

UNIVERSITY OF COPENHAGEN



## High genetic diversity revealed in barley (*Hordeum vulgare*) collected from small-scale farmer's fields in Eritrea

Backes, Gunter Martin; Orabi, Jihad; Wolday, Asmelash; Yahyaoui, Amor; Jahoor, Ahmed

*Published in:*  
Genetic Resources and Crop Evolution

*DOI:*  
[10.1007/s10722-008-9347-5](https://doi.org/10.1007/s10722-008-9347-5)

*Publication date:*  
2009

*Document version*  
Publisher's PDF, also known as Version of record

*Citation for published version (APA):*  
Backes, G. M., Orabi, J., Wolday, A., Yahyaoui, A., & Jahoor, A. (2009). High genetic diversity revealed in barley (*Hordeum vulgare*) collected from small-scale farmer's fields in Eritrea. *Genetic Resources and Crop Evolution*, 56(1), 85-97. <https://doi.org/10.1007/s10722-008-9347-5>

# High genetic diversity revealed in barley collected from small-scale farmer's fields in Eritrea

Gunter Backes<sup>1,5</sup>, Jihad Orabi<sup>1,4,5</sup>, Asmelash Wolday<sup>2</sup>, Amor Yahyaoui<sup>3</sup>, Ahmed Jahoor<sup>1</sup>

<sup>1</sup> Department of Agricultural Sciences, Faculty of Life Sciences, The University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg, Copenhagen, Denmark

<sup>2</sup> National Agricultural Research Institute (NARI), Ministry of Agriculture of the Government of Eritrea, P.O. Box 4627, Halhale, Asmara, Eritrea

<sup>3</sup> International Center for Agricultural Research in the Dry Areas (ICARDA), P.O. Box 5466, Aleppo, Syria

<sup>4</sup> General Commission for Scientific Agricultural Research (GCSAR), P.O. Box 113, Douma, Damascus, Syria

<sup>5</sup> These authors contributed equally to this work

## Abstract

Remarkable morphological variation has been found within small Eritrean barley fields. Barley was collected from fields approximately 50 m<sup>2</sup> in size. Spike shape, type, and colour were observed to vary both between and within fields. A set of 39 Simple Sequence Repeats (SSR) markers were used to explore the genetic diversity of the Eritrean barley collected from small-scale farmer's fields. Significant genetic diversity was found within the barley fields. Out of 240 spikes collected from 24 fields (10 spikes per field), only two spikes from geographically distant fields were genetically similar. Based on the SSR data, individual farmers' fields were found to possess 97.3% of the genetic variation present in the Eritrean barley. We discuss a strategy to improve the barley yield in Eritrea, and to facilitate the *in situ* conservation of barley genetic diversity.

Keywords: Barley, *Hordeum vulgare*, Eritrea, genetic diversity, conservation

---

Published 2009 in: Genetic Resources and Crop Evolution 56:85-97

Corresponding Author:

Ahmed Jahoor, Department of Agricultural Sciences  
Faculty of Life Sciences, The University of Copenhagen  
Thorvaldsensvej 40, DK-1871 Frederiksberg, Denmark  
Tel.: +45 3533 3466, Fax: +45 3533 3468, E-mail: aja@kvl.dk

## Introduction

Barley (*Hordeum vulgare* L.) ranks fourth in world cereal production. In Eritrea, barley is one of the most important staple food and feed crops, with an average yield of around 352 kg/ha (2001-2005, FAO 2007). It is grown predominantly in the Central Highland zone, where it is principally used in human nutrition as roasted kernels, porridge, soup, bread and beer, and as straw to feed livestock (Asfaw 2000). Eritrea is found in the Horn of Africa, an area identified as one of the centres of barley diversity, and is considered a region of barley domestication (Orabi et al 2007). Furthermore, barley from this region is a valuable resource for several desirable traits, such as disease resistance against powdery mildew (*Blumeria graminis* f.sp. *hordei*) (Negassa 1985) and scald (*Rhynchosporium secalis*) (Bjørnstad et al 2004; Grønnerød et al 2002). Therefore, landraces and farmers' varieties of Eritrean barley are expected to be an important source of genetic diversity in the future. Farmers have improved landraces over many generations by mass selection, without the use of cross breeding techniques (Fischbeck 2003). Thus, the amount of variability between individuals of a landrace is much higher than in modern varieties (Demissie et al 1998). While cross-breeding has increased both the yield and the quality of barley, it is apparent that genetic diversity, at least those responsible for morphological traits, has decreased considerably (Fischbeck 2003). This loss of diversity is caused by the fact that farmers are increasingly forced to grow genetically improved cross-bred varieties to fulfil their food and feed requirements, thereby, the remarkable diversity present in the landraces disappears (Backes et al 2003; Brantestam et al 2004). It is imperative to maintain a high level of genetic diversity within the gene pool of barley, as well as in other crops, to ensure the preservation of genetic resources required for future crop plant improvement (Kaplan 1998).

Determining the level of variation within, and among, barley populations is an essential step towards conserving genetic resources and developing future strategies for plant improvement. Several studies have been conducted to reveal the substantial level of genetic diversity within barley populations collected from farmers' fields (Hamza et al 2004; Hou et al 2005). Within the framework of these analyses, morphological markers (Alemayehu and Parlevliet 1997; Demissie and Bjørnstad 1996; Kebebew et al 2001; Konishi 2001; Lasa et al 2001) as well as DNA markers (Demissie et al 1998) have been employed. While morphological markers are inexpensive and easily implemented, DNA markers are not commonly affected by the environment and selection, and are also available in almost unlimited numbers. (Backes et al 2003; Ghebru et al 2002). Among the DNA markers, SSR markers have been used intensively to analyse genetic diversity. These markers are

favourable as they exhibit high locus-specificity, high levels of variability, robustness towards genotyping, and a co-dominant mode of inheritance (Matus and Hayes 2002; Russell et al 1997; Woodhead et al 2005).

The conservation of genetic resources can be accomplished either *in situ* (in their natural habitat) or *ex situ* (outside their natural habitat). Conservation *ex situ* in gene banks has been the historical mode of genetic preservation, with extensive collections generated by explorers and researchers that date back to the first part of the last century. More recently, there has been a stronger focus on *ex situ* conservation strategies. One of the reasons for this transition relates to the debate over ownership of these resources, and the increasing demand for gene banks in the countries of collection. As the political situation and/or availability of financial means required for the reliable storage of genetic resources is often limiting, *in situ* conservation has many advantages compared with *ex situ* conservation. Furthermore, with *in situ* conservation, the material is maintained in a dynamic manner and is therefore able to adapt to new threats such as new pathogens or new isolates/races of a given pathogen (Nevo 1998) and climate change. Finally, *in situ* conservation is immediately useful for the local population and is cost-efficient compared with *ex situ* conservation.

The objective of this study was to investigate the genetic diversity among single plant accessions of barley collected from Eritrean farmers' fields in different agro-ecological zones using SSR markers. Furthermore, we sought to elucidate the cause of the diversity observed structure and the implications for future improvement of barley yield as well as conservation of the existing genetic diversity found on farmers' fields in Eritrea.

## Materials and Methods

### Plant materials

A total of 240 single-spikes were collected from 24 farmer's fields, *i.e.* 10 spikes per field, located in main barley cultivation area in the central highland region of Eritrea (Fig. 1, Table 2). Different agro-climate zones were taken into consideration during the collection of the materials, ranging from the high plains and highland terraces to mid-highland plains. The collection was carried out over an area covering no more than 100 km in diameter.

