

**THE SEQUENCE VARIATION OF CHLOROPLAST DNA INTERGENIC
SPACER REGION BETWEEN *trnT*(UGU) AND *trnL*(UAA) 5'EXON OF
ERAGROSTIS TEF (Zucc.) Trotter AND *E. PILOSA* (L.)**

**A Thesis Presented to
the School of Graduate Studies
Addis Ababa University,**

**In Partial Fulfillment of
the Requirement for the Degree of Master of Science
in Biology (Applied Genetics).**

By

Tileye Feyissa

May, 1999

ACKNOWLEDGEMENTS

I am greatly indebted to Prof. Endashaw Bekele, my research advisor, for identification of the research problem and his unreserved help and guidance in the course of the study.

I wish to express my gratitude to Prof. R. G. Herrmann, Botanic Institute, University of Munich, Germany, who allowed me to work in his laboratory and for his continuous guidance and advice.

I am very grateful to the co-operative staff of the Botanic Institute, University of Munich, Germany, in general and to Dr. Rainer Maier, Miss. Anya Drescher and Miss. C. Schraff in particular for their help, constructive suggestions and material assistance in the course of the study.

I would like to acknowledge with thanks the assistance of the Biology Department, Addis Ababa University, for its encouragement and kind permission for my study leave.

I would also like to thank Debrezeit Agricultural Research Centre, Ethiopia for providing me with the study seed samples.

Finally, I would like to express my gratitude to the German Academic Exchange Service (DAAD), Botanic Institute of the University of Munich and the NUFU project for the expenses incurred in this study.

TABLE OF CONTENTS

CONTENT	PAGE
Acknowledgements	i
Table of Contents	ii
List of Tables	iii
List of Figures	iv
Abstract	v
1. INTRODUCTION	1
2. LITERATURE REVIEW	5
3. MATERIALS AND METHODS	20
3.1 Plant Material	20
3.2 Total DNA Extraction	22
3.3 Chloroplast DNA Extraction	23
3.4 The Polymerase Chain Reaction (PCR)	24
3.5 Sequencing of the PCR Products	25
3.6 Calculation of Genetic Parameters	28
3.6.1 Genetic distance	28
3.6.2 The number of segregating sites	29
3.6.3 Heterozygosity	29
4. RESULTS	31
4.1 Variations in <i>Eragrostis tef</i>	32
4.2 Variations in <i>Eragrostis pilosa</i>	32
4.3 Interspecific Relationships Between <i>E. tef</i> and <i>E. pilosa</i>	33
5. DISCUSSION	49
6. CONCLUSION	53
7. REFERENCES	55

LIST OF TABLES

TABLE	PAGE
Table 1. The list of the 35 cultivars of <i>E. tef</i> and the two lines of <i>E. pilosa</i> studied.	21
Table 2. Sequence of the six universal primers for the amplification and sequencing of three non-coding regions of cpDNA.	34
Table 3. The measure of genetic distance for two groups of <i>E. tef</i> and two groups of <i>E. pilosa</i> .	35

Table 3. The measure of genetic distance for two groups of *E. tef* and two groups of *E. pilosa*.

LIST OF FIGURES

FIGURE	PAGE
Figure 1. Genetic map of tobacco chloroplast DNA.	8
Figure 2. Positions and directions of universal primers used to amplify three non-coding regions of cpDNA.	36
Figure 3. The PCR product of the cpDNA for the region between <i>trnT</i> (UGU) and <i>trnF</i> (GAA) genes that have been amplified by the universal primers A and F.	37
Figure 4. The sequencing results that show the long strands of C's where variations are observed	38
Figure 5. The alignment of the sequences of the intergenic spacer region between <i>trnT</i> (UGU) and <i>trnL</i> (UAA)5' exon of cpDNA for all the 35 cultivars of <i>E. tef</i> and the two lines of <i>E. pilosa</i> .	40

ABSTRACT

This study was carried out to determine cpDNA sequences of the intergenic spacer region between *trnT*(UGU) and *trnL*(UAA)5' exon of *Eragrostis tef* and *Eragrostis pilosa* to establish the degree of similarities and variation of cpDNA at the intraspecific and interspecific level since this region shows high degree of mutation. The sequence of this region has been worked out for 35 cultivars of *E. tef* and two lines of *E. pilosa*. This was achieved by isolating total DNA by CTAB method and for one cultivar by isolating cpDNA and then followed by amplification of three non-coding regions which are adjacent to each other by PCR using universal primers of Taberlet *et al.* (1991) followed by sequencing of the PCR products. From the results of the sequences it has been found that the 33 cultivars of *E. tef* and one line of *E. pilosa* (*E. pilosa* 30-5) showed no variation. But two cultivars of *E. tef*, namely "Addisse" and "Key murrie", and one line of *E. pilosa* (*E. pilosa* 30-4) showed a deletion of one base. It is the same base that was deleted in all the three. The total number of bases for this intergenic spacer region is 400. But for cultivars Addisse, Key murrie, and *E. pilosa* 30-4 the total number of bases is 399.

