Sharma, J.R., (1998) Statistical and Biometrical Techniques in Plant Breeding. New Age International (P) Limited Publishers, New Delhi. 432p.

Singh, S.K., Singh, R.S., Maurya, D.M., Verma, O.P. (1987) Genetic divergence among lowland rice cultivars: Indian Agricultural Research Institute (IARI), New Delhi, India.

Table 1. Genotypes distribution in to their respective clusters based on UPGMA analysis.

Cluster	Total number of genotypes	Genotypes (variety number)
1	2	ETBW5336, ETBW5321
2	2	ETBW5327, HAR 3730
3	2	HAR2093/HAR604, CRAKER/ST825
4	2	HAR1522/KARIGA, HAR1522/ST826
5	2	HAR1522/ST825, HAR1522/HAR1709
6	2	ETBW5352, ETBW5341
7	2	ETBW5314, ETBW5311
8	2	ETBW5338, ETBW5332
9	3	ETBW5331, ETBW5345, ETBW5320
10	6	CMSS96MO 3653S-0Y, ETBW5354, ETBW5348, ETBW5341, ETBW5343, CMSS96YO2555S4-0Y
11	13	ETBW5312, ETBW5330, ETBW5329, ETBW5322, ETBW5306, ETBW5342, ETBW5339, ETBW5349, ETBW5318, ETBW5351, ETBW5333, ETBW5324, ETBW5307,
12	1	OZDA
13	1	ETBW5308
14	1	CMSW95MO03808S-0Y
15	1	CMSS96MO3653-0Y
16	1	CMSW95WM00808S-10BOY
17	1	ETBW5325
18	1	ETBW5317
19	1	ETBW5327
20	1	CMSS97MOO3165-010Y
21	1	HAR 1685
22	1	HAR1522/HAR1868

Table 2. Euclidean genetic distance ranges of bread wheat (*T. eastivum*) genotypes

Genetic distance	Frequency	Frequency (%)
<0.50	2	0.16
0.51-1.00	251	20.49
1.10-1.50	596	48.65
1.51-2.00	318	25.89
2.01-2.5	54	4.41
>2.5	4	0.4

Table 3. Vector loadings and percentage explained variation by the first nine PCs

Characters	Eigenvectors								
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
DH	-0.040	-0.275	0.386	0.390	0.013	-0.199	-0.148	0.370	-0.055
DM	0.006	-0.100	0.495	-0.090	-0.303	-0.235	-0.339	0.153	-0.248
PH	-0.264	0.251	0.005	0.092	-0.260	0.406	-0.163	0.030	0.402
GY	-0.367	-0.001	-0.285	-0.170	0.126	-0.016	-0.140	0.307	-0.244
TSW	-0.138	0.300	-0.173	0.288	-0.292	-0.109	-0.459	0.016	0.216
SSp	-0.114	-0.396	-0.258	-0.247	-0.242	-0.109	0.080	0.178	0.197
SL	-0.104	-0.361	-0.159	0.290	-0.195	0.294	-0.032	-0.154	-0.065
SPS	-0.012	-0.418	0.006	0.007	-0.328	0.439	-0.120	-0.124	-0.183
TPP	0.017	-0.380	-0.060	-0.245	0.159	-0.330	-0.332	-0.136	0.591
BY	-0.355	0.122	-0.045	-0.250	0.041	0.124	0.019	0.574	-0.002
HLW	-0.011	0.028	-0.204	0.128	-0.588	-0.439	0.499	0.080	0.030
PRO	0.421	0.057	-0.188	0.015	-0.035	0.025	-0.254	0.161	-0.024
WG	0.408	0.113	-0.252	0.016	-0.033	0.009	-0.108	0.136	-0.016
STA	-0.321	0.207	-0.053	0.283	0.042	-0.285	-0.152	-0.305	-0.114
SDS	0.383	0.046	-0.236	0.116	-0.070	0.031	-0.178	0.287	-0.086
HI	-0.169	-0.178	-0.435	0.039	0.142	-0.190	-0.231	-0.159	-0.415
GFP	0.039	0.217	0.097	-0.583	-0.364	-0.039	-0.202	-0.272	-0.217
Eigenvalue	3.181	1.097	1.645	1.386	1.050	1.083	1.134	1.567	1.349
Individual percentage	18.71	6.45	9.68	8.15	6.18	6.37	6.67	9.22	7.94
Cumulative percentage	18.71	25.17	34.85	43.00	49.18	55.55	62.22	71.43	79.37