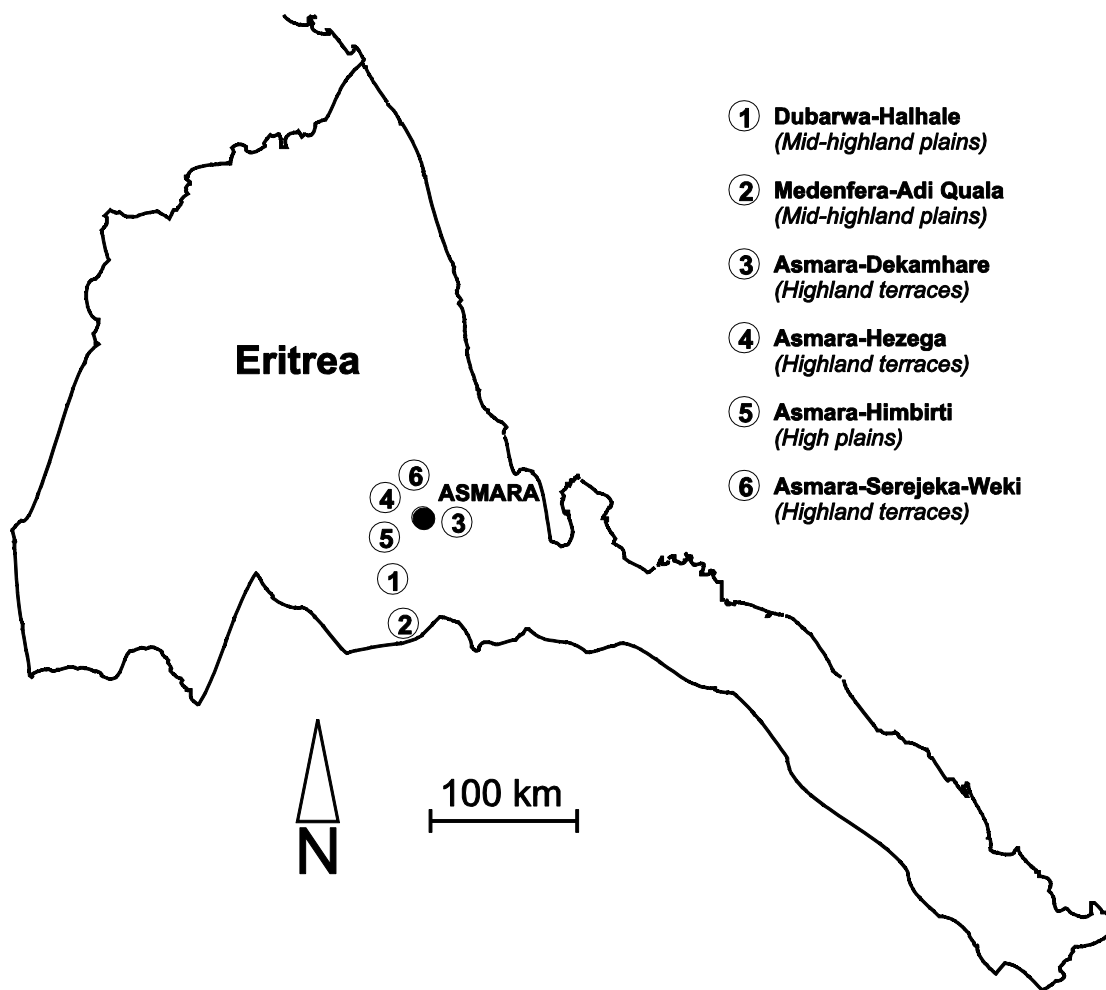


Fig. 1: The geographical distribution of the regions included in the collection of barley material (See also Table 2)

### DNA isolation

Single leaves from 10 day-old seedlings from each sample were freeze-dried for 48 hours. DNA was extracted using Cetyl Trimethyl Ammonium Bromide (CTAB) (100 mM TrisHCl pH 7.5, 666 mM NaCl, 10 mM EDTA, 140 mM  $\beta$ -mercaptoethanol, and 1% w/v CTAB) from the milled freeze-dried leaves according to the protocol described by Saghai-Marooif et al (1984) with some modifications.

### Microsatellites

All samples were tested with 39 nuclear microsatellite markers covering the whole barley genome (Becker and Heun 1995; Liu et al 1996; Ramsay et al 2000). The microsatellite

forward primers were labelled at the 5' end with fluorescent dyes, either 6-carboxyfluorescein (6-FAM, blue), tetrachloro-6-carboxyfluorescein (TET, green), or hexachloro-6-carboxyfluorescein (HEX or VIC, yellow) for the C filter, and either 5'-fluorescein phosphoramidite NED (yellow), HEX or VIC (green) or FAM (blue) for the D filter.

### PCR amplification

PCRs for nuclear microsatellites were performed in Thermo-Fast 96-well plates from ABgene in a final reaction volume of 10  $\mu$ l containing 100 ng template DNA, Mg free PCR buffer, X MgCl<sub>2</sub>, 0.25 units *Taq* polymerase (last three from Promega), 250  $\mu$ M dNTPs, and 1  $\mu$ M each of both forward and reverse primers. PCR reactions were carried out with a GeneAmp® PCR System 2700 thermal cycler from Applied Biosystem. PCR programs for each primer were performed according to Ramsay et al (2000), Liu et al (1996) and Becker and Heun (1995). PCR products were employed and analyzed on an ABI PRISM 377 DNA Sequencer using 4.5% (w/v) polyacrylamide denaturing gels (acrylamide:bisacrylamide 29:1). Fragment sizes were calculated semi-automatically with the computer software GeneScan 2.1 (Perkin Elmer/Applied Biosystems) and by comparing fragments with an internal size standard (ROX or TAMRA) labelled with *N,N,N,N*-tetramethyl-6-carboxyrhodamine.

### Computing

Unless otherwise stated, calculations were carried out using the R language and environment for statistical computing (v.2.3.1) with the base-package (R Development Core Team 2006), the “MASS” package (Venables and Ripley 2002) and the “ade4” package v.1.4-1 (Chessel et al 2005). The random mating or outcrossing rate (*t*) was calculated by the formula  $t = (1 - F_e)/(1 + F_e)$  (Crow and Kimura 1970), where  $F_e$  is the equilibrium inbreeding coefficient under partial selfing.  $F_e$  equals  $1 - H_e/H_r$ , where  $H_e$  is the equilibrium heterozygosity at a given locus or the observed heterozygosity, and  $H_r$  is the expected heterozygosity under random mating (Hardy–Weinberg equilibrium). The polymorphism information content (PIC) is a statistic to measure the informativeness of a given marker locus. PIC values for each marker locus were calculated according to Botstein et al (Botstein et al 1980) as follows:

$$\text{PIC} = 1 - \sum_{i=1}^n p_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2p_i^2 p_j^2$$

where  $p_i$  is the population frequency of the  $i$ -th *allele* and  $n$  is the number of alleles per marker. The genetic distance between individual lines and fields was calculated using Nei's Standard Genetic Diversity (Nei 1972; Pritchard et al 2000) as this dissimilarity coefficient assumes both mutation and genetic drift. The coefficient was transformed into an Euclidian statistic using the method of Caillez (Caillez 1983), included in the R *ade4*-package. Based on this distance, non-parametric Multi-Dimensional Scaling (npMDS) and an Analysis of Molecular Variance (AMOVA) were carried out, the first using the *MASS* package and the latter the *ade4* package in R. Further, the proportion of total alleles and the Average Gene Diversity (Nei 1973) were calculated for single fields, regions and for all 240 lines both as single calculations and as random repeats, using R-scripts. A Bayesian analysis of genetic structure was performed using the software "STRUCTURE" v.2.1 (Pritchard et al 2000), running 20 replicates of runs with a "burn-in period" of 10,000 rounds and 10,000 repeats according to burn-in and assuming an admixture model. The optimal number of groups was determined following the procedure of Evanno et al (2005)

## Results:

### Marker Alleles and Their Informativeness

Out of 39 SSR primer pairs analysed, 37 amplified a single band, while only two primers detected multiple loci (WMC1E8 on chromosome 1H with two loci and HvLTTPB on chromosome 3H with three loci). Three loci, HvLEU, WMC1E8a, and WMC1E8b, were excluded as the Eritrean barley was monomorphic for these markers. Further three loci, Bmac0218, Bmac0156 and Bmag0021, were excluded as they showed more than 5% missing data. This resulted in 36 marker loci yielding a total of 274 different alleles (Table 1). The number of alleles per locus ranged from three to 17 with an average of 7.6 alleles. The informativeness of the allele frequency was described by the number of alleles with a frequency  $\geq 0.3$  (dominant alleles), the number of alleles with a frequency  $\leq 0.05$  (rare alleles) and the PIC value. The PIC is influenced by the number and distribution of alleles. High PIC values result from numerous, equally distributed alleles. The marker HVM03 on chromosome 4H showed not only the highest allele number (17), but also the highest PIC-value (0.861), closely followed by the marker HvLTTPB on chromosome 3H (0.824), with only 9, though rather equally distributed alleles. The lowest PIC-value was found for EBmac0970 (0.066, 5H), which showed one dominant and 3 rare alleles, followed by HVLOX (0.109, 5H), the marker with the lowest number of alleles (3).