1. INTRODUCTION

The genus *Eragrostis* is a member of the tribe Eragrosteae, subfamily Eragrostoideae, of the family Poaceae (Costanza *et al.*, 1979). This genus constitutes about 300 species of which 250 are estimated to occur in the Old World (Willis, 1966 cited in Jones *et al.*, 1978).

The most common somatic chromosome number in *Eragrostis* is $2n=40$ though there are rare species which have diverged from this such as *E. aethiopica*, *E. cilianensis* and *E. bicolor* $2n=20$, *E. heteromera* $2n=41$, *E. barrelieri* and *E. mexicana* $2n=60$, *E. curvula* $2n=20, 40, 60, 70, 80$ (Jones *et al.*, 1978).

Among the members of this genus, *E. tef* is the most produced cereal in Ethiopia (Melakehail Mengesha, 1966). Five possible progenitors were suggested for this cereal by many authors. These are *E. pilosa*, *E. macilenta*, *E. aethiopica* or *E. pseudo-teff*, and *E. longifolia* (Costanza *et al.*, 1979). After accessing biochemical analysis of *E. tef* and some wild *Eragrostis* species, Endashaw Bekele and Lester (1981) concluded that *E. tef* probably originated from *E. pilosa* and also pointed the affinities of *teff* to *E. aethiopica*, *E. barrelieri* and perhaps also to *E. curvula* and *E. cilianensis*. Morphologically, *E. pilosa* is the most closely related species to *E. tef*. Both species are tetraploid (Jones *et al.*, 1978). Except being diploid, *E. aethiopica* is closely related to *teff* and *E. pilosa*. *E. pilosa* has the widest distribution of these five species and probably originated in the Mediterranean area (Rozhevitis, 1928 cited in Costanza *et al.*, 1979). It is annual and characterized by

extensive morphological diversity. In Ethiopia, this species is sympatric with tef and is a likely candidate for domestication (Costanza *et al.*, 1979).

E. tef (Zucc.) Trotter is the major cereal crop of Ethiopia that is adapted to a great diversity of climates and soil types. It is cultivated at altitudes ranging from 300 to 2800 m.a.s.l. although the best performance occurs at altitudes of 1700 to 2400 m.a.s.l. The varieties of tef grown at different altitudes may be different (Tadesse Ebba, 1969). According to Endashaw Bekele and Lester (1981) and Endashaw Bekele *et al.* (1995), this species is initially domesticated in Ethiopia from *E. pilosa* which is being collected as food by people in many parts of Africa during the times of famine, supporting the suggestion of Costanza *et al.* (1979) that selective planting and harvesting could have produced tef from *E. pilosa* or an *E. pilosa* like plant through domestication. As a food crop tef is mainly restricted to Ethiopia and covers about 30% of the total cultivated area (Endashaw Bekele *et al.*, 1995). This shows that the domestication of tef started during the time of food scarcity (Endashaw Bekele, 1978) and suggests that tef can resist different adverse environmental conditions. Tef grows in U.S.A, Australia, Pakistan and South Africa nowadays both for human consumption and animal feeds.

Tef originated in Ethiopia and widely distributed during the 19th century to India, South Africa, Australia, and British Guiana by military and political agents of British Empire (Tadesse Ebba, 1969) and is being grown as a forage crop in these countries. It is primarily grown in Ethiopia as human food, but sometimes used to prepare local alcoholic

drinks. In addition, the straw is a valuable animal feed mainly during the dry seasons and also as building materials since it sticks mud together.

Tef is a self-fertile sexually propagated annual that shows occasional outbreeding (Ponti *et al.*, 1978 cited in Endashaw Bekele and Lester, 1981). It is morphologically variable and the classification of cultivars is based on the degree and kind of pigmentation of the caryopses, glumes and vegetative parts, on panicle morphology, and on life cycle duration (Costanza *et al.*, 1979).

Endashaw Bekele (personal communication) studied the chloroplast DNA (cpDNA) using some restriction enzymes but the results indicated very limited polymorphism on the cultivars of tef that were studied. This was further carried on with the study of the non-coding regions of cpDNA, the 18S rRNA, and the transcription factor VPI sequence variation in tef by Endashaw Bekele *et al.* (in press). In their study on six tef cultivars, no intraspecific variation in the two chloroplast *trnL* intron and *trnL-trnF* spacer were found.

