

MULTIVARIATE DIVERSITY, HERITABILITY AND GENETIC ADVANCE IN *Tef* LANDRACES IN ETHIOPIA

HABTAMU AYALEW, TSIGE GENET¹, TADESSE DESSALEGN¹ and LANDUBER WONDALE

Adet Agricultural Research Center P. O.Box 08, Bahir Dar, Ethiopia

¹Bahir Dar University College of Agriculture and Environmental Sciences

Corresponding author: habtsh99@yahoo.com

(Received 30 May, 2011; accepted 11 September, 2011)

ABSTRACT

Characterisation of exiting genetic variability is a prerequisite for further crop improvement activity. This study was designed to assess genetic variability among randomly selected *Eragrostis tef*, Zucc. Trotter (*Tef*) genotypes from five administrative zones in the Amhara region in Ethiopia. The experiment was conducted in 2010 main cropping season at Adet Agricultural Research Center. All traits, except first inter-node length showed highly significant differences among the 37 lines. Number of productive tillers per plant, grain yield per plant, and biomass yield per plant showed high phenotypic coefficients of variation; 18.9, 17.5 and 16.9% in that order. Harvest index (15.1%) showed the highest genotypic coefficient of variation while the lowest (3.5%) was for days to maturity. Heritability in broad sense was highest for days to heading (80.7%), followed by culm length (72.4%). Grain yield and shoot biomass yield showed heritability values of 54.6 and 57.3%, and GAM values of 18.9 and 20.6%, respectively. The first three principal components (PCs) with eigenvalues greater than one explained 75% of the observed variation. Four PCs were effective in explaining 93% of the variation among zones. Cluster analysis grouped the 37 lines into five real clusters, while zones of collection were grouped into three major clusters. These data are useful for future *tef* breeding/crop improvement programmes and undertakings.

Key Words: Clusters analysis, principal component analysis, *Tef*

RÉSUMÉ

La caractérisation de la variabilité génétique existante est une précondition dans des activités d'amélioration culturale. La présente étude était conçue pour évaluer la variabilité génétique parmi les génotypes sélectionnés de *tef* *Eragrostis*, Zucc. Trotter (*Tef*) dans cinq zones administratives dans la région Amhara en Éthiopie. L'étude était conduite au cours de la saison culturale principale 2010 au Centre de Recherche Agricole Adet. Toutes les caractéristiques, excepté la longueur inter-nœud ont montré de différences significativement élevées parmi les 37 lignées. Le nombre de tiges par plante, le rendement en grains par plante, et le rendement en biomasse par plante ont montré des coefficients de variation phénotypique élevés de 18.9, 17.5 et 16.9% respectivement. L'index de récolte (15.1%) a montré un coefficient de variation le plus élevé pendant que le coefficient génotypique le moins élevé (3.5%) était celui du nombre de jours à la maturité. L'héritabilité au sens large était plus élevée pour les jours relatifs au *heading* (80.7%), suivi de la longueur du culm (72.4%). Le rendement en grains et le rendement en biomasse de tiges ont montré des valeurs d'héritabilité de 54.6 et 57.3%, et les valeurs de GAM de 18.9 et 20.6%, respectivement. Les trois premières composantes principales (PCs) avec des valeurs *eigen* plus grandes que un, ont expliqué 75% de la variation observée. Quatre PCs ont effectivement expliqué 93% de la variation parmi les zones. L'analyse par groupements, a groupé les 37 lignées en cinq groupements réels, pendant que les zones de collection étaient groupées en trois groupements majeurs. Ces données sont utiles pour dans le programme d'amélioration du *tef*.

Mots Clés: Analyse par groupement, analyse du composant principal, *Tef*

INTRODUCTION

Tef (*Eragrostis tef* (Zucc.) Trotter) belongs to the family Poaceae, subfamily Eragrostoideae, tribe Eragrosteae and genus *Eragrostis*. It is a C4, self-pollinated, chasmogamous annual cereal. It has a fibrous root system with mostly erect stems, although some cultivars are bending or elbowing types (Ayele, 1999). The genus contains about 350 species of which *tef* is the only cultivated species. The common name of the crop in Ethiopia is *tef* which seems to be derived from the Amharic word '*tefa*' to mean 'lost' to reflect the small size of the grain which disappears if scattered on the ground (Ketema, 1997).

Several endemic and non-endemic species of *Eragrostis*; some of which are considered to be the wild relatives of *tef*, are found in Ethiopia. Genetic diversity for *tef* exists nowhere else in the world, except in Ethiopia which indicates that *tef* is originated and was domesticated in Ethiopia (Ketema, 1997).

Fifty four *Eragrostis species* are found in Ethiopia, out of which fourteen are known to be endemic. Worldwide, Africa contributes 43% of the genus, while South America contributes 18%. Likewise, 12, 10, 9, 6 and 2% of the genus *Eragrostis* is from Asia, Australia, Central America, North America, and Europe, respectively (Ketema, 1997).

Gojam, Gonder, Wollo, Shewa and Welega are the major *tef* producing areas in Ethiopia. It is widely grown in both high potential and marginal production areas. These areas include most parts of the vertisols that suffer from water logging, and other soil types in different parts of the country that suffer from low moisture stress.

The Central Statistics Agency report (CSA, 2008) indicate the area covered by *tef* to be 2.56 million hectares covering 28.29% of the total land allotted to cereals. In the same report, an annual production of 2.99 million tonnes of *tef* grain was recorded contributing 21.85% of the total cereal produce.

Understanding the pattern and extent of genetic diversity in a population is pivotal to the success of any crop improvement programme. It can provide valuable information for plant breeders who are interested in introgressing

agronomically desirable traits into established cultivars or to select lines from the existing diversity.