Table 1: Marker loci , their chromosome location, number of alleles (total, dominant ( $f \geq 0.3$ ) and rare ( $f \leq 0.05$ )), their Polymorphism Information Content (PIC), the number of heterozygotes out of 240 lines (Het.) and the resulting estimated out crossing.

<i>Locus</i>	<i>Chrom.</i>	<i>Alleles</i>	$\geq 3.0$	$\leq 50.0$	<i>PIC</i>	<i>Het.</i>	<i>t</i>
Bmag0211	1H	6	2	2	0.532	2	0.007
HVM36	2H	5	2	2	0.631	11	0.034
Bmac0093	2H	5	1	3	0.311	21	0.072
Bmac0134	2H	13	1	9	0.688	2	0.005
EBmac0415	2H	6	1	4	0.253	5	0.024
HVM54	2H	6	2	2	0.627	12	0.038
Bmac0209	3H	7	1	4	0.427	2	0.010
Bmag0009	3H	9	1	7	0.264	5	0.038
Bmag0013	3H	10	1	5	0.700	-	-
Bmag0225	3H	7	2	4	0.607	-	-
HvLTPPBa	3H	7	2	5	0.484	-	-
HvLTPPBb	3H	7	2	4	0.61	-	-
HvLTPPBc	3H	9	0	2	0.824	-	-
HvM62	3H	6	1	5	0.132	10	0.161
Bmag0353	4H	11	2	7	0.709	10	0.029
Bmag0384	4H	4	1	2	0.195	7	0.020
HVM03	4H	17	0	10	0.861	13	0.080
HVM40	4H	11	1	6	0.694	13	0.037
HvM67	4H	4	2	2	0.417	4	0.016
Bmac0032	5H	12	1	8	0.727	1	0.003
Bmag0222	5H	10	1	5	0.714	11	0.030
Bmag0223	5H	8	2	4	0.631	0	0.000
EBmac0684	5H	6	1	5	0.155	2	0.015
EBmac0970	5H	4	1	3	0.066	0	0.000
HvLOX	5H	3	1	2	0.109	0	0.000
Bmac0018	6H	5	1	3	0.210	20	0.145
Bmac0040	6H	9	1	5	0.680	13	0.037
Bmac0316	6H	5	1	3	0.258	5	0.027
Bmag0173	6H	10	1	6	0.714	1	0.003
EBmac0806	6H	7	1	3	0.542	9	0.128
Bmac0273c	7H	8	2	4	0.645	7	0.021
Bmag0120	7H	9	1	6	0.495	9	0.033
Bmag0135	7H	10	0	4	0.778	2	0.005
Bmag0206	7H	6	2	3	0.587	1	0.003
HvCMA	7H	5	1	3	0.311	5	0.132
Bmag0125	7H	7	1	4	0.305	10	0.031



### Proportion of Alleles and Average Gene Diversity

The proportion of alleles (number of alleles in a group in relation to the total number of alleles) and the Average Gene Diversity (AGD) was calculated for every field. For the regions and for the total population, a 1,000-fold random sampling with a sample size of 10 lines was performed in order to obtain a reliable estimate of the 'average' for the proportion of alleles and the AGD. Table 2 shows the proportion of alleles and AGD values together with the number of 'private alleles' *i.e.* the alleles that only occur within a given field. Figure 2 and 3 illustrate the AGD values, the mean and the 95%-confidence interval of the mean for a 10-line random sampling in the whole population for the different fields and regions, respectively.

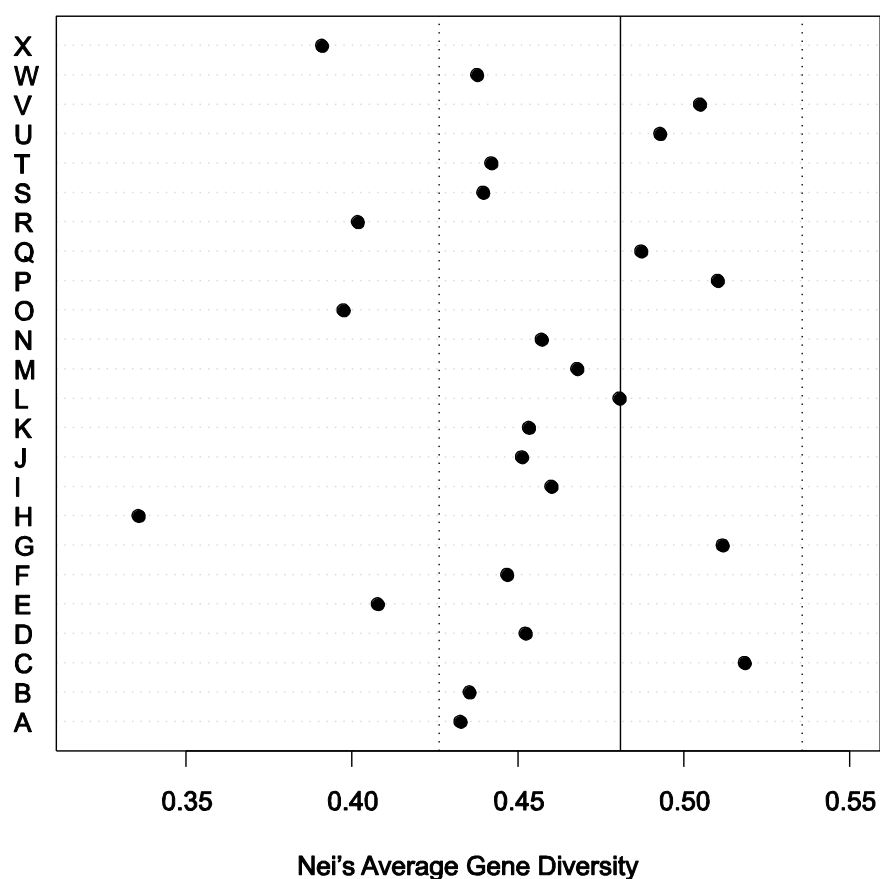


Fig. 2: Nei's Average Gene Diversity for the individual fields (10 lines per field). The vertical solid line is the mean of the random sampling for 10 lines (1,000 repeats) and the dotted lines are the limits of the respective 95%-confidence interval.

Table 2: Proportion of alleles, private alleles, genetic diversity and structure for regions and fields.