The general objective of this study is:

- to determine the intraspecific, interspecific and evolutionary relationships among *E. tef* and *E. pilosa*.
- to see the comparative rate of differentiation within and between these species using the presently and past studied biochemical, molecular and morphological characters.

The specific objective is: the sequencing of the intergenic spacer region between *trnT*(UGU) and *trnL*(UAA) 5' exon for 35 cultivars of *E. tef* and two entries of *E. pilosa* (*E. pilosa* 30-4 and *E. pilosa* 30-5). The data generated were analyzed in order to see the rate of intraspecific and interspecific polymorphism. The analysis was carried out by calculating genetic distance, the average number of pairwise nucleotide differences among DNA sequences and heterozygosity (nucleotide diversity).

2. LITERATURE REVIEW

Endashaw Bekele and Lester (1981) carried out quite a substantial work on the morphological and biochemical relationships of wild *Eragrostis* species with *E. tef*. They studied eleven accessions that had been chosen to represent the 35 cultivars of *E. tef* and 14 accessions of wild *Eragrostis* species, and attempted to identify the degrees of affinity among these species. These authors detected a wide range of variation in flavonoids and in proteins both amongst the wild species and the cultivars though the distinction between the two groups is fairly good. They considered *E. aethiopica* as a possible ancestor of *E. pilosa* and *E. tef* because *E. aethiopica* is placed between the cultivars and the wild species by Principal Component Analysis of the flavonoid data. By using flavonoid patterns, caryopsis and pollen morphology as markers, Endashaw Bekele (1986) studied the relationships between wild and cultivated species of *Eragrostis*. He found that groupings of these species by using morphological data does not fit well with that using flavonoid characters. The author suggested that this is due to the differences in the laws governing molecular evolution and phenotypic evolution, and pointed out that giving differential weights to different flavonoid characters could produce the already established groupings. In addition to this, the author also suggested that markers in biochemical speciation need to be established first.

After collecting 124 tef head selections from major tef producing areas of Ethiopia, Melakehail Mengesha *et al.* (1965) studied different characters and found significant

variations in plant height, panicle length, maturity, seed colour, seed yield, lodging and panicle type. However, insignificant variation was observed among different progenies originating from the same selection, suggesting self-pollination is a general rule for the species. Panicle length, plant height and seed yield were positively correlated. A significant negative correlation was obtained between panicle length and lodging resistance.

Tadesse Ebba (1975) classified *E. tef* into 35 cultivars by combining different characteristics. These cultivars are placed into six complexes based on morphological and qualitative characters:

- Very compact panicles and white or medium-brown grain.
- Fairly loose panicles, obpyramidal or with branches unequally ramified, and white grain.
- Very loose panicles and medium-brown grain.
- Spikelets and/or foliage dark red or purple.
- Plants with quick maturity and poor yields
- Light yellow-green to grey-green lemmas and white to orange-white seeds.

But more cultivars have been collected since then both by Tadesse Ebba and other researchers from remote parts of Ethiopia (Endashaw Bekele and Lester, 1981). Costanza *et al.* (1979) classified tef into three groups from 36 accessions based on 34 vegetative characters. By analyzing morphological characters and pigmentation of *E. tef* from 14

major sites in Ethiopia, Endashaw Bekele (1996) detected that each of the sites contained a very large number of different genotypes and suggested these regions for *in situ* gene conservation activities of tef.

Endashaw Bekele *et al.* (1995) extracted albumin, globulin, and prolamine fractions from seeds of 37 lines or cultivars of *E. tef* and observed the polymorphism of the seed proteins in these cultivars. From the result of this study they concluded that there is the possibility of discriminating many cultivars and lines of tef based on seed protein polymorphism. But since many of the lines are indistinguishable, these authors suggested the need for a joint inclusions of morphological and molecular data to increase the level of discrimination.

Considerable differences in amino acid composition between tef and wild *Eragrostis* species were found by Lester and Endashaw Bekele (1981) but *E. pilosa* was relatively similar to *E. tef*, especially in having more lysine than the rest of the wild species. In spite of the fact that many workers in the area agree on the view that *E. pilosa* was the wild progenitor of *E. tef*, the wild species which were the ancestral diploids of these two tetraploids remains unanswered.