To this end, there had been efforts in *tef* germplasm collection and characterisation in the country. Ketema (1997) reported that there were about 3,843 accessions and the need for further collection as there were areas not addressed. According to Demisie (2001), there were 4,395 *tef* accessions, out of which about 1,100 were from Amhara region. The existence of broad variation in many of the phenotypic traits has been demonstrated by several previous studies on *tef* germplasm (Ketema, 1993; Tefera *et al.*, 2003; Asefa *et al.*, 1999, 2001, 2003a; Chanyalew *et al.*, 2009).

Seed storage protein analyses showed polymorphism among *tef* cultivars (Bekele *et al.*, 1995). But Bai *et al.* (2000) reported a relatively low level of DNA polymorphism using RAPD markers. On the other hand, inter-simple sequence repeats (ISSR) analysis showed great diversity among *tef* cultivars (Asefa *et al.*, 2003b). However, these efforts were not all inclusive since Ethiopia has a huge endowment of *tef* genetic diversity. In fact activities to characterise, classify and identify the regional genetic wealth are minimal. Therefore, this research was undertaken to assess the genetic diversity, heritability and genetic gain of *tef* lines from the Amhara region in Ethiopia.

MATERIALS AND METHODS

Experimental site. The field experiment was conducted at Adet Agricultural Research Center in 2010 main cropping season. Adet is located at 37°29' E and 11°16' N in the Amhara National Regional State at an altitude of 2240 metres above sea level. It receives mean annual rain fall of 1230 mm.

Experimental procedures. Landraces from the former province of Gojam, Gondar, and Wollo, were accessed from Debrezeit Agricultural Research Center. Improved varieties from both late (Dz-Cr-387, Dz-01-3186, Dz-01-2423) and early (Dz-Cr-37) sets were included, totaling the test materials to 37 (Table 1).

TABLE 1. *Tef* germplasm lines and their areas of collection in Ethiopia

Germplasm	Zone	District
Dz-01-2580	N/Gondar	Dabat
Dz-01-2721	N/Wolo	Kobo
Dz-01-2613	S/Gondar	D/Tabor
Dz-01-2608	S/Gondar	A/Zemen
Dz-01-2611	S/Gondar	A/Zemen
Dz-01-2572	N/Gondar	A/Arkay
Dz-01-2610	S/Gondar	A/Zemen
Dz-01-2706	E/Gojam	Bichena
Dz-01-2606	S/Gondar	A/Zemen
Dz-01-2582	N/Gondar	Dabat
Dz-01-2631	S/Gondar	Ebinat
Dz-01-2602	S/Gondar	Ebinat
Dz-01-2729	N/Wolo	Kobo
Dz-01-2691	E/Gojam	Bichena
Dz-01-2653	W/Gojam	Mecha
Dz-01-2627	S/Gondar	D/Tabor
Dz-01-2576	N/Gondar	A/Arkay
Dz-01-2654	W/Gojam	Mecha
Dz-01-2569	S/Gondar	A/Arkay
Dz-01-2705	E/Gojam	Bichena
Dz-01-2609	S/Gondar	A/Zemen
Dz-01-2688	E/Gojam	Bichena
Dz-01-2708	E/Gojam	Bichena
Dz-01-2781	N/Wolo	Degem
Dz-01-2629	Gondar	Na
Dz-01-2619	S/Gondar	A/Giorgis
Dz-01-2666	S/Gondar	D/Tabor
Dz-01-2598	W/Gojam	Dembecha
Dz-01-2594	S/Gondar	Ebinat
Dz-01-2686	E/Gojam	Bichena
Dz-01-2614	S/Gondar	D/Tabor
Dz-01-2643	W/Gojam	Adet
Dz-01-2678	E/Gojam	Lumame
Dz-Cr-387	Quncho	Improved
Dz-01-3186	Esub	Improved
Dz-01-2423	Dima	Improved
Dz-Cr-37	Tsedey	Improved

N/ = north; S/ = south, E/ = east, W/ = west, Na/ = not available

Analysis of variance based on the ANOVA skeleton (Gomez and Gomez, 1984)

Source of variation	df	Mean squares	Expected mean squares
Block	(f-1)		
Genotypes(g)	(g-1)	MS _g	$\delta_g^2 + r\delta_e^2$
Error	(g-1)(f-1)	MSe	δ_e^2

Where, MS_g = mean square of genotypes, MSe = mean square of error, δ_e^2 = expected error variance (MSe), δ_g^2 = Genotypic variance component

The experiment was laid out in randomised complete block design with three replications. A plot size of 1 m x 0.2 m was used with 1 m and 2 m inter row and between block distance, respectively. Based on the recommended seed rate of 30 kg ha⁻¹, 0.6 g of seed was broadcasted on the surface of the plots. At tillering, the plants in each plot were thinned to 5 cm plant-to-plant spacing to ease individual plat inspection.

Data collection and analysis. Main tillers of five randomly selected plants in each plot were pre-tagged and evaluated for fourteen quantitative traits (Table 3). Days to heading, and maturity were estimated on plot basis. Prior to the actual data analysis, traits that deviated from normal distribution were transformed using different transformation methods. The first and second inter-node lengths were normalised by square root transformation while; number of nodes per culm, number of productive tillers per culm and shoot biomass yield were log₁₀ transformed. Grain yield was double log₁₀ transformed, all satisfying the requirement of normal distribution.

For all analyses, SAS Version 8 statistical software was used. Analysis of variance was conducted based on the ANOVA skeleton below (Gomez and Gomez, 1984).