<i>Region and topography</i>	<i>Field</i>	<i>Row type</i>	<i>Prop. alleles</i>	<i>Private alleles</i>	<i>AGD</i>	<i>A:I:B</i>
1 Dubarwa-Halhale <i>Mid-highland plains</i>	A. Halhale E	2rs	0.380	6	0.433	5:0:5
	B. Halhale W	2rs	0.365	3	0.435	6:0:4
	C. Adi Hizbay	2rs	0.445	1	0.518	6:0:4
	D. Emni Teslim	2rs	0.398	2	0.452	5:0:5
	E. Dubarwa	2rs	0.361	0	0.408	3:1:6
	F. Adi Geret W	2rs	0.391	3	0.447	5:0:5
	G. Amader W	2rs	0.445	5	0.511	6:0:4
	Avrg./sum region 1		0.393	2.9	0.449	36:1:33
2 Medenfera-Adi Quala <i>Mid-highland plains</i>	H. Abi Adi	2rs	0.336	1	0.336	2:0:8
	I. Adi Ketay	2rs	0.401	1	0.460	4:2:4
		Avrg./sum region 2		0.357	1.0	0.396
3 Asmara–Dekamhare <i>Highland terraces</i>	J. Adi Hawsha + 4	2rs	0.376	1	0.451	8:0:2
	K. Adi Hawsha + 8	2rs	0.405	2	0.453	5:0:5
	L. Wikerti	2rs	0.412	4	0.481	8:0:2
	M. Dekamhare + 8	2rs	0.409	1	0.468	8:0:2
	N. Kurbaria	2rs	0.376	2	0.457	6:0:4
	O. Harien	6rs	0.361	2	0.397	4:6:0
	Avrg./sum region 3		0.419	2.0	0.473	39:6:15
4 Asmara–Hezega <i>Highland terraces</i>	P. Adi Sogdo W	2rs	0.442	8	0.510	6:2:2
	Q. Mekerka E	2rs	0.431	2	0.487	5:2:3
	R. Adi werhiseb	2rs	0.368	2	0.401	10:0:0
	S. Adi Gerbu E	2rs	0.391	4	0.439	10:0:0
	T. Tse Azega	2rs	0.391	2	0.442	7:0:3
	U. Tsaeda Kristian	2rs	0.423	1	0.492	7:0:3
		Avrg./sum region 4		0.416	3.2	0.463
5 Asmara–Himbirti <i>High plains</i>	V. Mbarda E	2rs	0.442	2	0.505	7:0:3
	W. Himbirti E N	2rs	0.409	2	0.438	5:1:4
		Avrg./sum region 5		0.415	2.0	0.465
6 Asmara–Serejeka–Weki <i>Highlands terraces</i>	X. Quandaba E	6rs	0.339	5	0.391	3:7:0
All 240 lines	n.		0.426	2.6	0.486	141:21:78

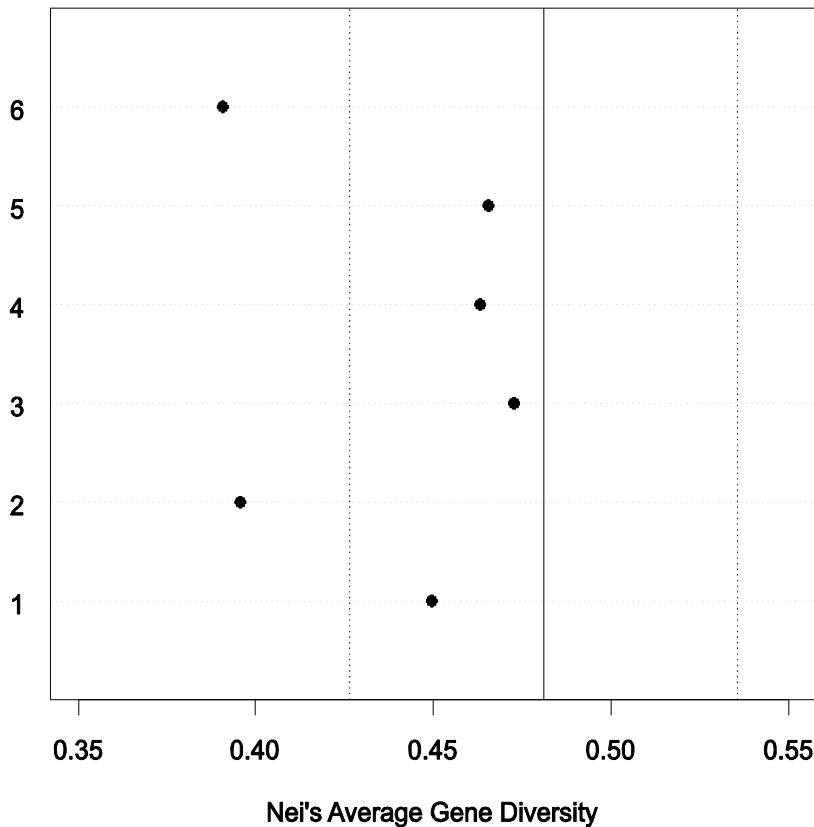


Fig. 3: Nei's Average Gene Diversity for the individual regions except region 6: mean of a random sampling of 10 lines per region, 1,000 repeats). The vertical solid line is the mean of the random sampling for 10 lines (1,000 repeats) and the dotted lines are the limits of the respective 95%-confidence interval.

Most of the fields (Fig. 2) and regions (Fig. 3) lie within the confidence interval of the respective random sampling, implying that they contain the same diversity as a sample of the corresponding size of all lines. Fields E, H, O, R and X and regions 2 and 7 had significantly lower average gene diversities than would be expected from random sampling. Region 7 is represented only by the field X, and region 2 is represented by 2 fields (H and I), where H has the lowest AGD value of all fields (0.336). It is striking that the two fields with 6-row barley (O and X) had significantly lower AGD values than the random sample. No region showed a higher AGD value than the mean of the random samples, while this was the case for fields C, G, P, Q, U and V. This might be a random effect due to the relatively small sample size. While there is a clear positive relationship between the

proportion of alleles and the AGD value, the number of private alleles is independent from these two values.

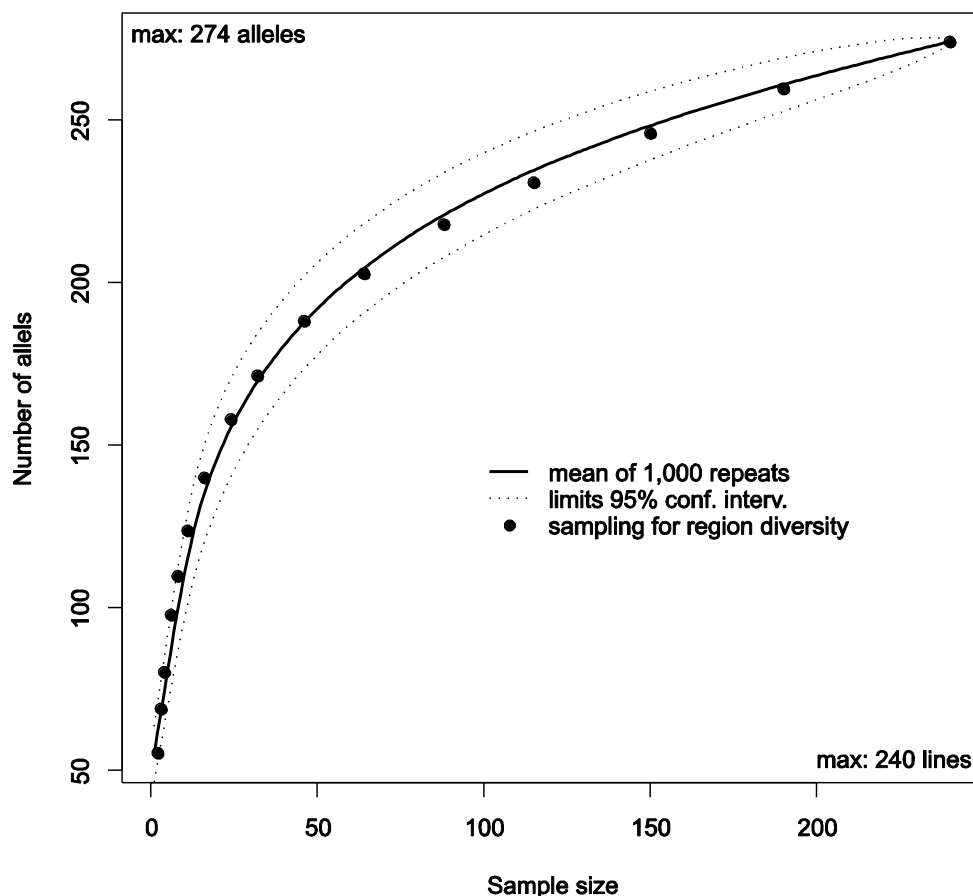


Fig. 4: Relationship between the number of detected alleles over 36 SSR-loci and the sample size as the results of a random sampling of variable number of lines (1,000 repeats per sample size, randomly over all 240 lines). The dotted line gives the 95% confidence intervals while the points reflect an alternative sampling changing the region after the random-selection of each line

The effect of the sample size on the number of detected alleles was evaluated using 1,000 replicates for each sample size. The mean and the 95% confidence interval are presented as solid and dotted lines, respectively (Fig. 4). A sample size of 50 lines yielded a mean of 192 of 274 alleles (70%), and half of all lines, 120 accessions, yielded 237 alleles (86%). In order to obtain at least 95% of all alleles with a 2.5% error probability, 211 out of 240 lines would have to be sampled. Alternative sampling was performed for maximized variability in relation to regions or fields for selected sample sizes. This was done by avoiding

sampling from the same field/region in one round of sampling. For the method optimizing sample variability in relation to the field, the sampling resulted in approximately the same values as for the random sampling (data not shown). When optimizing regional variability, for small sampling sizes, there is a minor increase in the number of alleles in the sample. These data are plotted as points in Fig. 4.