Mulu Ayele (1994-1995) explained the existence of significant genotype X environment interaction after carrying out combined analysis of variance for all environments that have been worked in. Hailu Tefera *et al.* (1990) concluded that phenotypic variability for

Variation of nuclear ribosomal RNA genes were studied among 314 plants representing 28 accessions of *E. tef* by using 10 restriction enzymes (Pillay, 1997). In the same experiment the structural similarity in rDNA restriction map of *E. tef* with other grass species in the location of restriction sites and its size, especially for the coding region, were observed. However, in the intergenic spacer regions of some accessions a single restriction site polymorphism, a loss of *Sacl* site was detected. The author suggested that since the 18S region exhibited differences as compared to other grasses, this region may be of diagnostic value for the species. This is the other indicative of the existence of molecular genetic variation within this species. Wang *et al.* (1992) pointed out that even for morphologically very similar species the genetic variation study at the molecular level, particularly nuclear RFLP can provide a high intraspecific variation. Vosman *et al.* (1992) after identification of highly polymorphic nuclear DNA regions in tomato, suggested that fingerprints can be used to determine genetic relationships between cultivars since closely related cultivars show more similarity in fingerprint pattern than cultivars of more remote types.

Hailu Tefera and Peat (1996a) studied gene action for some quantitative traits of *E. tef* and found that grain yield was controlled by additive and dominance x dominance form of gene interaction. Tareke Berhe *et al.* (1989) after making a cross between "Fesho" and "Key murrie" and getting additive effects for brown pigment in F2, reported that seed colour is controlled by duplicate genes, B1 and B2. Hailu Tefera and Peat (1996b) evaluated selection methods for grain yield in the F2 and F3 generations of *tef* and concluded that segregating materials of *tef* could be handled by a combination of modified

bulk population method and modified pedigree method of breeding. For tef breeding program emphasis has been given to the seed yield-related traits. However, due to its great importance straw yield should also be considered (Mulu Ayele *et al.*, 1994-95). Since tef is endemic to Ethiopia, its major diversity is found only in this country (Seyfu Ketema, 1997). Therefore, the genetic improvement of tef primarily depends on the germplasm resource available in the country. However, satisfactory efforts are not made to collect and utilize these resources (Kebebew Assefa *et al.*, in press).

The breeding program is successful if there is genetic variation and there are possibilities to select that variation (Shah *et al.*, 1994). Therefore, although intraspecific variation is useful to get high yield producing and disease resistant variants, the wild *Eragrostis* species should also be considered to maximize the benefits. Once the relationships between the wild species and the cultivars are constructed, the selection and screening will be easier.

CpDNA restriction fragment length polymorphism (RFLP) and sequence analyses are widely used to study variations among plants (Soltis *et al.*, 1989). The presence of this unique DNA molecule in chloroplast has led many scientists to study its structure and expression intensively. These studies are highly accelerated with the development of gene cloning and DNA sequencing techniques that have been introduced in the mid 1970s (Sugiura, 1992). Bedbrook and Bograd (1976) constructed the first physical map of

cpDNA for maize and Bedbrook *et al.* (1977) cloned the chloroplast gene for the first time. These studies and others established a new field, "chloroplast molecular biology", and the organization and expression of these genomes were among the most extensively studied fields in plant molecular biology (Sugiura, 1992). The cpDNA of vascular plants is circular, and most are 120-180 Kb in size. This is small compared with the genome of the simplest free-living prokaryote. One explanation for this is that the evolutionary history of the chloroplast precursor since its first initial endosymbiosis, have involved many genes from the prokaryote to be transferred to the nuclear genome where they have become active (Weeden, 1981 cited in Ayliffe and Timmis, 1992). Alternatively, the nuclear genes which control chloroplast development were at first duplicated in the cell as a result of the endosymbiosis and those in the ancestral chloroplast were lost (Ayliffe and Timmis, 1992). Chloroplasts contain 22-900 genome copies that are found in association with the inner membrane of the envelope or with internal thylakoid membranes in aggregates of 10-20 DNA molecules (Mullet, 1988).

These DNA molecules enabled chloroplasts to harbour their own translational apparatus including rRNA and tRNA species, all of which are chloroplast-encoded (Sugiura, 1992). Although these chloroplasts possess their own genetic system, and a number of their components are encoded in their genomes (Shimida and Sugiura, 1989), chloroplast biogenesis needs the cooperation of nuclear and chloroplast genetic systems (Gordon *et al.*, 1981). The coding potential of chloroplast genome is modest as compared to that of the nuclear genome but it plays an essential role in the biogenesis of the organelle.

Available biochemical data suggest that the chloroplast genome encodes part of the energetic and translational processes of the organelle. Moreover, when interspecific genome/cpDNA hybrids are constructed even from closely related species, serious developmental disturbances may be observed. Interspecific genome/cpDNA hybrids provide several unique insights for the study of genome /cpDNA interactions, particularly concerning their specificity and evolution (Gordon *et al.*, 1981).