The linear model of the experiment was:

$$Y_{ij} = \mu + g_i + r_j + \varepsilon_{ij}$$

Where, Y_{ij} =observed variation, μ =overall mean, g_i =effectof the i^{th} genotype, r_j =effect of the j^{th} replication, and ε_{ij} =experimental error.

The genotypic and phenotypic variance components were estimated based on the following formula (Sharma, 1998).

$$\sigma^2_g = \frac{MS_g - MS_e}{r}, \sigma_e^2 = \text{error mean square}$$

$$(MS_e), \text{ and } \sigma_p^2 = \sigma_g^2 + \sigma_e^2$$

Where, MS_g = mean square due to genotypes, MS_e = environmental variance (error mean square), σ_p^2 = phenotypic variance, and r = the number of replication. Phenotypic (PCV), genotypic (GCV), and environmental (ECV) coefficients of variation were computed as the square roots of the respective variances divided by their mean and converted in to percent.

Heritability in broad sense was computed as the ratio of genotypic variance to phenotypic variance, *i.e.*, $H = GCV/PCV * 100\%$ (Falconer, 1989). Expected genetic advance for each character at 5% selection intensity ($K=2.056$) was computed using the methodology described by Johnson *et al.* (1955). Genetic advance, as percent of mean, was also calculated to compare the extent of predicted genetic gain of different traits under selection.

For multivariate analysis, the mean values of the 37 lines were used. Prior to multivariate analysis, data were standardised to mean zero and variance one to avoid differences in measurement scale. Principal component analysis was computed to reduce the number of variables into a few uncorrelated components that can explain much of the variability. It also helps to identify traits that load most in explaining the observed variation. PCs with eigenvalues greater than unity were considered important in explaining the variability. The PROC PRINCOMP of SAS version 8 (SAS Institute, 2000) was used for principal component analysis. Euclidean distance between each line for mean values of each variable was used as a measure of genetic divergence for clustering.

Cluster analysis was done on the genetic distance matrix of the lines to envision their grouping pattern. The PROC CLUSTER of SAS Version 8 (SAS Institute, 2000) using Unweighted Pair Group Method using Arithmetic Average linkage (UWPGMA) was employed. The number of clusters was determined based on the pseudo F and t^2 values.

RESULTS

Hierarchical analysis of variance for the different zones showed significant differences for different traits (Table 2). In accessions from east Gojam, highly significant ($p < 0.01$) differences were observed for phenological traits, *i.e.*, days to heading, days to maturity and grain filling period. Significant differences ($P < 0.05$) were recorded for plant height, peduncle length and culm length among these genotypes. However, yield and yield related traits like number of productive tillers per plant, panicle length, biomass yield, grain yield and harvest index were not significantly different for these genotypes (Table 2). But for accessions from west Gojam, differences in biomass yield, grain yield, harvest index, days to heading and days to maturity were highly significant ($P < 0.01$).

Genotypes from south Gondar were significantly different for most of the traits (Table 2). Differences in days to heading, days to maturity and grain filling period, plant height, panicle length, culm length, and peduncle length were highly significant ($P < 0.01$). But yield related traits like biomass yield, grain yield and grain filling period were not significantly different. Unlike collections from Gojam, differences in number of nodes per culm, number of productive tillers per plant, and second inter-node length were highly significant in south Gondar.

Improved varieties were significantly different for days to heading and days to maturity (Table 2). Highly significant differences were also recorded for plant height, panicle length, peduncle length, and culm length. But these varieties were not significantly different for biomass yield and grain yield. For days to heading, Quncho was significantly different from the rest of the varieties except Etsub. Tsedey was the earliest of the varieties confirming earlier findings (standard check for earliness) in this zone. The same scenario was observed for days to maturity, except that the latest maturing was Etsub, though not significantly different from Quncho. Based on plant height, the four varieties were grouped into three; Quncho being the longest followed by Etsub. There were significant

TABLE 2. Mean values, variances and variation coefficients of the 37 *tef* lines in Ethiopia

	DH	DM	GFP	PH	PL	SIL	FIL	PDL	CL	NNP	NPT	BMY	GY	HI
Mean	62.86	112.1	49.23	93.33	40.15	9.71	4.99	21.80	31.39	3.49	7.37	19.54	4.64	0.23
MST	83.7**	54.7**	64.9**	214**	67.5**	22.5**	62.3**	0.01**	0.09**	0.09**	0.035*	0.19**	0.08**	0.08 ^{ns}
MSE	6.166	8.315	11.18	33.357	14.213	5.305	7.01	0.002	0.02	0.018	0.019	0.094	0.034	0.059
CV (%)	3.94	2.57	6.79	6.19	9.39	10.57	8.43	19.6	11.49	11.41	16.92	8.79	5.89	10.93

DH = days to heading, DM = days to maturity, GFP = grain filling period, PH = plant height, PL = panicle length, SIL = second internode length, FIL = first internode length, PDL = peduncle length, CL = culm length, NNP = number of nodes per culm, NPT = number of productive tillers per plant, BMY = biomass yield, GY = grain yield, HI = harvest index
*, **, and ns indicate significant (5%), highly significant (1%), and not significant, respectively

differences between Quncho and the rest of the varieties, and also Etsub and the remaining two varieties for culm length. But for biomass yield grain yield and number of productive tillers per plant, differences were only numeric. Generally, there was strong evidence that the 37 accessions were highly variable for all the quantitative traits except first inter-node length.