### Degree of out-crossing

The degree of out-crossing was estimated from the observed number of heterozygotes and the expected heterozygosity under Hardy-Weinberg conditions for each single marker. Five markers (Bmag0013, Bmag0225 and HvLTPPBa-c) were excluded as they showed individuals with ambiguous results in relation to heterozygosity. The estimated outcrossing rate  $t$  ranged from 0.000 to 0.161 with an average value of 0.040.

### Population structure

Out of the 240 lines, only 2 showed the same pattern, one line from field B (region 1) and one line from field I (region 2). Another pair of lines from field X (region 6) potentially had the same pattern as both patterns were identical except where there was missing data for the marker HVM36. Distance data were calculated between the individual lines and the results were visualized in a non-parametric Multi-Dimensional Scaling (npMDS, Fig. 5). The npMDS had a stress of 35.6. The only clustering that could be detected visually is marked in Figure 5 by the two ellipses designated as 'cl.1' (cluster 1) and 'cl.2' (cluster 2). In addition, lines from the field X are nearly exclusively occurring in the upper right corner of Figure 5. This shows that these lines have a relatively short distance between them. This was also confirmed by the low AGD-value for this field (Table 2, Fig. 2). Nine of the lines from field H, which had the lowest AGD-value, are grouped in cluster 2 (Fig. 5).

The Bayesian analysis for genetic structure resulted in the best  $\Delta K$ -value for  $K = 2$  (two groups). As there was a clear increase in the value of  $\ln \Pr(X|K)$  from  $K = 1$  (no grouping: -19397) to  $K = 2$  (two groups: -16575), a division into two groups can be accepted. The lines were distributed according to the ancestry estimates in one of the two groups or, if the ancestry estimate was between 0.4 and 0.6, lines were assigned to an 'intermediate group'. The largest group 'A' consisted of 141 lines, 'B' comprised 78 lines, while 21 intermediate lines ('I') were identified. The number of lines per group for each field and region are shown in the last column of Table 2. A comparison of the structural groups with the groups identified visually (Fig. 5) showed that cluster 1 and 2 together compose the structural group 'B' (including only a single intermediate line), while the rest of the lines belong to

the structural group 'A' or show intermediate ancestry. The intermediate lines are shown in bold letters in figure 5. Region 1, and especially region 2, which both are found in the Mid-highland Plains and are situated close to Ethiopia (Figure 1), showed a much higher proportion of lines from the group B' (Table 2) as expected from the average of all regions. Further, it is striking that the two fields with 6-row barley ('O' and 'X') had a high proportion of intermediate lines (13 out of 21).

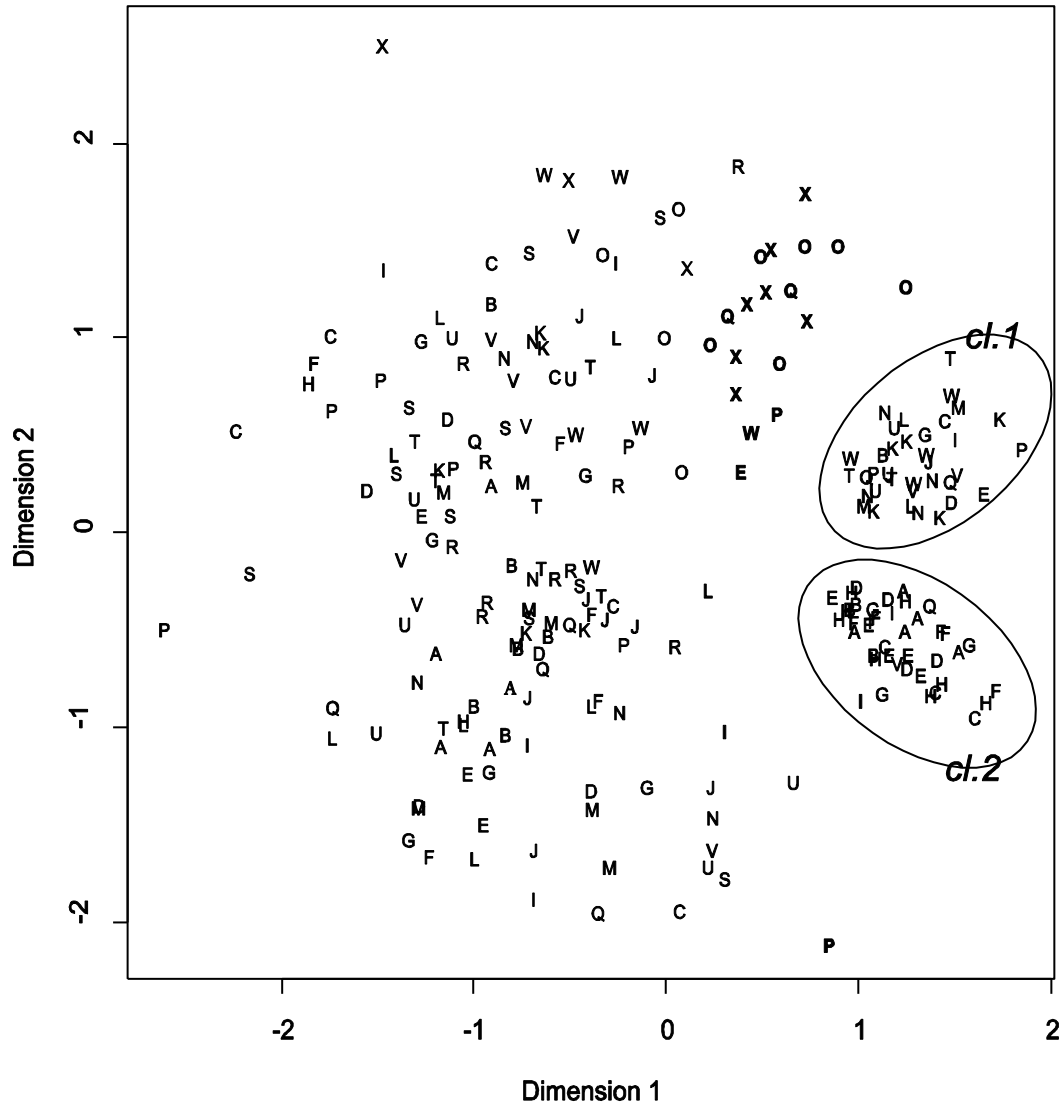


Fig. 5: Results of a two-dimensional non-parametric MDS over individual lines based on Nei's Standard Genetic Distance (Nei 1973). The symbol for the lines corresponds to their field code (see Table 2). 'cl.1' and 'cl.2' are line-cluster that we identified visually in the plot. Lines identified as 'intermediate genotypes' in the cluster analysis are shown in bold.

Allele proportions and AGD-values for the visually detected groups were calculated. In cluster 1 (37 lines), 45% of all alleles were found and an AGD value of 0.263 was calculated. This was significantly lower than the values for a random sample of the same size with 65% and 0.520, respectively. In cluster 2 (43 lines), even lower values were found (42% of the alleles and an AGD-value of 0.229). The respective values for the random sample were 67% and 0.522. The last group containing the rest of the lines did not deviate significantly from the random sample.

The result of an additional npMDS for the distances between the fields (stress: 9.97) is shown in Figure 6. Here, the two fields with 6-row barley (O and X) are clearly separated from the rest of the fields.