The simplicity of gene cloning, the lower rate of silent substitutions, and the availability of a number of data made cpDNA molecule to be more suitable for classification of cytoplasms and phylogenetic study among divergent plant species (Wolfe *et al.*, 1987). The sequences of cpDNA plays a vital role in determining the evolutionary history of organisms (Fitch and Margoliash, 1967). However, many methodological problems had limited the study of the determination of DNA sequences until the past decade.

The introduction of the polymerase chain reaction (PCR) technique in the mid-1980s solved these methodological problems (Clegg, 1993). This new technique enabled the amplification of specific regions of the DNA which could then be sequenced by dideoxy nucleotide sequencing. This enabled the production of a great wealth of sequence data. Many scientists, particularly students of plant evolution, have been quick to exploit the advantages. They used this opportunity to produce a detailed phylogeny of the seed plants based on sequences of the chloroplast-encoded large sub unit of ribulose-1,5-bisphosphate carboxylase (*rbcL*) (Clegg, 1993). Although sequences of this gene are more preferred

to study phylogenetic relationships among divergent plants (Shinozaki *et al.*, 1986; Golenberg *et al.*, 1990), genes found in different regions on the physical map show a different mode and tempo of sequence evolution (Ogihara *et al.*, 1991).

Sequence analysis of cpDNA was first achieved in 1986 when the entire sequences of this molecule from tobacco (*Nicotiana tabacum*; Shinozaki *et al.*, 1986) and the liverwort, *Marchantia polymorpha* (Ohyama *et al.*, 1986a), were reported. Since then the complete sequences of cpDNA from rice (*Oryza sativa*; Hiratsuka *et al.*, 1989), [*Euglena gracilis* (Hallik *et al.*, 1993), black pine (*Pinus thunbergii*; Wakasugi *et al.*, 1994), root parasite *Epifagus virginiana* (Wolfe *et al.*, 1992)] cited in Maier *et al.* (1995) and maize (*Zea mays*; Maier *et al.*, 1995) have been determined. The sequences of specific regions of cpDNA for many plants have also been done, but mostly the identification and expression analysis of many chloroplast genes have been done with several representative higher plants and green algae (Sugiura, 1992).

From comparative studies of cpDNAs based on restriction fragment and sequence analysis it is found that cpDNAs from most plants are composed of a large single-copy (LSC) and a small single-copy (SSC) region separated from each other by a large inverted repeat (IR) which is present in two copies per genome. The presence of this IR (10-30 Kb) results in a doubling of the rRNA genes and any other genes included within this region. A great deal of the chloroplast genome size variation between species is due to the variation in the size of the IR. Although this IR is common to most higher plants,

several species of the Viciae tribe of the legume family possess only one copy of this region (Palmer and Thompson, 1982; Ogiwara *et al.*, 1992).

From restriction analysis and nucleotide sequencing, cpDNA is found to be conservative in nature as compared to nuclear DNA (Palmer *et al.*, 1985; Wolfe *et al.*, 1987). As a result, both RFLP and sequence comparison of cpDNA have provided a powerful tool that is used to investigate plant phylogenetic studies at higher taxonomic levels (Clegg, 1993). Studies of RFLP have produced restriction maps for the whole cpDNA of many species and this enabled the studies of gross structural evolution. But many reports show that a number of mutations have been observed in the chloroplast genome. These are structural changes such as inversions, rearrangements of gene order, and insertions/deletions, as well as base substitutions (Ogiwara *et al.*, 1988). The restriction enzyme recognition-site changes appear to have equal distribution along the chloroplast genome except for the IR, where alterations are rarely observed (Palmer *et al.*, 1985). However, length mutations occur more frequently at specific sites on the genome. This suggests that there are "hot spots" that generate variation indicating that some directed mechanism operates to introduce the rearrangements in the chloroplast genome (Palmer *et al.*, 1985). One possibility is that AT-rich regions in the genome are associated with the recombination hot-spots which cause genome instability (Shih *et al.*, 1984). Some additional characteristic features of hot spot were found from the precise sequence analysis of the region. Furthermore, a large number of deletions/insertions have also accumulated in the region. The region is highly AT-rich and contains many direct and

inverted repeats that may contribute to the instability of this segment of the chloroplast segment (Ogihara *et al.*, 1992). These authors discussed the possible mechanisms of cpDNA rearrangements and the significance of short repeated sequences for structural alterations of the chloroplast genome based on the above analyses. Moreover, they have examined the possibility that the rearranged DNA segment had been translocated to other position(s) both inside the chloroplast genome as well as to other organelles.

