Classifying Ethiopan Tetraploid Wheat (*Triticum turgidum* L.) Landraces by Combined Analysis of Molecular and Phenotypic Data

Negash Geleta[1*] and Heinrich Grausgruber[2]

[1]Wollega University, Department of Plant Sciences, P. O. Box 395,

[2]BOKU- University of Natural Resources and Applied Life Sciences, Department of Applied Plant Sciences and Plant Biotechnology, Institute of Agronomy and Plant Breeding, Vienna, Austria, A-1180;

*Corresponding Author

Negash Geleta

e-mail: ayananegash@yahoo.com

Abstract

The aim of the study was to investigate the extent of the genetic diversity among genebank accessions of Ethiopian tetraploid wheat (*Triticum turgidum* **L.)** using microsatellite markers, qualitative and quantitative data. Thirty-five accessions of Ethiopian tetraploid wheat (*T. turgidum* L.) landraces were grown in the greenhouse at IFA Tulln, Austria during spring 2009 for DNA extraction. The same accessions were already grown in spring 2008 at BOKU Vienna, Austraia for their phenotypical characterisation. DNA was extracted from each approximately one month old plant according to Promega (1998/99) protocol. A total of 10 µl reaction mixture per sample was used for DNA amplification by PCR. The amplified mixture was loaded to PAGE (12%) containing TE buffer (1×) in CBS electrophoresis chambers and run in an electric field for 2 hrs. The fragments were visualized by scanning with Typhoon Trio scanner. Six and ten quantitative and qualitative morphological traits data respectively were used for combined analysis. Genetic variation was significant within and between wheat species and within and between altitudes of collection site. Genetic distances ranged from 0.21 to 0.73 for all accessions while it ranged from 0.44 within *Triticum polonicum* to 0.56 between *T. polonicum* and *T. turgidum*. Genetic distance between regions of collection ranged from 0.51 to 0.54 while for altitudes it ranged from 0.47 (≤2200 m) to 0.56 (≤2500 m). Cluster analysis showed that *T. polonicum* accessions were grouped together whereas *T. durum* and *T. turgidum* formed mixed clusters indicating *T. polonicum* as genetically more distinct from the other two species. We suggest combined analysis of molecular and morphological data for a better classification of accessions.

Keywords: Cluster analysis, Gower distance, microsatellite marker, *Triticum*

INTRODUCTION

Microsatellites are tandemly repeated short DNA sequences that are favoured as molecular-genetic markers due to their high polymorphism index (Mun *et al*., 2006). Tandem repeat in DNA is a sequence of two or more contiguous, approximate copies of a pattern of nucleotides and tandem repeats occur in the genomes of both eukaryotic and prokaryotic organisms (Sokol *et al*., 2006). Microsatellite markers are the best DNA markers so far used for genetic diversity studies and fingerprinting of crop varieties. Microsatellites motifs are conserved in species and their unique behaviour abundance, co-dominance, robustness and easiness for PCR screening make them the best DNA markers for the evaluation of crop genetic diversity. Furthermore microsatellite markers have many advantages for tracing pedigrees because they represent single loci and avoid the problems associated with multiple banding patterns obtained with other marker systems (Powell *et al*., 1996). However, developing microsatellite markers for a plant species requires prior knowledge of its genomic sequences, lack of which makes this technology very expensive and time consuming (Yu *et al*., 2009).

MATERIALS AND METHODS

Plant Material

Thirty-five accessions of Ethiopian tetraploid wheat (*T. turgidum* L.) landraces (Table 1) were grown in the greenhouse at IFA Tulln, Austria during spring 2009. Ten seeds per accession were planted in order to have enough plants per accession for DNA extraction. The same accessions were already grown in spring 2008 at BOKU Vienna, Austria for their phenotypical characterisation.