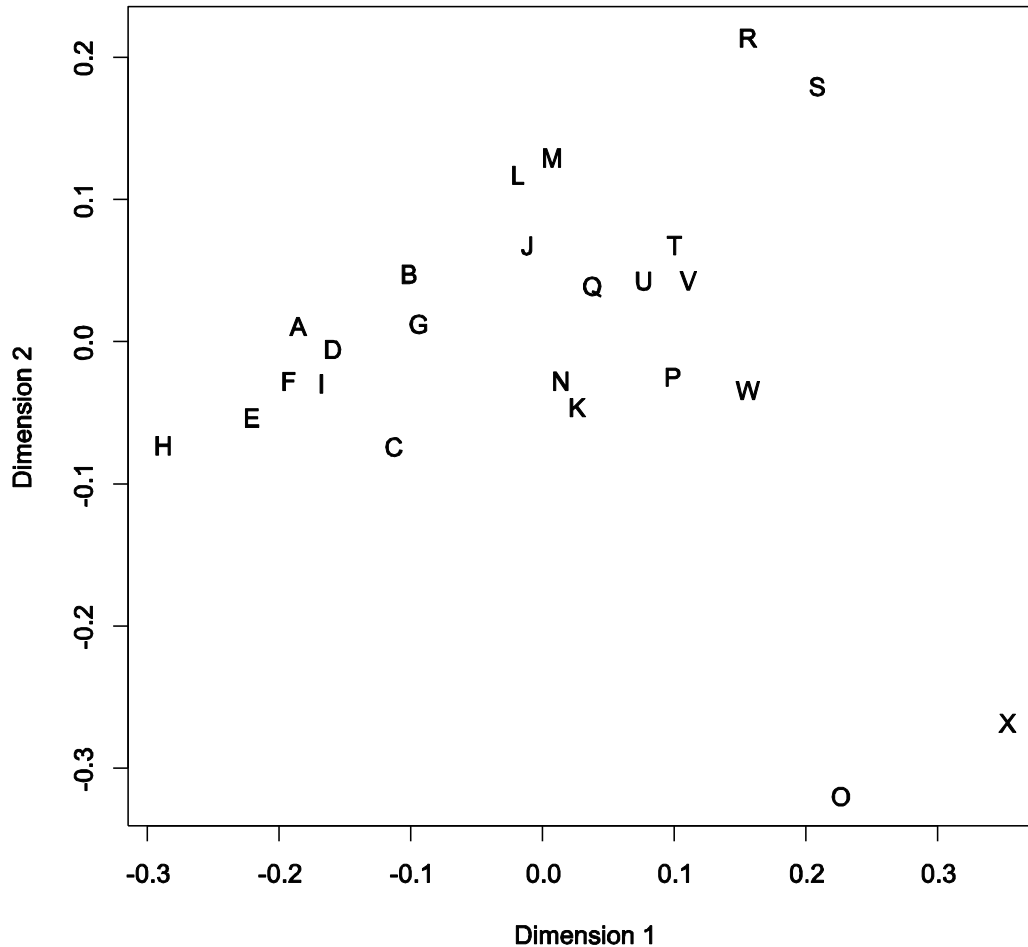


Fig. 6: Results of a two-dimensional non-parametric MDS over fields based on Nei's Standard Genetic Distance (Nei 1973). For the field codes see Table 2.

Finally, the structure was analysed using an Analysis of Molecular Variance (AMOVA) and the results are shown in Table 3. The main part of the variance (97.3%) was contributed by the lines within the field, while regions were responsible for a minor part of the variance (2.5%) and the effect of the field within the regions was not significant.

Table 3: Results of the Analysis of Molecular Variance (AMOVA)

	D.F.	SSQ	MSQ	Var.	%Var.	$\Phi$	p( $\Phi$ )
Between regions	5	150.3	30.06	0.39	2.5	0.230	0.001
Within regions between fields	18	281.8	15.66	0.04	0.2	0.002	0.218
Within fields	216	3303.1	15.29	15.29	97.3	0.024	0.001
Total	239	3735.2	15.63	15.72	100.0		

D.F. = degrees of freedom, SSQ = sums of squares, MSQ = mean sums of squares, Var. = variance, % Var. = percent variance,  $\Phi$  =

## Discussion:

### Genetic Diversity of Barley in Eritrea

The Horn of Africa region is considered to be a centre of origin and diversity for many cultivated crops including barley (Vavilov N.I., 1931). A noticeable morphological diversity was observed in the material collected in the main barley cultivation areas in Eritrea. The data obtained with the 36 SSR loci from 240 lines show an average allele number of 7.6 (Table 1) and an Average Gene Diversity (AGD) of 0.489 (Table 2). In order to achieve a meaningful comparison with published data, a subset of the 240 lines corresponding to the line number in each respective reference was taken 100 times and the average of these random subsets was compared with data from the reference. Hamza et al (2004) studied 15 SSR loci on 26 Tunisian barley accessions and found on average 3.6 alleles per marker and an AGD of 0.45. The corresponding values from our data are 4.4 alleles and an AGD of 0.51, higher than that found for Tunisian barley. Analysing 101 European landraces and varieties grown in Northern Europe over the last 100 years, Russel et al (2000) found 7.5 alleles per locus and an AGD of 0.57. Taking the same number of



lines in the present data, 6.3 alleles and an AGD of 0.53 would be obtained which is slightly lower. Markedly higher values for the number of alleles and AGD were observed by examining wild barley. Ivandic et al (2002) applied 33 SSR loci to 39 wild barley accessions from Israel, Turkey and Iran and obtained 12.9 alleles per locus and an AGD of 0.79. With the data from the present analyses, an average number of alleles of 5.0 and an AGD value of 0.52 would be obtained using the same number of lines. This comparison demonstrates that the Eritrean barley shows a remarkable diversity, taking into account that the collection was done in an area not more than 100 km in diameter. Nevertheless, at least for genomic SSRs, the genetic diversity does not reach the high values observed in wild barley. The similar environmental factors within the rather small collection area might confer an advantage on a limited number of alleles. The situation is different for chloroplast SSRs, where the genetic diversity of the Eritrean lines was reported to be at least as high as that for wild barley lines (Orabi et al 2007). This might be caused by the differences in selective pressure on the two types of SSRs.

### **Structure of the Genetic Diversity in Eritrean Barley**

Commonly, a field planted with a modern barley variety will show a within-field variation close to 0. The distribution of the variation from the material collected in Eritrea, revealed by the AMOVA (Table 3), is therefore striking. Molecular variance at the field level was 97.3%, while almost no variance was assigned to the 'within regions between fields' level and the differences between regions contributed only 2.5%. Similar observations were made for barley in Ethiopia, both on the phenotypic and the marker level. Alemayehu and Parleviet (1997) observed six agronomic traits in 18 Ethiopian barley landraces and found a large degree of variation within the landraces. Demisse and Bjørnstad (1998) analyzed 153 RFLP bands in 43 barley landraces from Ethiopia and concluded that the population is the main source of variability. Aside from barley, other crops in the Horn of Africa have a similar genetic structure. Studying 28 Eritrean sorghum landraces with 15 SSR markers, Ghebru et al (2002) found 50% of the genetic variation within the landraces. For wheat, Bekele (1984) found that the largest part of the variation for 14 morphological/agronomic traits (65-70%) occurred within populations. The latter article circumvents the term 'landrace' to avoid misunderstandings and uses 'population' instead, a term applied mainly in relation to the wild relatives of our crop plants. Even though the term 'landrace' is often used to describe a mixture of different agro-types with diverse genetic background grown together in one field (e.g. Lakew et al 1997, Belay et al 1993, Teklu et al 2006) it seems problematic to describe the field populations in Eritrea as 'landraces', as the term 'race' insinuates a subpopulation that can be discriminated from other 'races' by specific

characteristics. Harlan (1975) wrote: "Landraces have a certain genetic integrity. They are morphologically recognizable; farmers have names for them and different landraces are understood to differ in adaptation to soil type, time of seeding, date of maturity, height, nutritive value, use and other properties." At least this definition of a landrace certainly does not describe a field population in Eritrea.

