Morphological and Molecular Characterization of *Lepidium sativum* population collected from Ethiopia

Said Mohammed¹* and Kassahun Tesfaye²

¹Biology Department, College of Natural and Computational Sciences, Debre Birhan University, Ethiopia.
²Institute of Biotechnology, Addis Ababa University, Ethiopia.

INTRODUCTION

The genus *Lepidium* L. comprises about 150 species distributed worldwide. In tropical Africa, only nine species are found. The genus *Lepidium* belongs to the family Brassicaceae. The garden cress, *Lepidium sativum* L., is a fast growing annual herb native to Egypt and West Asia (Zhan et al., 2009). Medicinal plants are excellent sources of unknown chemical substances with therapeutic effects (Rao, 2004). *L. sativum* seeds contain flavonoids, coumarins, sulphur, glycosides, triterpenes, sterols and various imidazole alkaloids (Radwan et al., 2007; Agarwal and Verma, 2011; Datta et al., 2011). Ethno-medicinal uses of *L. sativum* are: leaves are used as salad, cooked with vegetables, curries and also used as fodder for cattle (Moser et al., 2009; Patel et al., 2009;
Raval et al., 1999) software, and very few studies have used this method with arithmetic mean of coefficient (Jaccard, 1908). The study used markers outside sustainable conservation and use were used to compare the population and generate phenogram using NTSYS (UPGMA) (Sneath and Sokal, 1973) was used to analyze and calculate Jaccard's similarity coefficient. AMOVA was used to calculate variation among populations as number of polymorphic loci, percent polymorphism, gene diversity (H) and Shannon diversity index (I).

**MATERIALS AND METHODS**

**Tissue harvest and DNA extraction**

The experiment was designed to characterize these accessions using inter simple sequence repeat (ISSR) markers. Borsch et al. (2003) procedures were used.

**Primer selection and optimization**

The ISSR marker assay was conducted at Genetics Laboratory of the Microbial, Cellular and Molecular Biology Program Unit, College of Natural Sciences, Addis Ababa University, Addis Ababa. A total of 10 primers, obtained from the Genetic Research Laboratory (Primer kit UBC 900) and primers used by Kim et al. (2002) were used for the initial testing of primers variability and reproducibility.

**PCR and gel electrophoresis**

The polymerase chain reaction was conducted in Biometra 2003 T3 Thermo cycler. PCR amplification was carried out in a 25 µl reaction mixture containing 1 µl template DNA, 13.45 µl H2O, 5.60 µl dNTP (1.25 mM), 2.6 µl Taq buffer (10XH buffer S), 1.25 µl MgCl2 (50 mM), 0.6 µl primer (20 pmol/µl) and 0.5 µl Taq Polymerase (3 U/µl). The amplification program was 4 min preheating and initial denaturation at 94°C, then 40 x 15 s at 94°C, 1 min primer annealing at (45/48°C) based on primers used; 1.30 min extension at 72°C and the final extension for 7 min at 72°C. The PCR reactions were stored at 4°C until loading on gel for electrophoresis. The amplification products were differentiated by electrophoresis using an agarose gel (1.67% agarose with 100 ml 1xTBE) and 8 µl amplification product of each sample with 2 µl loading dye (6 times concentrated) was loaded on gel. DNA marker 100 bp was used to estimate molecular weight and size of the fragments. The electrophoreses were done for 3 h at constant voltage of 100 V. The DNA was stained with 10 mg/ml ethidium bromide which were mixed with 250 ml distilled water for 30 min and washed with distilled water for 30 min (Table 1).

**Statistical analysis**

The bands were recorded as discrete characters, presence ‘1’ or absence ‘0’ and ‘?’ for missing data. Based on recorded bands, different software were used for analysis. POPGENE version1.32 software (Yeh et al., 1999) was used to calculate genetic diversity for each population as number of polymorphic loci, percent polymorphism, gene diversity (H) and Shannon diversity index (I). Analysis of molecular variance (AMOVA) was used to calculate variation among and within population using Arealinquim version 3.01 (Excoffier et al., 2006). NTSYS-pc version 2.02 (Rohlf, 2000) and Free Tree 0.9.1.50 (Pavlicek et al., 1999) software were used to calculate Jaccard’s similarity coefficient (Table 2).

The unweighted pair group method with arithmetic mean (UPGMA) (Sneath and Sokal, 1973) was used to analyze and compare the population and generate phenogram using NTSYS- pc version 2.02 (Rohlf, 2000).