DNA extraction

DNA was extracted according to Promega (1998/99) protocol. DNA was extracted from each approximately one month old plant. Ten to fifteen centimetres long young leaves were taken and chopped in 2-ml Eppendorf tubes (Eppendorf AG, Hamburg, Germany) and left open to dry for four days in plastic bags containing silica gel. The dried leaves were grounded and leaf tissues were lysed by adding 600 µl of nucleic lysing solution to each of the tubes. The tubes were vortexed for 1-3 minutes to wet the cell uniformly and incubated in hot water at 65°C for 15 min. Ribonucleic acids (RNAs) were dissolved by adding 3 µl (4 mg ml^{-1}) RNase solution. Mixing was done by inverting the tubes 2-5 times. The mixture was incubated at 37°C for 15 min and then cooled at room temperature. 200 µl protein precipitation solution was added to each sample and vortexed vigorously for 20 s and then centrifuged for 3 min at 16000× g. The precipitated proteins formed a tight pellet. The supernatant was carefully removed and transferred to another new 1.5 µl micro centrifuge tube containing 600 µl room tempered isopropanol. The solution was gently mixed for each sample by inversion until a thread like mass of DNA strand was visible. Then the mixture was centrifuged at 16000× g for 2 min at room temperature. The supernatant was carefully decanted for each sample. 600 µl of room tempered ethanol (70%) was added and the tubes were gently inverted several times to wash the DNA and then centrifuged at 16000× g for 2 min at room temperature. The ethanol was carefully decanted and the tube containing the sample was inverted on clean absorbent paper and the pellet was air dried for 15-20 min. 100 µl TE buffer solution was added to re-hydrate the DNA and incubated at 65°C for 1 hr. For subsequent use of DNA in PCR, it was diluted by 1:50 (v/v) DNA/dH_2O.

Polymerase chain reaction (PCR)

A total of 10 µl reaction mixture per sample was used for DNA amplification by PCR. The 10 µl PCR mixture contained 0.025 µl forward primers (10 µM), 0.25 µl (10

µM) reverse primers, 0.225 µl of fluorescent M-13 labelled tail of 10 µM (HEX or FAM), 5 µl GoTaq®Green master mix (Promega Corporation, Madison, USA) (a, b), and 1.2 µl dH_2O. GoTaq®Green master mix (a, b) contains dNTPs (dATP, DGTP, dCTP and dTTP), $MgCl_2$ and reaction buffers at optimal concentrations for efficient amplification of DNA templates by PCR. GoTaq®Green master mix (a, b) (Flanagan et al. 2005) is a premixed ready to use solution containing a non-recombinant modified form of TaqDNA polymerase that lacks 5′→3′exonuclease activity. It also contains two dyes (blue and yellow) that allow monitoring of progress during electrophoresis. PCR program SSR M13 was used for amplification. The following temperatures and times were used for PCR amplification of genomic DNA: (1) 95°C for 2 min (to heat the lid); (2) 95°C for 45 s to denature the double stranded DNA; (3) 68°C for 45 s to anneal the primers to the single stranded DNA; (4) 72°C for 1 min for TaqDNA polymerase to extend the primers. Steps 2 to 4 were repeated for 7 times; (5) 95°C for 45 s to denature the DNA; (6) 54°C for 45 s to anneal the primers to the single stranded DNA; (6) 72°C for 1 min for TaqDNA polymerase to extend the primer ends and steps 5 to 6 were repeated 30 times; (7) further extension of primers was done at 72°C for 5 min by TaqDNA polymerase; (8) finally the reaction was stopped and cooled at 8°C.

Polyacrylamide gel electrophoresis (PAGE) and scanning

The amplified mixture was loaded to PAGE (12%) containing TE buffer (1×) in CBS electrophoresis chambers (C.B.S. Scientific Co., Del Mar, USA) and run in an electric field for 2 hrs. The fragments were visualized by scanning with Typhoon Trio scanner (GE Healthcare Europe GmbH, Regional Office Austria, Vienna).