Estimation of genetic parameters. Phenotypic coefficients of variation ranged from 4.35 (days to heading) to 24.5 % (harvest index) (Table 3). Relatively higher PCV values were observed for number of productive tillers per plant, grain yield per plant, and biomass yield per plant. Moderate to high PCV values were also observed for panicle, peduncle and culm length. PCV values were lower for days to heading, days to maturity and second inter-node length than the rest of the traits. Harvest index showed the highest GCV, while days to maturity was the one with the lowest GVC value. Environmental coefficients of variation ranged from 2.6% (days to maturity) to 19.3% for harvest index. Comparable ECV values (14.4%) were observed for grain yield and shoot biomass yield.

Heritability in broad sense was highest for days to heading, followed by culm length (Table 3). Generally, most of the traits showed moderate to high heritability values. Grain yield and shoot biomass yield showed heritability values of 54.6% and 57.3%, respectively. Grain filling period, plant height, and days to maturity also had high heritability values. Genetic gain was also divergent across the test materials. Culm length showed the largest GMA (Table 3) value; while first inter-node length was the lowest.

Principal component analysis. Of the total variation, 81% was explained by the first four PCs out of the fourteen PCs employed. Eigen values were greater than one only for the first three PCs, which together explained 75% of the observed variation. Therefore, three PCs were retained to explain the observed variation without losing a substantial variability to be explained. The first PC explained about 35%, the second 23%, and the third explained 17% of the variation.

Grain filling period, culm length, number of nodes per culm, number of productive tillers per

TABLE 3. Root mean square, treatment mean squares, environmental, genotypic, and phenotypic variance, phenotypic, genotypic, and environmental coefficients of variation, heritability in broad sense, expected genetic gain, and expected genetic gain as percent of mean of the 14 traits

VAR	RMSE	MST	EV	GV	PV	PCV	GCV	ECV	H	GA	GAM
DH	2.48	83.650	6.166	25.828	31.994	9.024	8.11	3.962	80.73	9.39	14.98
DM	2.88	54.697	8.315	15.461	23.776	4.350	3.51	2.573	65.03	6.52	5.82
GFP	3.34	64.982	11.180	17.934	29.114	10.961	8.60	6.793	61.60	6.83	13.88
PH	5.77	213.826	33.357	60.156	93.514	10.361	8.31	6.188	64.33	12.79	13.70
PL	3.77	67.490	14.213	17.759	31.972	14.083	10.50	9.389	55.55	6.46	16.08
FIL	0.24	0.081	0.059	0.007	0.066	11.598	3.88	10.930	11.19	0.06	2.67
SIL	0.18	0.084	0.034	0.017	0.050	7.213	4.15	5.897	33.16	0.15	4.92
PDNL	2.30	22.475	5.305	5.723	11.028	15.235	10.98	10.567	51.90	3.54	16.26
CL	2.65	62.270	7.010	18.420	25.430	16.070	13.68	8.437	72.43	7.51	23.93
NPT	0.14	0.035	0.020	0.005	0.025	18.999	8.58	16.953	20.38	0.07	7.96
HI	0.44	0.005	0.002	0.001	0.003	24.503	15.11	19.287	38.04	0.04	19.17
MBY	0.14	0.092	0.020	0.024	0.044	16.923	12.50	11.405	54.58	0.24	18.99
GY	0.13	0.089	0.018	0.024	0.041	17.479	13.23	11.420	57.32	0.24	20.60
NNP	0.36	0.194	0.130	0.021	0.151	11.121	4.18	10.307	14.10	0.11	3.22

*DH=days to heading, DM days to maturity, GFP=grain filling period, PH=plant height, PL=panicle length, PDL=peduncle length, CL=culm length, BMY=biomass yield, GY=grain yield, HI=harvest index, NPT=number of productive tillers per plant, NNP=number of nodes per culm, FIL=first internode length, SIL=second internode length

plant and harvest index showed greater loadings in the first PC (Table 4). Similarly, days to maturity, plant height, panicle length, biomass yield and grain yield contributed in the second PC; while days to heading, first inter-node length, second-inter node length and peduncle length were significant in the third PC.

Principal components analysis on the zonal average data was effective in that only the first four PCs were able to explain about 93% of the total variation. The first PC explained 40%, the second 33%, and the third 12% of the variations observed among the different zones of collection and the two sets (early and late sets) of improved cultivars.

Grain filling period, culm length, number of productive tillers per plant, biomass yield and grain yield showed greater loadings in the first PC (Table 4). Similarly, days to heading, days to maturity, plant height, panicle length, and peduncle length, were important in the second PC; while the second inter-node length, number of nodes per culm and harvest index were significant in the third PC. In the fourth PC, peduncle length and number of productive tillers were important traits.

Cluster analysis. Average linkage, single linkage, and complete linkage were tried, but clearer and more understandable clustering was achieved by average linkage clustering (Fig. 1). The number of clusters was determined based on the pseudo F and t^2 values which classified the test materials in to five real clusters each being able to be classified further. The number of accessions in each cluster varied from two in cluster five to eighteen in cluster one.

Cluster one consisted of 18 genotypes, which were mixes of lines from four zones. Ten of them were from South Gondar, 3 from north Gondar, 2 from north Wollo and 2 from east Gojam (Table 5). Similarly, cluster two was composed of lines from four zones and two improved varieties, while cluster three had 4 lines from Gojam, 2 lines from south Gondar, and the other 2 improved varieties. Members of the fourth cluster were only from east Gojam, while that of cluster five were from north Gondar and west Gojam, one from each. Cluster one was more heterogeneous than others. Lines in cluster two were the earliest, the shortest; in plant height panicle length, both first and second inter-node lengths; and culm length, and the least yielding ones in biomass. Cluster

TABLE 4. Eigen values, total variance, cumulative variance, and eigenvectors for 14 quantitative characters in 37 *tef* genotypes and, zones of collection