In both of these structural types, i. e, possessing and lacking the IR, genome rearrangements have led to a considerable variety of chloroplast genomes. Therefore, very similar gene organizations are expected to be found only in the chloroplast of very closely related plant species (Bonnard *et al.*, 1985). Some evidences show that the frequency of cpDNA rearrangement is higher in genomes lacking IR than those possessing the IR. Therefore, this IR structure maintains the structural stability of the chloroplast genome. Broad bean and pea are phylogenetically closely related but since their cpDNA lacks the IR, the gene organization differs considerably in these species (Palmer and Thompson, 1982).

Most of the rearrangements in chloroplast genomes are due to inversions and deletions (Bonnard *et al.*, 1985). In several cases, however, especially in broad bean and pea, duplicated genes (which is the case for some tRNA genes) are present (Mubumbila *et al.*, 1984, 1985 cited in Bonnard *et al.*, 1985). By making comparison of chloroplast gene organization and nucleotide sequences from different higher plants, the theory that

different chloroplast genomes have evolved by rearrangements of a common ancestor DNA molecule have been confirmed (Bonnard *et al.*, 1985). But the mechanisms of structural alteration of the cpDNA has still remained unknown (Ogihara *et al.*, 1988).

To determine the entire coding potential of the cpDNA, a complete sequence determination is required (Hiratsuka *et al.*, 1989). Therefore, from complete sequences of some chloroplast genomes and the identification of genes in many plants it is estimated that the genome codes for approximately 123 genes. These genes are classified into three groups. The first group includes the genes encoding gene products involved in photosynthesis. The second group consists of genes required for transcription, translation, or replication of the chloroplast genetic information. The third genes include putative sequences for an NADH oxidoreductase complex and 19 open reading frames (more than 100 codons of unknown function) (Ohyama *et al.*, 1986b). Without the knowledge of the function of the unidentified open reading frames (ORFs) no final conclusions can be made concerning the full coding potential of the chloroplast genome (Mullet, 1988). Many of these ORFs have been identified by comparing their predicted translation products with known sequences of chloroplast or prokaryotic proteins. However, several ORFs could not be identified. But by using evolutionary filtering (Wolfe *et al.*, 1987) it is possible to identify those unidentified ORFs probably that code for genuine chloroplast proteins.

The determination of DNA sequences from two or more chloroplast genomes followed by comparison of the corresponding regions provide us more accurate information on the rearrangements that have occurred during evolution (Zurawski and Clegg, 1987 cited in Ogihara *et al.*, 1988). Since sequences which are conserved over large evolutionary distance presumably may perform some necessary functions, they are inferred to be evolving under constraint (Hiratsuka *et al.*, 1989). From the world-wide agricultural importance of monocots, especially cereals, and their evolutionary distance from dicots, after determination of the complete sequence of the rice cpDNA, Hiratsuka *et al.* (1989) compared it to other complete cpDNA sequences in order to identify conserved and non-conserved genes. In addition, these authors have found that rice cpDNA shows similar rearrangements observed in other grasses.

Since most of the cpDNA sequences are concentrated on the coding regions, especially the *rbcL* gene, much is known about the evolution of chloroplast genes and genome structure but little about the sequence changes in various regions of the genome particularly the non-coding sequences. Detailed studies of nucleotide divergence are limited except for the comparisons of entire cpDNAs (Zurawski *et al.*, 1984). Intergenic spacer regions have been omitted from the sequence comparisons, mainly because sequences have completely changed among divergent plants such that the processes of sequence changes can not be precisely traced. From this point of view, sequence comparisons among closely related cpDNAs give a vital information because their sequences have not greatly diverged. Thus, this allows study of the evolutionary

constraints operating on both the genes and the intergenic spacer region (Ogihara *et al.*, 1991).

The intergenic spacer region is the most variable portion of the rDNA cistron and is sometimes referred to as the nontranscribed spacer (NTS), a misnomer, since large portions of it are transcribed (Bonnard *et al.*, 1985). The 18s and 26s rDNA are conserved and hence have phylogenetic importance at higher taxonomic levels. In contrast to this, intergenic spacer sequences often contain sufficient variation thus allowing investigation of genetic relationships between plant varieties, cultivated varieties (cultivars), populations, and individuals (Nickrent and Patrick, 1998). This region shows a great variation among different plant groups ranging in size from about 1 to 12 Kb (Nickrent and Patrick, 1998). This length variation results from at least one, and often many, tandem or dispersed subrepeat domains. The advent of PCR greatly facilitating broad-scale sequencing studies of many genes; however, the length of plant intergenic spacer sequences and the lack of universally conserved sites among different species within intergenic spacer regions may explain the relative paucity of these sequences (Nickrent and Patrick, 1998).

Morton and Clegg (1993) stated that the cpDNA of *Oryza sativa* contains 31% of non-coding region. Since these non-coding regions exhibit highest frequency of mutation, both in terms of nucleotide substitution and insertion/deletion mutations, such regions may be useful in determining relationships at both the interspecific and intraspecific levels. By

amplification of the region of interest by PCR technique followed by sequencing or restriction digestion of the PCR product, it is possible to carry out such studies.