Microsatellite loci

Microsatellite loci were selected based on available information. Out of 30 micro satellite loci only 11 of them gave polymorphic bands that can be scored as either 0

or 1. However, the microsatellite markers *Xgwm181* and *Xgwm340* are located on the same chromosome arm, i.e. 3BL, very near to each other (Röder et al. 1998). Hence, only fragments from *Xgwm340* were considered for the analysis. Chinese Spring wheat was used as size standard marker. The microsatellite primers are presented in Table 2.

Molecular and phenotypic data

Data from the 10 microsatellite markers were recorded in a binary way (0 or 1). Zero means no allele for the locus while 1 means there is an allele. In total 42 alleles were present. Quantitative data of six morphological traits, i.e. days to heading, spike density, awn length, thousand kernel weight, yellow pigment content and protein content which were used for the combined analysis. Furthermore ten qualitative traits included beak shape, beak length, glume colour, awn color, glume hairiness, seed color, seed size, seed shape, vitreousness and seed plumpness were used.

Statistical Analysis

Gene diversity among accessions for microsatellite markers was calculated according to Nei (1973):

$$Gene\,diversity = 1 - \sum P_{ij}^2 \ ,$$

where P_{ij} is the frequency of the j^{th} allele for the i^{th} locus summed across all alleles of the locus. The gene diversity coefficient is also referred to as the allelic polymorphic information content according to Anderson *et al*. (1993). Data from SSR marker, qualitative and quantitative traits were combined and analysed modified after Franco *et al*. (1997a). Regions with only a few number of accessions were pooled together and four groups were formed, i.e. Northern (Eritrea, Tigray, Welo, Gonder, Gojam), Central (Shewa) and Southern (Arsi, Kefa, Gamu Gofa) Ethiopia. Accessions with no available information of their original collection site were pooled

together in one group. Similarly, altitudes of collection sites were classified as ≤2200 m, ≤2500 m, ≤2800 m, >2800 m and genotypes with no available information. Genetic distances between accessions, within and between species, within and between regions, and within and between altitudes were computed using Gower's distance (Gower, 1971). Using the dissimilarity distances between accessions a GLM analysis of variance was run for species, regions and altitudes to check significances between these effects and in order to obtain means and standard errors. Hierarchical cluster analysis was performed for all genotypes using the dissimilarity matrix of Gower's distance and the Ward fusion method. All analyses were carried out using SAS Vers. 9.1 software (SAS Institute, Cary, USA).

RESULTS

The used microsatellite markers revealed a total of 42 alleles. The number of alleles per locus ranged from two for *Xgwm160* and *Xgwm344* to six for *Xgwm135.* Genetic diversity ranged from 0.09 (*Xgwm344*) to 0.62 (*Xgwm294*) (Table 2). Based on combined data Gower's dissimilarity ranged from 0.21 between *ID 5585* and *ID 241997-1* (*T. turgidum*) to 0.73 between *ID 241982-2* and *ID 209774* (*T. turgidum* and *T. polonicum*, respectively). Analysis of variance of the Gower dissimilarity matrix showed that the difference within and between species and altitudes were significant ($P<0.0001$), whereas the differences within and between regions were not significant ($P>0.05$) (Table 3). Mean dissimilarities within and between species, regions and altitudes are presented in Tables 4, 5 and 6, respectively. At species level the dissimilarity ranged from 0.44 (within *T. polonicum*) to 0.56 (between *T. polonicum* and *T. turgidum*). On the other hand, within species variability was higher for *T. durum* and *T. turgidum* genotypes. Within region dissimilarity ranged from 0.51 for

Central Ethiopia to 0.53 for accessions of unknown origin while between regions dissimilarity ranged from 0.51 between Central and Southern Ethiopia to 0.54 between accessions of unknown origin and Northern and/or Southern Ethiopia. Generally, accessions of unknown origin had higher within and between regions dissimilarities. The most probable reason is that these accessions have been collected in different regions of Ethiopia.