Characters	Eigenvectors for 37 lines				Eigenvectors for zones			
	Prin 1	Prin 2	Prin 3	Prin 4	Prin 1	Prin 2	Prin 3	Prin 4
Days to heading	-0.29	0.26	-0.32	0.01	-0.15	0.36	-0.34	-0.19
Days to maturity	0.03	0.38	-0.24	0.13	0.21	0.34	-0.28	-0.01
Grain filling period	0.36	0.06	0.14	0.11	0.39	0.10	-0.05	0.15
Plant height	-0.26	0.38	0.26	0.17	-0.24	0.37	0.12	-0.01
Panicle length	-0.18	0.43	0.06	-0.14	-0.22	0.37	0.12	-0.16
1 st internode length	-0.05	0.17	0.53	-0.28	-0.20	0.17	0.59	0.26
2 nd internode length	-0.14	-0.12	0.44	-0.37	-0.27	-0.22	0.02	0.18
Peduncle length	0.13	0.02	0.38	0.77	0.09	0.28	-0.15	0.65
Culm length	-0.36	0.24	0.19	0.00	-0.32	0.27	0.20	-0.13
Number of nodes/ culm	-0.36	0.13	-0.16	0.25	-0.25	0.29	-0.32	-0.15
Productive tillers/plant	0.30	0.22	-0.21	-0.18	0.35	0.06	-0.05	-0.41
Biomass yield/plant	0.29	0.39	-0.03	-0.08	0.31	0.29	0.11	0.16
Grain yield/plant	0.33	0.35	0.02	-0.11	0.33	0.26	0.18	0.11
Harvest index	0.33	0.14	0.16	0.03	0.27	0.09	0.45	-0.40
Eigenvalue	4.83	3.28	2.37	0.86	5.60	4.60	1.60	1.10
% of total variation	0.35	0.23	0.17	0.06	0.40	0.33	0.12	0.08
% cumulative variance	0.35	0.58	0.75	0.81	0.40	0.73	0.85	0.93

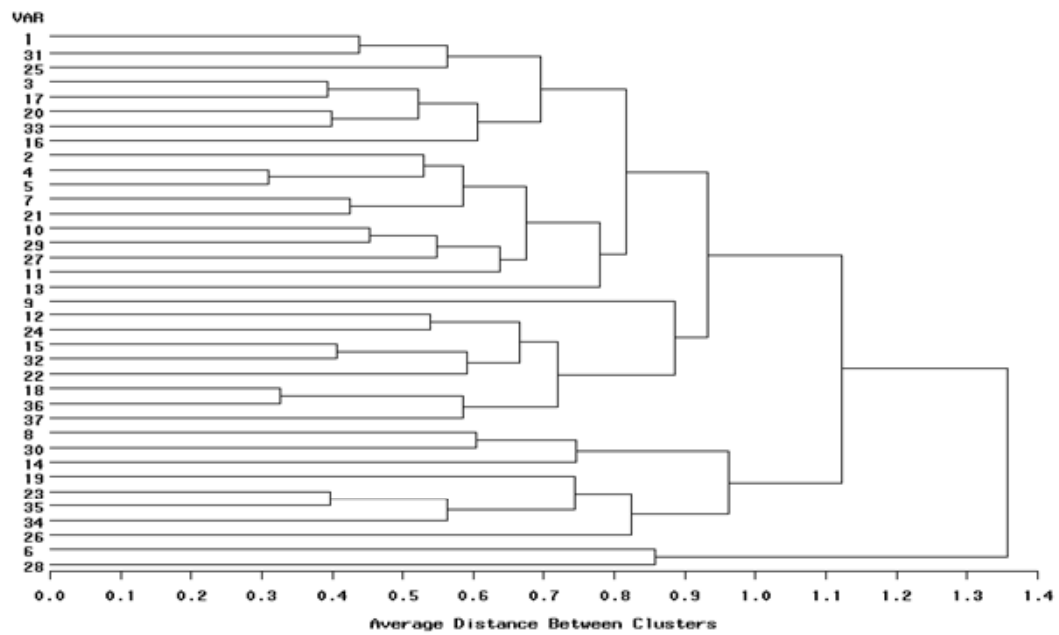


Figure 1. Dendrogram of 37 genotypes based on Euclidean distance and average linkage clustering.

TABLE 5. Distribution of 37 tef genotypes over five clusters by zone of origin, and improved varieties

Zone/variety	Cluster					Total
	1	2	3	4	5	
E/Gojam	2	1	1	3	-	7
W/Gojam	-	3	-	-	1	4
S/Gondar	10	2	2	-	-	14
N/Gondar	3	-	-	-	1	4
N/Wolo	2	1	-	-	-	3
Gondar	1	-	-	-	-	1
Dima	-	1	-	-	-	1
Tsedey	-	1	-	-	-	1
Etsub	-	-	1	-	-	1
Quncho	-	-	1	-	-	1
Total	18	9	5	3	2	37

E = East, W = West, S = South, N = North

three consisted of 5 lines in which the longest average peduncle length was recorded. The period required to fill was also shortest in this cluster. Average plant height was highest for lines in cluster three.

The latest lines, both in days to heading and days to maturity were clustered in cluster five (Table 6). Culm length and number of nodes per culm were also highest in this cluster. Low grain yield and harvest index were characteristics of the cluster four. Of all clusters, cluster five consisted of lines that were high yielder both in biomass and grain yield. The number of productive tillers was also highest in this cluster, and conformed with the recorded biomass and grain yield; whereas the number of nodes per culm was the lowest.

Inter-cluster distance. The longest genetic distance was measured between clusters four and five while the shortest one was measured between clusters three and four (Table 7). All traits studied were highly significantly different, except for first inter-node length.