Taberlet *et al.* (1991) designed three pairs of universal primers (Table 2) for amplification of the following three non-coding regions of cpDNA that are located in the large single copy region.

- An intergenic spacer between *trnT* (UGU) and the *trnL* (UAA)5'exon,
- the *trnL* (UAA) intron, and
- intergenic spacer between the *trnL* (UAA)3'exon and *trnF* (GAA).

Each of these non-coding regions are flanked by conserved sequences of the tRNA genes and the primers are placed on the conserved regions and ultimately amplification is accomplished. Vogel *et al.* (1998) amplified the intergenic spacer *trnL* (UAA)-*trnF* (GAA) by using universal primers E and F for broad bean.

3. MATERIALS AND METHODS

3.1. Plant Material

The seeds of 35 cultivars of *E. tef* and two lines of *E. pilosa* were obtained from Debrezeit Agricultural Research Centre, Ethiopia. They are listed in Table 1. The plants were grown in the green house for about 2-3 weeks for the total DNA isolation, and for 4-6 weeks to get enough leaf material, for cpDNA isolation.

Table 1. The list of the 35 cultivars of *E. tef* and the two lines of *E. pilosa*.

<u>Abbreviation</u>	<u>Name</u>	<u>Code</u>
Ep4	<i>E. pilosa</i> 30-4	
Ep5	<i>E. pilosa</i> 30-5	
cv1	Ada	T-58
cv2	Addisse	T-54
cv3	Adonsis	T-9
cv4	Alba	T-23
cv5	Balemi	T-49
cv6	Betten	T-45
cv7	Bunniye	T-108
cv8	Bursa	T-87
cv9	Curati	T-119
cv10	Dabbi	T-1
cv11	Denkeye	T-102
cv12	Dschanger	T-27
cv13	Enatit	T-61
cv14	Fesho	T-51
cv15	Gea lamie	T-111
cv16	Gommadie	T-30
cv17	Gorade	T-105
cv18	Gotarie	T-101
cv19	Hamarawi murri	T-106
cv20	Hatalla	T-135
cv21	Janno	T-116
cv22	Karadabi	T-3
cv23	Key agacham	T-6
cv24	Key murrie	T-71
cv25	Magna	T-24
cv26	Murrie	T-112
cv27	Purpurea	T-4
cv28	Rossea	T-47
cv29	Rubicunda	T-17
cv30	Shewagimira	T-140
cv31	Trotrianna	T-138
cv32	Tulunassi	T-69
cv33	Varigata	T-68
cv34	Viridis	T-109
cv35	Zuccanginna	T-139

3.2. Total DNA Extraction

The total DNA is isolated following the hexadecyltrimethylammoniumbromide (CTAB) method of Doyle and Doyle (1990). The leaf tissue was harvested and either used immediately or stored at -70 °C after freezing with liquid nitrogen. The isolation buffer containing 2% CTAB, 1.4M NaCl₂, 20 mM EDTA, 100 mM Tris-HCl, pH 8.0, was preheated in the microwave. After cooling of the isolation buffer, 2-Mercaptoethanol was added to a final concentration of 100 mM. 0.05 to 0.1g of fresh leaf tissue was ground to a powder in liquid Nitrogen in a clean mortar and pestle and mixed with 400 µl isolation buffer in 1.5 ml eppendorf tubes. After incubation of the crude extract at 60 °C for 30 min, equal volume (400 µl) of chloroform-isoamyl alcohol (24:1; v/v) was added. The tubes are vortexed and centrifuged at 10,000 rpm (Eppendorf centrifuge) for 10 min at 10 °C. The supernatant was transferred to new eppendorf tubes (1.5 ml). This extraction with Chloroform-isoamyl alcohol (24:1; v/v) was done twice. The final supernatant was transferred to new eppendorf tubes (1.5 ml) and mixed with 2/3 Vol. of cold isopropanol. The DNA was precipitated after leaving the sample at room temperature (RT) from 30 min to 1hr followed by centrifugation at 12,000 rpm for 20 min at 4 °C. To get more precipitate, the sample could be left at RT for several hours to overnight. The isopropanol was removed gently without losing the DNA precipitate and 1ml 70% EtOH was added and left in the refrigerator overnight followed by centrifugation at 10,000 rpm for 10 min at 4 °C. This is the washing step. The supernatant was poured off carefully and the DNA pellet was air dried at RT. The DNA

pellet was dissolved in 50 μ l of distilled deionized H₂O.