To further examine the patterns of variation among individual samples on 3D, a principal coordinated analysis (PCO) was performed based on Jaccard’s coefficient (Jaccard, 1908). The

---

### Table 1. List of primers, annealing temperature, primer sequence, amplification quality and repeat motives used for optimization

<table>
<thead>
<tr>
<th>Primer</th>
<th>Annealing temperature (°C)</th>
<th>Primer sequence</th>
<th>Amplification quality</th>
<th>Repeat motives</th>
</tr>
</thead>
<tbody>
<tr>
<td>810</td>
<td>45</td>
<td>GAGAGAGAGAGAGAT</td>
<td>Monomorphic</td>
<td>Dinucleotide</td>
</tr>
<tr>
<td>812</td>
<td>45</td>
<td>GAGAGAGAGAGAGAA</td>
<td>Polymorphic, reproducible</td>
<td>Dinucleotide</td>
</tr>
<tr>
<td>818</td>
<td>48</td>
<td>CACACACACACACAAAG</td>
<td>Monomorphic</td>
<td>Dinucleotide</td>
</tr>
<tr>
<td>824</td>
<td>48</td>
<td>TCTCTCTCTCTCTCTCG</td>
<td>Monomorphic</td>
<td>Dinucleotide</td>
</tr>
<tr>
<td>834</td>
<td>45</td>
<td>AGAGAGAGAGAGAGAGYT</td>
<td>Polymorphic, reproducible</td>
<td>Dinucleotide</td>
</tr>
<tr>
<td>844</td>
<td>45</td>
<td>GAGAGAGAGAGAGAGYT</td>
<td>No Amplification</td>
<td>Dinucleotide</td>
</tr>
<tr>
<td>872</td>
<td>38</td>
<td>GATAGATAGATAGATA</td>
<td>No Amplification</td>
<td>Tetra- nucleotide</td>
</tr>
<tr>
<td>873</td>
<td>45</td>
<td>GACAGACAGACAGACA</td>
<td>Polymorphic, reproducible</td>
<td>Tetra- nucleotide</td>
</tr>
<tr>
<td>878</td>
<td>45</td>
<td>GATGAAAATGATGGAT</td>
<td>No Amplification</td>
<td>Tetra- nucleotide</td>
</tr>
<tr>
<td>880</td>
<td>45</td>
<td>GAGAGAGAGAGAGAGA</td>
<td>Polymorphic, reproducible</td>
<td>Penta- nucleotide</td>
</tr>
</tbody>
</table>

Table 2. Banding patterns generated using the four selected primers, their repeat motifs, amplification patterns and number of scored bands.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Repeat motif</th>
<th>Amplification quality</th>
<th>Number of scored bands</th>
</tr>
</thead>
<tbody>
<tr>
<td>812</td>
<td>(GA)8A</td>
<td>Excellent</td>
<td>14</td>
</tr>
<tr>
<td>834</td>
<td>(AG)8YT</td>
<td>Excellent</td>
<td>11</td>
</tr>
<tr>
<td>873</td>
<td>(GACA)4</td>
<td>Excellent</td>
<td>16</td>
</tr>
<tr>
<td>880</td>
<td>(GGAGA)3</td>
<td>Excellent</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>53</td>
</tr>
</tbody>
</table>


Table 3. Number of scorable bands (NSB), number of polymorphic loci (NPL), percent polymorphism (PP), genetic diversity (H) and Shanon index information (I) of 85 L. sativum accessions based on all primers used.

<table>
<thead>
<tr>
<th>Population</th>
<th>NSB</th>
<th>NPL</th>
<th>PP</th>
<th>H±SD</th>
<th>I±SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amhara</td>
<td>35</td>
<td>66.04</td>
<td>0.24±0.19</td>
<td>0.35±0.28</td>
<td></td>
</tr>
<tr>
<td>Oromia</td>
<td>27</td>
<td>50.94</td>
<td>0.17±0.19</td>
<td>0.26±0.28</td>
<td></td>
</tr>
<tr>
<td>Tigray</td>
<td>35</td>
<td>66.04</td>
<td>0.24±0.19</td>
<td>0.35±0.27</td>
<td></td>
</tr>
<tr>
<td>SNNPR</td>
<td>25</td>
<td>47.17</td>
<td>0.18±0.21</td>
<td>0.27±0.30</td>
<td></td>
</tr>
<tr>
<td>Somali</td>
<td>24</td>
<td>45.28</td>
<td>0.18±0.21</td>
<td>0.26±0.30</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>182</td>
<td>343.39</td>
<td>1.27±1.20</td>
<td>1.87±1.72</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. The number of polymorphic loci (NPL), percent polymorphism (PP), genetic diversity (H) and Shannon information index (I) among the five regions of Ethiopia.

Analysis of molecular variance (AMOVA)

Analysis of molecular variance was carried out on the overall ISSR data score of L. sativum accessions without grouping by region or geographic location. AMOVA revealed high percentage of variation (94%) that is attributed to within population variation while the remaining variation is due to among population variation (6%). The variation was found to be highly significant at \( P = 0.00 \). The result shows that there is high gene flow or seed flow among population in different region; this resulted in low genetic variation and differentiation among population (Table 5).