Cluster analysis on the different germplasm sources, *i.e.*, the five zones of collection, and improved varieties was also done based on the correlation matrix and pair wise Euclidean distances among the entities. As a result three major clusters were determined based on the

TABLE 6. Mean values of the 14 quantitative traits in the five clusters

Cluster	N	DH	DM	GFP	PH	PL	SIL	FIL	PDL	CL	NNP	NPT	BMV	GY	HI
1	18	62.39	113.89	51.50	93.62	40.42	9.50	4.80	22.36	30.84	3.51	7.82	22.65	5.67	0.25
2	9	59.11	107.33	48.22	84.12	34.96	9.18	5.07	21.47	27.68	3.33	6.86	12.59	2.66	0.21
3	5	65.87	109.40	43.54	104.65	44.83	10.80	5.27	22.73	36.50	3.61	5.63	15.40	3.22	0.21
4	3	73.11	117.56	44.44	101.97	45.38	9.69	5.73	17.33	39.26	3.95	5.67	13.19	2.00	0.15
5	2	61.17	115.84	54.67	90.95	41.52	10.82	4.53	21.07	28.37	2.99	12.57	42.74	11.80	0.28

DH = days to heading, DM = days to maturity, GFP = grain filling period, PH = plant height, PL = panicle length, SIL = second internode length, FIL = first internode length, PDL = peduncle length, CL = culm length, NNP = number of nodes per culm, NPT = number of productive tillers per plant, BMV = biomass yield, GY = grain yield, HI = harvest index

TABLE 7. Dissimilarity matrix calculated using Euclidean distance measure among the six clusters

CN	1	2	3	4	5
1	0.0				
2	17.5	0.0			
3	18.3	26.1	0.0		
4	21.9	29.9	13.1	0.0	
5	22.3	35.2	36.4	39.2	0.0

pseudo F and t^2 test criterion. The first cluster consisted of four zones of germplasm collection namely south Gondar, north Gondar, north Wollo and one accession from Gondar whose exact zone of collection was not available.

The second and third clusters consisted of three entities each. In the second cluster were the improved varieties Quncho, and Etsub, and the mean values for east Gojam, while the last cluster was composed of the improved varieties, Dima and Tsedey, and the mean value of lines from West Gojam.

DISCUSSION

There was a large diversity for all traits, except for the first inter-node length (Table 2); but this variability was not uniform across zone of collection indicating the possibility of selection for different traits in different zones. For instance, lines from east Gojam were highly variable for plant height, panicle length and for phenological traits; while those from west Gojam were significantly different for biomass and grain yield. However, in this particular study the sample size was not uniform across zones of collection, so a different scenario may be observed especially for those claimed non-significant if samples are increased and equally sized to validate the zonal comparison. The present study is in agreement with many of previous findings in *tef* germplasm (Ketema 1993; Tefera *et al.*, 2003; Asefa *et al.*, 1999, 2001, 2003a; Chanyalew *et al.*, 2009).

Improved varieties were significantly different for days to heading and days to maturity (Table 2). This was however, predictable in that improved lines were included from both late and medium maturing varieties. Highly significant differences were also recorded for plant height,

panicle length, peduncle length and culm length. But these varieties were not significantly different for biomass yield and grain yield. This is also true because improved lines are highly selected especially for grain yield so that much of the variability is lost through selection.

Relatively higher GCV (12.5-15.11) values were observed in harvest index, grain and biomass yield and culm length; while days to heading, grain filling period, plant height, panicle length and number of productive tillers per plant showed intermediate GCV values. Similar findings with relatively higher GCV values were reported by Tefera *et al.* (2003), Asefa *et al.* (1999) and Chanyalew *et al.* (2009). Unlike these findings, the present study showed highest GCV in harvest index, which was intermediate in all those studied. Discrepancies between GCV and PCV was high for harvest indices and number of productive tillers per plant (Table 3), indicating the high proportion of ECV which is the contribution of non-additive gene effects. Therefore, selection for these traits may not be proportional to the observed variation as the proportion of additive gene effect is low. This was manifested by the low heritability value for these traits, though heritability alone does not determine the predicted improvement of a trait.

Days to heading showed the highest heritability value followed by culm length (Table 3). Grain and shoot biomass yield, grains filling period, plant height, and days to maturity had high heritability values. However, this does not mean that these values guarantee success in selection because resemblance between relatives is controlled by the proportion of the additive genes, not by all of the genetic variation (Falconer and Mackay, 1996). Asefa *et al.* (2001) reported similar kind of heritability values for days to heading and peduncle length. But, in this particular report heritability value for grain yield and biomass yield were lower than the ones recorded in the present study. This confirms that heritability values vary with the nature of the test materials and the area where the experiment is conducted.

Heritability, coupled with genetic gain at specified selection intensity, is a reliable tool to guide selection. As such, in this study, a change in rank was observed among lines with high