For altitude, within altitude dissimilarity ranged from 0.47 (≤2200 m) to 0.56 (≤2500 m) while between altitudes dissimilarity ranged from 0.49 between ≤2200 m and accessions of unknown altitude and between ≤2800 m and >2800 m to 0.55 between ≤2200 m and ≤2500 m. Clustering of genotypes using Gower's dissimilarity matrix grouped the 35 genotypes into 6 subgroups (Figure 1). The most remarkable result of the dendrogram is that almost all *T. polonicum* accessions are grouped together, indicating the indigenous evolution of this tetraploid wheat species. *T. durum* and *T. turgidum* accessions were randomly mixed together throughout all clusters.

Discussion

In the present study of combined analysis of molecular marker and quantitative and qualitative phenotypic data variation within and between tetraploid species of Ethiopian origin was evident. Due to the larger number of *T. durum* and *T. turgidum* genotypes variation within these two species were higher than within *T. polonicum*. Genetic dissimilarity within *T. polonicum* was lower than within the other two species. The lower variation within *T. polonicum* genotypes is most probably due to the fewer number of investigated genotypes and the narrower, more indigenous evolution of this species. Therefore, dissimilarity between *T. polonicum* and the other two species is significantly higher than within dissimilarity. The higher variation within T. durum

and T. turgidum and the random mixing of these species in the clusters following cluster analysis of Gower's dissimilarity matrix is not astonishing considering the different developments in wheat taxonomy. Dorofeev *et al*. (1979) clearly differentiated between *T. durum* and *T. turgidum* at species level, whereas MacKey (1988) classified durum wheat as convariety of subspecies *turgidum* of species *turgidum*, i.e. *T. turgidum* subsp. *turgidum* convar. *durum*, van Slageren (1994) followed this idea at the subspecies level, i.e. *T. turgidum* subsp. *durum*, and Kimber & Sears (1987) classified all tetraploid wheats with a BA genome as *T. turgidum* (for a *Triticum* comparative classification table see http://www.k-state.edu./wgrc/Taxonomy/comptri.html).

Conclusion

The present data was enough to depict variation within and between species. Combing molecular with phenotypic data might be more promising. Although within region and between regions dissimilarities were not significant, accessions of unknown origin were responsible for higher dissimilarities. The most probable reason for this observation is that these accessions were collected in different regions. From our results we conclude that accessions of the Ethiopian genebank with no available information about their collection sites are the most variable group and, therefore, can be valuable sources for crop improvement programmes despite the fact that more or less no passport data about their origin is available. From the results of the present study the combined use of molecular markers and phenotypic data is suggested as a promising way for the characterization of genebank accessions.

Acknowledgement

The authors are grateful to Prof. Tamas Lelley, IFA Tulln, Austria, for providing the laboratory facilities for analysis of the molecular part. This work was part of the PhD study for first author and financed by Austrian Agency for International Cooperation in Education and Research.

References

Alamerew S., Chebotar, S., Huang, X., Röder, M., Börner, A. (2004). Genetic diversity in Ethiopian hexaploid and tetraploid wheat germplasm assessed by microsatellite markers. *Genetic Resources and Crop Evolution* 51:559-567.

Dorofeev, V.F., Filatenko, A.A., Migushova, E.F., Udaczin, R.A., Jakubziner, M.M. (1979). Wheat, Vol. 1. In: Dorofeev, V.F., Korovina, O.N. (Eds.), Flora of cultivated plants. Kolos, Leningrad, Russia (in Russian).

Flanagan, L., Wheeler, S., Koeff, M., Knoche, K. (2005). GoTaq® Green Master Mix: from amplification to analysis. Promega Notes 91 (9/05):13-16. [Available online: http://www.promega.com/pnotes/91/12972_13/euro/12972_euro.pdf; verified 12 October 2009].

Kimber, G., Sears, E.R. (1987). Evolution in the genus *Triticum* and the origin of cultivated wheat. Oxford Publication, Madison, WI. pp 154-164.

Promega (1998/99). Technical manual: Wizard®Genomic DNA Purification Kit.