Asfaw (2000) points out that barley farmers in Ethiopia have very elaborate names for the different types of barley based on their specific morphological traits. The field populations in Eritrea we analysed always comprised a mixture of these types. This might be a situation provoked consciously by the farmer in order to stabilise yields. Another factor leading to this population composition might be that farmers revitalise their seed stocks from time to time with seeds from friends and relatives (Asfaw 2000), thereby decreasing the genetic variation between fields and regions while simultaneously increasing the within-field variation. A further reason for the structure observed might be that yield levels are relatively low and the farmers therefore have to keep a higher proportion of seed for the next sowing season. Consequently, selection is low and together with a modest outcrossing rate, higher within-field diversity could be obtained (Alemayehu and Parleviet 1997). The outcrossing rate estimated from the collected material was 3.8% using 39 SSR loci (Table 1). This is relatively high compared with the rate found in the literature. Abdel-Ghani et al (2004) found an average out-crossing rate of 1.17% (0-3.4%) for 13 wild barley lines and 0.59% (0-3.5%) for 12 barley landraces using 4 SSR markers. With 22 isoenzyme loci, Brown et al 1978 estimated the average out-crossing rate of 26 wild barley populations to be 1.7% (0-9.6%). All these factors might have resulted in a population structure which is often found among wild barley. Turpeinen et al (2003) analyzed 94 wild barley lines from 10 locations within Israel with AFLP markers and found the within-population variation to be 69% vs. a value of 31% for the between-population variation. Similar results for wild barley were also obtained by Baum et al (1997) and Huang et al (2002).

While the high overall diversity and the huge field level diversity as discussed above reflect the general situation in Eritrea, differences at the field level, and to a minor extent at the regional level, were also detected. As shown in Table 2 and Figure 2, it is evident that the 6-row barley (field O and X) generally showed lower AGD-values than the average of the all types. This might be caused by a relatively recent introduction into the area or recent mutation, both of which could cause a bottleneck. Recent introductions might also play a role in the patterns observed with the Bayesian structure analysis. The closer the region of collection is to Ethiopia, the higher the relative proportion of lines assigned to the 'B' group. Thereby, an introgression from Ethiopia could explain the different genetic

structure. The reduced diversity in these lines could then be explained by a bottleneck-effect during the selective introgression and subsequent distribution of only a relatively small subsample of barley lines from Ethiopia. The intermediate positioning of the 6-row barley in the Bayesian analysis and the MDS (Fig. 5) is more difficult to interpret. A possible explanation could be that these lines were introgressed earlier and hybridized with the original Eritrean barley. Consequently, the introgressed lines might have superior yields.

### **Conservation and yield improvement**

The results of the present analysis and the conclusions we draw highlight a dilemma for Eritrean barley breeding between conservation of diversity and yield improvement. It is obvious that Eritrean barley farmers consciously maintain extremely high within-field diversity. This heterogeneity provides greater stability, especially under 'low-input' conditions (Wolfe 2000). Therefore, the present agricultural system in Eritrea is ecologically balanced and is well adapted to local conditions. Furthermore, the system constitutes a stunning example of *in situ* conservation of highly diverse genetic material, even though threats such as the introgression of lines with reduced genetic diversity can already be recognised. On the other hand, a yield of 352 kg/ha (2001-2005, FAO 2007) is very low and Eritrea urgently needs to increase the yield of barley. The use of genotypes with higher yield potential could contribute to increased barley production. However, the introduction of high yielding varieties could be risky both for the genetic diversity present in Eritrea and for yield stability. An example for the genetic erosion is the introduction of the six-row barley variety 'Rihane' to Tunisia in 1983. At present, this variety occupies about 60% of the total barley cultivation area and has, accordingly, superseded the local landraces (Hamza et al 2004). With the genetic diversity in the field vanishing, the risk of severe yield reduction in a single year would be increased, either by fungal pathogens, insect pests or unfavourable environmental conditions. In addition, cultural traditions are adapted to this cropping system. For all these reasons, a specific approach has to be found in Eritrea. One way would be to select superior lines from the existing material and to bulk them as mixtures, as proposed by Ceccarelli and Grandó (2000). An alternative, as discussed above, would be to select lines with superior traits, with respect to yield, disease resistance and quality, representing the genetic diversity at the allele level using molecular marker studies. In the next step, subgroups would have to be generated within these subgroups and multiple crosses between the selected lines would have to be made in order to produce composite crosses. Subsequently, the F<sub>1</sub> should be grown in bulk, and negative selection conducted in subsequent years in cooperation with local farmers. This approach

combines the composite-crosses advocated by Philips and Wolfe (2005) with the participatory plant breeding promoted by Ceccarelli and Grando (2007), and should ensure long-term high yield and yield stability, with optimised local adaptation and increased local acceptance of the product. While this method aims to maximise genetic diversity, it is at the same time an *in situ* conservation approach to preserving the rich barley gene pool at the Horn of Africa for future generations.

*In situ* conservation (at farmer's fields) is in many cases superior to *ex situ* conservation (in gene banks) and is therefore given priority by the 1992 Rio Convention on Biological Diversity. It allows continued adaptation to changing abiotic and biotic conditions (Nevo, 1998), and avoids dependence on local gene banks that are vulnerable to unstable political conditions and limited resources. On the other hand, adverse climatic conditions, pests and diseases can destroy efforts put *in situ* conservation into practice. Therefore, a local gene bank should be established as an *ex situ* back-up for the *in situ* conservation and this gene bank should be supported by the international community. This rich and unexploited gene pool of cultivated barley is not only part of the heritage of mankind, but can also be very useful for meeting current or new challenges in barley breeding.

## Acknowledgements

We would like to thank the Eritrean farmers who allowed us to sample the spikes from their fields. We also thank Mr. T. Ghebremania for his advice on the collection sites in Eritrea, and Mr. M.H. Matougui (Consultant -ICARDA) who assisted with the barley collection. Financial support for this research was provided by Danish International Development Assistance (DANIDA) and the Islamic Development Bank.

## References

- Abdel-Ghani AH, Parzies HK, Omary A, Geiger HH (2004) Estimating the outcrossing rate of barley landraces and wild barley populations collected from ecologically different regions of Jordan. *Theor Appl Genet* 109:588-595.
- Alemayehu F, Parlevliet0 JE (1997) Variation between and within Ethiopian barley landraces. *Euphytica* 94:183-189.
- Asfaw Z (2000) The barleys of Ethiopia. In: Genes in the field. Brush SB (ed) International Development Research Center, Ottawa, pp. 77-108.

- Backes G, Hatz B, Jahoor A, Fischbeck G (2003) RFLP diversity within and between major groups of barley in Europe. *Theor Appl Genet* 122:291-299.
- Baum BR, Nevo E, Johnson D, Beiles A (1997) Genetic diversity in wild barley (*Hordeum spontaneum* C. Koch) in the Near East: a molecular analysis using Random Amplified Polymorphic DNA (RAPD) markers. *Genet Resour Crop Evol* 44:147-157.
- Becker J, Heun M (1995) Barley microsatellites - allele variation and mapping. *Plant Mol Biol* 27:835-845.
- Belay G, Tesemma T, Becker HC, Merker A (1993) Variation and interrelationships of agronomic traits in Ethiopian tetraploid wheat landraces. *Euphytica* 71:181-188.
- Bekele E (1984) Analysis of regional patterns of phenotypic diversity in the Ethiopian tetraploid and hexaploid wheats. *Hereditas* 100:131-154.
- Bjørnstad Å, Grønnerød S, Key JM, Tekauz A, Crossa J, Martens H (2004) Resistance to barley scald (*Rhynchosporium secalis*) in the Ethiopian donor lines 'Stuedelli' and 'Jet', analyzed by partial least squares regression and interval mapping. *Hereditas* 141:166-179.
- Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of A Genetic-Linkage Map in Man Using Restriction Fragment Length Polymorphisms. *Am J Hum Genet* 32:314-331.
- Brantestam AK, von Bothmer R, Dayteg C, Rashal I, Tuveesson S, Weibull J (2004) Inter simple sequence repeat analysis of genetic diversity and relationships in cultivated barley of Nordic and Baltic origin. *Hereditas* 141:186-192.
- Brown AHD, Zohary D, Nevo E (1978) Outcrossing Rates and Heterozygosity in Natural-Populations of *Hordeum spontaneum* Koch in Israel. *Heredity* 41:49-62.
- Caillez F (1983) The analytical solution of the additive constant problem. *Psychometrika* 48:305-310.
- Ceccarelli S, Grando S (2000) Barley landraces from the Fertile Crescent: a lesson for plant breeders. In: Brush SB (ed) *Genes in the field*. International Development Research Center, Ottawa, pp. 51-76.
- Ceccarelli S, Grando S (2007) Decentralized-participatory plant breeding: an example of demand driven research. *Euphytica* 155:349-360.
- Chessel D, Dufour A-B, Dray S, Lobry JR, Ollier S, Pavoine S, Thioulouse J (2005) ade4: analysis of environmental data : exploratory and euclidean methods in environmental sciences. Available from: <http://pbil.univ-lyon1.fr/ADE-4> (accessed June 2007)
- Crow JF, Kimura M (1970) *An introduction to population genetics theory*. Harper & Row Publishers, New York.