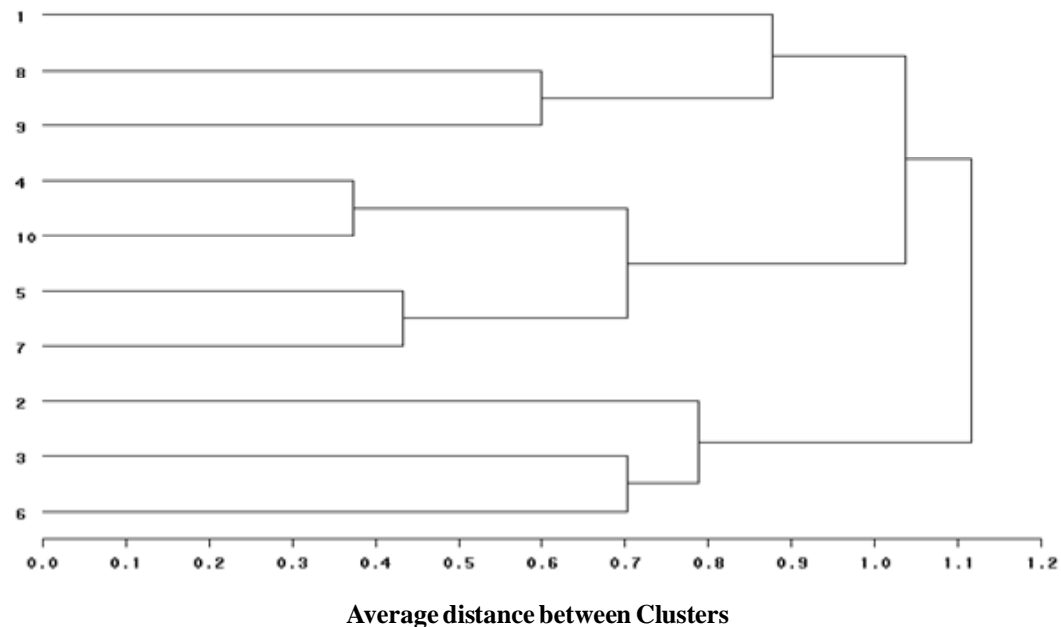
heritability values, rendering only small genetic gain in selection. High GA values were recorded in culm length, grain yield and harvest index different from the rank for heritability values. A trend similar was reported by Asefa *et al.* (1999). Generally, lines in this study showed figures of significant magnitude pointing to the possibility of screening promising lines for yield.

Principal component analysis reduced the 14 variables (characters) to three uncorrelated components (principal components) each having eigenvalues greater than one. The first PC explained about 35% of the variation, which was mainly due to variations in grain filling period, culm length, number of nodes per culm, number of productive tillers per plant and harvest index (Table 4). Similarly, days to maturity, plant height, panicle length, biomass yield and grain yield had greater loadings (contribution) in the second (23%) PC; while days to heading, first inter node-length, second inter-node length and peduncle length were significant in the third principal

component. The result of PCA is also in agreement with variance analysis results in that it was only 75% of the total variation that was explained by the first three PCs. This indicates the complexity in variability in these lines. The first five PCs were reported to explain 79% of variation in sorghum (*Sorghum bicolor*) germplasm by Ayana *et al.* (1999); while in mustard (*Brassica carinata*) 96.01% of variation was explained by the first five PCs each having eigenvalues greater than unity (Genet *et al.*, 2005).

Principal components analysis on the zonal average data and the four improved varieties was effective in that only the first four PCs were able to explain about 93% of the total variation. The first PC explained 40%, the second 33% and the third 12% of the variations observed among the different zones of collection and the two sets (early and late sets) of improved cultivars. Similarly, the first four PCs explained 93.9% of regional variation in *tef* germplasm (Asefa *et al.*, 2003a).

Zone



*Dima = 1, E/Gojam = 2, Etsub = 3, N/Gondar = 4, N Wolo = 5, Qunch = 6, S/Gondar = 7, Tseday = 8, W/Gojam = 9, Gondar = 10

Figure 2. Dendrogram of zones of collection and improved varieties.

Unweighted Pair Group Method using Arithmetic Average linkage (UWPGMA) grouped the 37 *tef* lines in to five real clusters each consisting 2 to 18 lines (Fig. 1). The number of clusters was determined based on the pseudo F and t^2 values such that the pseudo F reaches its pick and at the same time it is larger than values before and after it in the list, while the pseudo t^2 being at its minimum. This procedure also fulfills the way of determining the number of clusters such that each cluster has to have at least two members (Ayana *et al.*, 1999).

Cluster one was the most heterogeneous consisting of lines from all the five zones, except west Gojam (Table 5). Ten of the 14 lines from south Gondar were grouped in cluster one in which two lines from east Gojam and north Wollo each, and three from north Gondar were included. A line from Gondar was also included in this cluster, though its zone of collection was not known. Short first inter-node length was the distinguishing feature of this cluster. This cluster was second to cluster five in biomass yield, grain yield and harvest index. Generally, clustering does not strictly apply to the zones of collection in that lines were mixed in all clusters except the third cluster, which solely contained lines from east Gojam. Similar findings were reported by Ayana *et al.* (1999) on sorghum and Asefa *et al.* (2003a) on *tef* accessions in that a clear cut differentiation of lines in to their region of origin was not apparent. However, the improved varieties Dima and Tsedey were in a group, while Quncho and Etsub were grouped together in another cluster. Cluster two which encompassed Dima and Tsedey, had nine lines which were aggregates from all zones of collection, except north Gondar. This cluster was characterised by the earliest, the shortest; in plant height panicle length, both first and second inter-node lengths and culm length. They were the least yielding ones in biomass. This indicates the possibility of developing lines adapted to drought prone areas or where early cessation of rainfall is prevalent. Early maturing lines can also be developed to fit a double cropping system.

CONCLUSION

This study has revealed the presence of large genetic diversity both within and among zones of *Tef* genotypes collection. It also indicates the need for giving equal emphasis to all zones during collection as they are diverse in the germplasm pool they house. Hybridisation can also be done among lines from divergent clusters to create heterotic lines and to enrich the existing diversity.

ACKNOWLEDGEMENT

Adet Agricultural Research Centre is acknowledged for financing this study. Debrezeit Agricultural Research Centre provided landraces; while the improved varieties were accessed from Adet Agricultural Research Centre, *tef* breeding project.