Table 4. Mean values of characters in the 22 clusters

Traits	Clusters																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
DH	71.4	80.83	75.92	77.74	77.81	79.60	81.40	69.45	68.15	70.49	65.39	78.41	71.48	63.26	57.79	57.40	58.26	54.98	61.08	66.41	68.71	82.84
DM	123.00	126.2	123.75	130.50	132.2	123.00	125.00	118.5	122.33	123.58	119.54	105.0	129.00	122.0	105.0	106.0	122.5	105.0	106.0	115.50	108.0	134.5
PH	83.31	86.43	55.16	85.97	85.83	78.46	81.09	95.60	98.20	90.48	80.83	82.09	82.87	97.19	105.1	79.41	80.35	85.73	86.85	81.29	85.61	79.51
GY	641.75	874.2	423.77	607.02	337.0	641.42	999.07	843.0	1038.8	871.42	786.87	964.7	1239.8	731.0	842.0	841.1	920.5	824.3	817.3	1057.1	884.5	613.3
TSW	31.02	31.22	22.61	30.09	32.29	34.68	30.25	29.98	33.16	32.27	30.85	35.21	30.54	27.19	33.30	29.46	29.94	36.16	32.07	33.55	27.21	26.05
SSp	52.32	47.19	49.45	41.76	37.14	46.95	52.72	57.74	44.57	47.95	48.12	48.88	59.82	36.64	47.32	51.86	55.54	41.57	55.66	43.28	45.97	46.14
SL	7.10	8.00	8.72	7.48	6.31	8.00	8.33	9.28	7.75	8.80	7.56	8.80	8.45	6.88	7.89	8.76	6.61	7.24	8.34	7.09	7.61	8.44
SPS	17.75	17.75	19.74	16.00	18.00	18.25	18.75	20.99	16.45	18.63	17.54	18.00	20.50	16.50	18.49	16.50	20.51	17.50	16.50	15.50	17.49	17.99
TPP	4.70	3.95	4.83	2.06	1.76	3.52	3.52	3.35	3.09	2.87	2.40	3.07	4.38	3.59	1.83	3.49	4.18	2.04	3.71	2.41	2.50	2.94
BY	1575	2000	1075	2050	1450	1475	2600	1900	2433	2100	1915	1800	2350	2050	2100	2000	2000	1900	1800	2000	2250	1600
HLW	72.85	72.40	69.80	76.30	71.50	74.00	70.07	72.80	69.08	74.21	73.28	73.60	75.20	64.10	73.60	71.30	75.40	75.60	78.80	73.90	71.20	75.30
PRO	11.85	9.97	12.50	11.00	11.12	12.57	9.35	10.70	10.88	10.71	11.34	11.20	9.75	11.45	11.35	12.70	11.35	12.80	10.10	10.05	10.10	11.65
WG	31.525	29.80	33.42	28.45	28.45	34.55	19.57	27.15	26.03	26.58	30.31	31.35	21.35	29.80	30.15	35.35	30.70	36.70	24.70	25.95	26.20	29.45
SDS	65.25	67.72	62.77	65.27	65.60	65.00	65.27	64.40	66.93	66.56	65.34	67.25	66.95	64.40	35.50	64.30	64.50	65.50	67.30	67.90	66.35	64.70
STA	38.32	29.40	46.07	35.50	36.95	48.75	28.17	32.32	31.65	34.81	37.32	36.05	29.50	34.30	40.05	53.25	36.05	55.55	28.35	31.70	31.95	37.80
HI	40.37	43.81	39.09	29.42	23.78	43.25	38.58	45.37	42.84	41.33	41.37	55.16	55.69	35.62	40.09	42.20	45.99	43.52	45.49	54.11	39.27	43.93
GFP	51.39	45.77	47.04	52.06	53.85	43.77	44.15	49.04	53.95	53.44	54.15	27.55	57.15	58.49	46.75	48.91	63.98	49.18	46.12	49.58	40.30	50.35

DH = Days to heading, DM = Days to maturity, PH = Plant height, GY = Grain yield, TSW = 1000 Seed weight, SSp = Seeds per spike, SL = Spike length, SPS = Spike lets per spike, TPP = Tiller per plant, BY = biological yield, HLW = Hecto liter weight, PRO = protein, WG = Wet gluten, STA = Starch, SDS = Sedimentation volume test, HI = Harvest index, GFP = Grain filling period