- Demissie A, Bjørnstad Å (1996) Phenotypic diversity of Ethiopian barleys in relation to geographical regions, altitudinal range, and agro-ecological zones: as an aid to germplasm collection and conservation strategy. *Hereditas* 124:17-29.
- Demissie A, Bjørnstad Å, Kleinhofs A (1998) Restriction fragment length polymorphisms in landrace barleys from Ethiopia in relation to geographic, altitude and agro-ecological factors. *Crop Sci* 38:237-243.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611-2620.
- FAOSTAT (2007) FAO Statistics Division, Rome, Italy. <http://faostat.fao.org>. Cited 27 June 2007
- Fischbeck G (2003) Diversification through breeding. In: von Bothmer R, van Hintum T, Knüpffer H, Sato K (eds) *Diversity in barley (Hordeum vulgare)*. Elsevier Science, Amsterdam, pp. 29-52.
- Ghebru B, Schmidt RJ, Bennetzen JL (2002) Genetic diversity of Eritrean sorghum landraces assessed with simple sequence repeat (SSR) markers. *Theor Appl Genet* 105:229-236.
- Grønnerød S, Marøy AG, Bjørnstad Å, MacKey J, Tekauz A, Penner GA, Bjørnstad Å (2002) Genetic analysis of resistance to barley scald (*Rhynchosporium secalis*) in the Ethiopian line 'Abyssinian' (CI668). *Euphytica* 126:235-250.
- Harlan JR (1975) Our Vanishing Genetic Resources. *Science* 188:618-621.
- Hamza S, Ben Hamida W, Rebai A, Harrabi M (2004) SSR-based genetic diversity assessment among Tunisian winter barley and relationship with morphological traits. *Euphytica* 135:107-118.
- Hou Y-C, Yan Z-H, Wei Y-M, Zheng Y-L (2005) Genetic diversity in barley from west China based on RAPD and ISSR analysis. *Barley Genet Newsl* 35:9-22.
- Huang Q Y, Beharav A, Li YC, Kirzhner V, Nevo E (2002) Mosaic microecological differential stress causes adaptive microsatellite divergence in wild barley, *Hordeum spontaneum*, at Neve Yaar, Israel. *Genome* 45:1216-1229.
- Ivandi V, Hackett CA, Nevo E, Keith R, Thomas WTB, Forster BP (2002) Analysis of simple sequence repeats (SSRs) in wild barley from the Fertile Crescent: associations with ecology, geography and flowering time. *Plant Mol Biol* 48:511-527.
- Kaplan JK (1998) Conserving the world's plants. *Agr Res* 46:4-9.
- Kebebew F, Tsehaye Y, Neilly T (2001) Morphological and farmers cognitive diversity of barley (*Hordeum vulgare* L. [Poaceae]) at Bale and North Shewa of Ethiopia. *Genet Resour Crop Evol* 48:467-481.

- Konishi T (2001) Genetic diversity in *Hordeum agriocrithon* E. Åberg, six-rowed barley with brittle rachis, from Tibet. *Genet Resour Crop Evol* 48:27-34.
- Lakew B, Semeane Y, Gebre H, Grando S, van Leur JAG, Ceccarelli S (1997) Exploiting the diversity of barley landraces in Ethiopia. *Genet Resour Crop Evol* 44:109-116.
- Lasa JM, Igartua E, Ciudad FJ, Codesal P, Garcia EV, Gracia MP, Medina B, Romagosa I, Molina-Cano JL, Montoya JL (2001) Morphological and agronomical diversity patterns in the Spanish barley core collection. *Hereditas* 135:217-225.
- Liu ZW, Biyashev RM, Saghai-Marooof AS (1996) Development of simple sequence repeat DNA markers and their integration into a barley linkage map. *Theor Appl Genet* 93:869-876.
- Matus IA, Hayes PM (2002) Genetic diversity in three groups of barley germplasm assessed by simple sequence repeats. *Genome* 45:1095-1106.
- Negassa M (1985) Genetics of resistance to powdery mildew in some Ethiopian barleys. *Hereditas* 102:123-138.
- Nei M (1972) Genetic distance between populations. *Am Nat* 106:283-&.
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci Unit States Am* 70:3321-3323.
- Nevo E (1998) Genetic diversity in wild cereals: Regional and local studies and their bearing on conservation ex situ and in situ. *Genet Resour Crop Evol* 45:355-370.
- Orabi J, Backes G, Wolday A, Yahyaoui A, Jahoor A (2007) The Horn of Africa as a centre of barley diversification and a potential domestication site. *Theor Appl Genet* 114:1117-1127.
- Phillips SL, Wolfe MS (2005) Evolutionary plant breeding for low input systems. *Journal of Agricultural Science* 143:245-254.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.
- R Development Core Team (2006) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Ramsay L, Macaulay M, Ivanissevich SD, MacLean K, Cardle L, Fuller J, Edwards KJ, Tuveesson S, Morgante M, Massari A, Maestri E, Marmiroli N, Sjakste T, Ganai M, Powell W, Waugh R (2000) A simple sequence repeat-based linkage map of barley. *Trends Plant Sci* 156:1997-2005.
- Russell JR, Fuller JD, Macaulay M, Hatz BG, Jahoor A, Powell W, Waugh R (1997) Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theor Appl Genet* 95:714-722.

- Russell J, Ellis RP, Thomas WTB, Waugh R, Provan J, Booth A, Fuller J, Lawrence P, Young G, Powell W (2000) A retrospective analysis of spring barley germplasm development from 'foundation genotypes' to currently successful cultivars. *Mol Breed* 6:553-568.
- Saghai-Marouf MA, Soliman KM, Jørgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population-dynamics. *Proc Natl Acad Sci Unit States Am* 81:8014-8018.
- Teklu Y, Hammer K, Huang XQ, Röder MS (2006) Analysis of microsatellite diversity in ethiopian tetraploid wheat landraces. *Genet Resour Crop Evol* 53:1115-1126.
- Turpeinen T, Vanhala T, Nevo E, Nissila E (2003) AFLP genetic polymorphism in wild barley (*Hordeum spontaneum*) populations in Israel. *Theor Appl Genet* 106:1333-1339.
- Vacilov NI (1931) The problem of the origin of the world's agriculture in the light of the latest investigations. Available from: <http://www.marxists.org/subject/science/essays/vavilov.htm> (accessed June 2007).
- Venables WN, Ripley BD (2002) *Modern Applied Statistics with S*. Springer, New York.
- Wolfe MS (2000) Crop strength through diversity. *Nature* 406:681-682.
- Woodhead M, Russell J, Squirrell J, Hollingsworth PM, Mackenzie K, Gibby M, Powell W (2005) Comparative analysis of population genetic structure in *Athyrium distentifolium* (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. *Mol Ecol* 14:1681-1695.