REFERENCES

- Assefa, K., Merker, A. and Tefera, H. 2003a. Multivariate analysis of diversity of *tef* [*Eragrostis tef* (Zucc.) Trotter] germplasm from western and southern Ethiopia. *Hereditas* 138:228-236.
- Assefa, K., Merker, A. and Tefera, H. 2003b. Inter simple sequence repeat (ISSR) analysis of genetic diversity in *tef* [*Eragrostis tef* (Zucc.) Trotter] *Hereditas* 139:174-183.
- Assefa, K., Tefera, H., Merker, A., Kefyalew, T. and Hundra, F. 2001. Variability, heritability and genetic advance in phen-omorph and agronomic traits of *tef* [*Eragrostis tef* (Zucc.) Trotter] germplasm from eight regions of Ethiopia. *Hereditas* 134:103-113.
- Assefa, K., Yu, J.-K., Zeid, M. Belay, G., Tefera, H. and Sorrells, M.E. 2010. Review: Breeding *tef* [*Eragrostis tef* (Zucc.) Trotter]: conventional and molecular approaches. *Plant Breeding*. doi:10.1111/j.1439-0523.2010.01782.x.
- Assefa, K., Ketema, S., Tefera, H. Nguyen, H.T., Blum, A., Ayele, M., Bai, G., Simane, B. and Kefyalew, T. 1999. Diversity among

- germplasm lines of the Ethiopian cereal tef [*Eragrostis tef* (Zucc.)Trotter]. *Euphytica* 106:87-97.
- Ayana, A. and Bekele, E. 1999. Multivariate analysis of morphological variation in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea. *Genetic Resources and Crop Evolution* 46:273-284.
- Ayele, M. 1999. Genetic diversity in tef (*Eragrostis tef* (Zucc.) Trotter) for osmotic adjustment, root traits, and Amplified Fragment Length Polymorphism. PhD Thesis, Texas Tech University, USA.
- Bai, G., Ayele, M., Tefera, H. and Nguyen, H.T. 2000. Genetic diversity in tef [*Eragrostis tef* (Zucc.) Trotter] and its relatives as revealed by random amplified polymorphic DNAs. *Euphytica* 112:15-22.
- Bekele, E., Fido, R.J., Tatham, A.S. and Shewry, P.R. 1995. Heterogeneity and polymorphism of seed proteins in tef (*Eragrostis tef*). *Hereditas* 122:67-72.
- Central Statistical Authority (CSA), 2008. The Federal Democratic Republic of Ethiopia Central Statistical Agency Agricultural Sample Survey (2007/2008) (September - December 2007) Volume I . Report on Area and Production of Crops (Private Peasant Holdings, Meher Season). Addis Ababa, Ethiopia.
- Chanyalew, S., Tefera, H. and Singh, H. 2009. Genetic variability, heritability, and trait relationships in recombinant inbred lines of tef (*Eragrostis tef* (Zucc.) Trotter). *Research Journal of Agriculture and Biological Sciences* 5(4):474-479.
- Demisie, A. 2001. *Tef* genetic resources in Ethiopia. pp. 27-31. In: Tefera, H., Belay, G. and Sorrells, M. (Eds). *Narrowing the Rift: Tef Research and Development*. 16-19 October 2000. Debre Zeit, Ethiopia.
- Falconer, D.S. and F.C. Mackay, 1996. *Introduction to Quantitative Genetics*. 4th Edition. Long man group Ltd. London, England. 464p
- Falconer, D.S. 1989. *Introduction to Quantitative Genetics*. 3rd Edition. Long Man group Ltd., England. 438p
- Genet, T., Labuschagne, M.T. and Hugo, A. 2005. Genetic relationships among Ethiopian mustard genotypes based on oil content and fatty acid composition. *African Journal of Biotechnology* 4(11):1256-1268.
- Gomez, K.A. and Gomez, R.A. 1984. *Statistical procedures for Agricultural Research*. 2nd Edition. John Wiley and Sons Inc. New York, Chichester, Brisbane, Toronto, Singapore.
- Johanson, H.W., Robinson, H.F. and Comstock, R.E. 1955. Estimation of genetic and environmental variability in soybeans. *Agronomy Journal* 47: 314 - 318.
- Ketema, S. 1993. Tef (*Eragrostis tef*) breeding, genetic resources, agronomy, utilization and role in Ethiopian agriculture. Institute of Agricultural Research, Addis Ababa, Ethiopia.
- Ketema, S. 1997. Tef. [*Eragrostis tef* (Zucc.) Trotter.] promoting the conservation and use of underutilized and neglected crops. Institute of Plant Genetics and Crop Plant Research, Gatersleben/ International Plant Genetic Resources Institute, Rome, Italy.
- Mohammadi, S.A. and Prasanna, B.M. 2003. Analysis of Genetic Diversity in Crop Plants-Salient Statistical Tools and Considerations. Review and interpretation. *Crop Science* 43:1235-1248.
- SAS Institute, 2000. SAS/STAT user's guide. SAS Institute Inc., Cary, North Carolina, U.S.A.
- Sharma, J.R. 1998. *Statistical and Biometrical Techniques in Plant Breeding*. New age int. publ. New Delhi, India.
- Tefera, H., Assefa, K., Hundura, F., Kefyalew, T. and Tefera, T. 2003. Heritability and genetic advance in recombinant inbred lines of tef [*Eragrostis tef*]. *Euphytica* 13:91-96.
- Wondimu, A. and Tekeba, F. 2001. Utilization of Tef in Ethiopian Diet. pp. 239-244. In: *Narrowing the Rift: Tef Research and Development*. Tefera, H., Belay, G. and Sorrells, M. (Eds.). 16-19 October 2000. Debre Zeit, Ethiopia.