Clustering analysis

UPGMA and Neighbor Joining tree construction methods was used to construct dendrogram for six populations and 85 individuals based on 53 PCR bands amplified by two di-nucleotides (812 and 834), one penta nucleotides (880) and one tetra nucleotide (873). The dendrogram derived from neighbor-joining analysis of the whole ISSR data with 85 L. sativum accessions showed four distinct clusters and two sub-clusters within each major cluster. Most of the individual accessions collected from the same region tend to spread all over the tree without forming their own grouping. The wider distribution of L. sativum in population from Somali region. No unique bands were observed for either the accessions or the populations (Table 3).

Among the L. sativum accessions evaluated using ISSR markers, samples from Tigray and Amhara exhibited the highest gene diversity (H = 0.24), whereas samples from Oromia had H = 0.17 from SNNPR H = 0.18 and Somali H = 0.18 gene diversity values. The average gene diversity for the total population (H) was 0.27 (Table 4).

Primer 873 showed highest gene and Shannon diversity (0.36 and 0.53, respectively) and primer 812 was the least (0.20 and 0.31, gene and Shannon diversity, respectively) (Table 3).

RESULTS

Genetic diversity analysis

Of the total 53 loci scored, 81.13% (43) were observed to be polymorphic. From all the populations studied, Amhara and Tigray were 66.04%, Oromia 50.94%, SNNPR 47.17% and Somali 45.28% percent polymorphic. Amhara and Tigray showed more percent polymorphism; while the least polymorphism was detected calculation of Jaccard’s coefficient was made with PAST software version 1.18 (Hammer et al., 2001). The first three axes were used to plot the three dimensional PCO with STATISTICA version 6.0 software (Hammer et al., 2001; Statistica soft, Inc.2001).
Table 5. Analysis of molecular variance (AMOVA) of L. sativum accessions in Ethiopia without grouping.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>Percentage of variation</th>
<th>Fixation</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among populations</td>
<td>4.122</td>
<td>0.02834</td>
<td>6.00</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>Within population</td>
<td>34.765</td>
<td>0.44387</td>
<td>94.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>38.888</td>
<td>0.47221</td>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. ISSR fingerprint generated from 16 individual accessions using primer 873.

accession all over the tree shows the low divergence among populations from different localities. UPGMA analysis based on regions of collection of L. sativum revealed three major groups. The first cluster contains Oromia, Amhara and Tigray; while the second cluster contains SNNPR and individual from unknown origins. The final major cluster contains the Somali group. However, UPGMA with individual accessions showed intermixing of individuals to different groups, except in two groups where individuals from Oromia clustered together (Figure 2).

**DISCUSSION**

**Molecular diversity and its implications for improvement and conservation**

In the present study, ISSR was used for the first time to assess genetic variation of L. sativum populations from Ethiopia. This method provides an alternative choice to other system for obtaining highly reproducible markers without any necessity for prior sequence information for various genetic analyses. Because of the abundant and rapidly evolving SSR regions, ISSR amplification has the potential of illuminating much larger number of polymorphic fragments per primer than any other marker system used such as RFLP or microsatellites. ISSRs are regions that recline within the microsatellite repeats and offer great potential to determine intra-genomic and inter-genomic diversity as compared to other arbitrary primers, since they reveal variation within unique regions of the genome at several loci simultaneously. Several property of microsatellite such as high variability among taxa, ubiquitous occurrence and high copy number in eukaryotic genome make ISSRs extremely useful marker for variability analysis (Morgante et al., 2002) (Figure 1).

In this study, bulk sampling approach was chosen, since it permits representation of the vast accession by optimum number of plants. Yang and Quiros (1993)
Figure 2. UPGMA based dendrogram for 6 *L. sativum* populations using 4 ISSR (2 di, 1 penta and 1 tetra nucleotide) primers.

Figure 3. Two dimensional representation of principal coordinate analysis of genetic relationships among 85 accessions of *L. sativum* accessions using ISSR data.
reported that bulked samples with 10, 20, 30, 40 and 50 individuals had resulted in the same RAPD profiles as that of the individual plant constituting the bulk sample. Gilbert et al. (1999) also reported that pooling of DNA from individuals within accessions is the most appropriate strategy for assessing large quantities of plant material and concluded that 2-3 pools of five genotypes is sufficient to represent the genetic variability within and between accessions in the lupin and similar collections. Edossa et al. (2010) used bulked samples for diversity assessment in lentil collected from Ethiopia. The technique revealed higher genetic diversity, and, therefore, validated the usefulness of bulk sample analyses. Dagmawi (2011) also used bulked sample in germplasm diversity study of sesame populations, and found moderate genetic diversity of both Ethiopian and exotic populations.