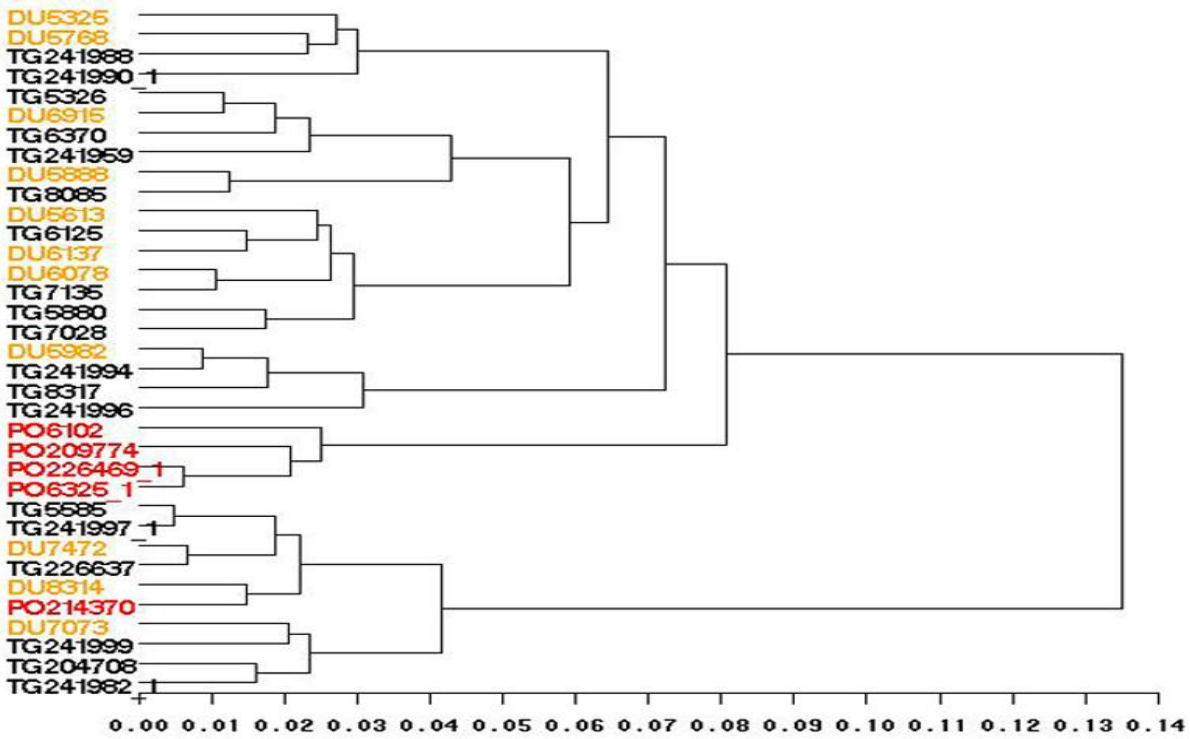
Tables and Figures

Table 1. Accession codes and regions and/or altitudes of collection sites of Ethiopian tetraploid wheat landraces

Accession	Region	Altitude of collection site
T. durum		
5325	Kefa	2667
5613	Shewa	2400
5768	Shewa	2300
5888	Shewa	2920
5982	Shewa	2930
6078	Arsi	2740
6137	Shewa	2670
6915	Gojam	2030
7073	Arsi	2480
7472	Welo	2920
8317	Gamu Gofa	2680
T. polonicum		
6102	Shewa	2430
209774	–	–
214370	Shewa	1975

226469-1	–	–
6325-1	–	–
T. turgidum		
5326	Kefa	–
5585	Shewa	2650
5880	–	–
6125	Shewa	2720
6370	–	–
7028	Arsi	2880
7135	Shewa	2820
8085	–	–
8314	Gamugofa	–
204708	Eritrea	2400
226637	–	–
241959	Gojam	2125
241988	Welo	2845
241994	Tigray	2965
241996	Tigray	2445
241999	Shewa	3030
241982-1	Gonder	3080
241990-1	Welo	2445
241997-1	Tigray	2445

247



248

249 **Figure 1.** Cluster analysis for 35 genotypes of tetraploid wheats using Gower's

250 distance dissimilarity matrix.

251