The present study shows that out of 53 loci generated by four primers, two di, one penta and one tetra; 43 of them were polymorphic with 81.13% polymorphism. In regions based analysis, Amhara and Tigray showed higher percent polymorphism (66.04%); while, SNNPR and Somali showed least polymorphism with 47.17 and 45.28%, respectively. The same patterns of diversity were observed with gene diversity and Shannon index. Generally, *L. sativum* populations from Amhara and Tigray showed higher diversity than the other regions.

Edossa et al. (2010) studied the morphological and molecular diversity of Ethiopian lentil (*Lens culinaris* Medikus) using four ISSR primers and found 59.57% polymorphism with higher percent variation attributed within populations (56.28%). Gezahegne et al. (2009) studied wild and cultivated rice species of Ethiopia using six ISSR primers and reported 38.3 and 28.3% polymor-
Phylogeny of wild and cultivar rice species, respectively. Moreover, higher proportion of genetic diversity was observed within populations of rice (Gezahegne et al., 2009). Hence, the present study showed higher percent polymorphism and higher proportion of diversity within population of *L. sativum* comparable with that of Edossa et al. (2010) and Gezahegne et al. (2009).

In general, Amhara and Tigray had good genetic diversity than Oromia, SNNPR and Somali. But this has to be further studied using proper sampling strategy and multilocation comparison.

AMOVA analysis resulted in high genetic diversity within population (94%) and very low genetic diversity among population (6%). This could be due to high seed exchange among different regions and markets which could lead to intermix of populations between regions. Unlike other landraces of cultivated plants, *L. sativum* in Ethiopia is not restricted to a given area rather it is wildly exchanged among local community and markets. This shows that there is very high gene flow between populations and regions. Jiang et al. (2012) who studied the genetic diversity of *Chimonanthus grammatus* populations by using ISSR marker showed that there was 73.6% within population variation, whereas the rest 26.4% was due to among population variation. As compared to the present study, there was less gene flow. Jiang et al. (2012) recommended that gene flow, genetic drift and evolutionary history might have important influence on genetic structure and diversity of a given population.

*L. sativum* is both self and cross pollinated plant (Quirós and Cárdenas, 1998). Hence, the proportion of genetic variation is dependent on the type of pollination that the species undergoes. If the species has large proportion of cross pollination, then we expect high genetic variation within population and less divergence among populations. In addition to pollination, behavior of insects; market exchange could facilitate gene flow among regions which could result in higher percent variation within population and less genetic structure. This is also supported by the spread of individual accessions on UPGMA, NJ and PCO graphs.

Dendrogram of the present study by using UPGMA of Jaccard’s coefficient of similarity showed Amhara and Oromia populations of *L. sativum* were closely related. Based on this study, the samples with unknown origins could have been probably collected from southern part of Ethiopia since they closely clustered with the SNNPR population.

The Somali population had its own lineage far from the other populations and diverted as outliers. Genetic distance is a measure of the allelic substitutions per locus that have occurred during the separate evolution of two populations or species. Smaller genetic distances indicate a close genetic relationship whereas large genetic distances indicate a more distant genetic relationship. Crosses between distantly related individuals are expected to give better offspring than those between closely related genotypes. Therefore, prior knowledge of the genetic distance between genotypes or accessions is important in designing breeding program.

Genetic diversity of plant populations is largely influenced by factors such as reproduction system, genetic drift, evolutionary history and life history (Loveless and Hamrick, 1984). In broad-spectrum, outcrossing species have higher levels of genetic diversity than selfing and clonal plants (Rossetto et al., 1995).

**Conclusions**

Analysis of molecular variance for the accessions studied showed that the highest proportion of genetic variation was attributed to within population than among population. It is also highly significant. This confirms that there was a high level of gene flow and low level of genetic differentiation. Based on the UPGMA data, the Amhara, Tigray and Oromia accessions were clustered into one group, whereas the SNNPR and the unknowns in the other cluster. Samples from Somali formed a distinct cluster showing that it is distantly related to accessions from the entire regions.

**Conflict of interest**

The authors declare that they have no competing interests.

**ACKNOWLEDGEMENT**

The authors would like to thank Institute of Biodiversity Conservation for providing germplasms with full passport data of *L. sativum* assembled from Ethiopia. Technical support from Professor Klaus Mummenhoff and Mr. Mistiru Tesfaye during laboratory and field work is highly acknowledged.

**REFERENCES**


Quirós CF, Cárdenas RA (1998). MACA (Lepidium meyenii Walp.). University of California, USA.


Rehman N, Khan AU, Alkharfy KA, Gilani AH (2010). Pharmacological Basis for the Medical use of Lepidium sativum in Airways Disorders. Department of Biological and Biomedical Sciences, the Aga Khan University.