3.3. Chloroplast DNA Extraction

Chloroplast DNA was extracted using the isolation method of Palmer (1986) with some modifications. Prior to extraction plants were kept in the dark for 1-3 days to reduce chloroplast starch levels. Young healthy destarched leaves were cut into small pieces and about 30 g were placed in 150 ml of cold isolation buffer containing 330 mM sucrose, 50 mM Tris-Acetate, pH 7.2, 10 mM EDTA, 1 mM DTT (fresh), 0.5% PVP. The leaves in the isolation buffer were homogenized in 1 litre clean waring blender for 3-5 seconds at high speed. The homogenates were filtered through silk guaze and double layered Miracloth followed by centrifugation at 5,000 rpm for 5 min at 4 °C. The supernatant was discarded and the pellet was resuspended in 5 ml of washing buffer consisting of 330 mM sucrose, 50 mM Tris-Acetate, pH 7.2, 10 mM EDTA, 1 mM DTT using paint brush. The resuspended pellet was loaded onto sucrose step gradients consisting of 15 ml 60% sucrose, 25 mM Tris-HCl, pH 8.0 which has been overlaid by 15 ml 20% sucrose, 25 mM Tris-HCl, pH 8.0 by using sucrose gradients mixer and centrifuged in an ultracentrifuge (SW-28rotor) at 24,000 rpm for 45 min at 4 °C. The green band of chloroplast at 20-60% sucrose interface was removed with a wide bore pipette and then diluted with 2 Volumes of 25 mM Tris-HCl, pH 8.0. After centrifugation at 5000 rpm for 5 min at 4 °C, the pellet was resuspended in TE buffer containing 50 mM Tris-HCl, pH 8.0, 20 mM EDTA and SDS was added to a final

concentration of 1% followed by incubation at 60 °C for 10 min. Proteinase K (2 mg/ml) was then mixed in and incubated overnight. Equal volume of phenol was added and mixed by vortexing followed by centrifugation at 10,000 rpm for 5 min at 10 °C. The phenol was removed two times with equal volume of Chloroform-isoamyl alcohol (24:1 v/v) by centrifugation at 10,000 rpm for 5 min at 10 °C. The supernatant (upper phase) contains the DNA and the lower phase is phenol so that the upper phase is transferred to new eppendorf tube. The cpDNA is then precipitated with 0.7 volume of isopropanol, 0.1 volume of 10 mM NaOAc, 2.5 volume EtOH. The pellet was washed with 70% EtOH by centrifugation at 10,000 rpm for 10 min at 4 °C and allowed to air dry. Then dissolved in 100 μ l of distilled deionized H₂O.

3.4. The Polymerase Chain Reaction (PCR)

About 100 ng of DNA, 5 μ l 10X PCR reaction buffer (750 mM Tris-HCl, pH 9.0 at 25 °C, 200 mM (NH₄)₂SO₄, 0.1% (w/v) Tween 20), 5 μ l 25 mM MgCl₂, 5 μ l 2 mM dNTPs, 1 μ l 10 pmol/ μ l of universal primers A and F of Taberlet *et al.* (1991), 0.5 μ l Taq DNA polymerase (2 U/ μ l), and H₂O were added to a final volume of 50 μ l in 0.2 ml eppendorf tubes. The reaction assay was mixed by pipetting up and down. The adding and mixing were done on ice to prevent unspecific amplifications. The template DNA used was from the total DNA extracts except for one cultivar of *E. tef* (Rossea T-47). For this cultivar the template DNA used for amplification was from cpDNA extracts.

The PCR protocol was as follows: hot start at 94 °C for 3 min, followed by 30 cycles of denaturation at 94 °C for 20 sec, annealing at 49 °C for 40 sec with a ramping phase of 1 °C/sec and elongation at 72 °C for 1 min and 20 sec, final elongation step at 72 °C for 3 min and subsequent cooling to 10 °C. Then it was put on ice immediately. 10 μ l of the PCR product was mixed with 2 μ l of DNA loading buffer (60% glycerol, 0.1% Bromophenol blue, 0.1% Xylenecyanol in 1X TBE buffer) and analyzed on a 1% agarose gel that has been mixed with Ethidiumbromide, and then visualized under Ultraviolet light. A 1kb DNA Molecular Weight Marker was used as length standard (Fig. 2).

3.5. Sequencing of the PCR Product

About 200 ng of the PCR product was mixed with 0.7 μ l DMSO, 1 μ l of 1 pmol/ μ l of the fluorescent dye-labeled universal primer (Amersham Pharmacia Biotech) and distilled-deionized H₂O was added to a final volume of 14 μ l. The universal primer used for the sequencing was primer B. The primer was covered with aluminium foil and put on ice until used. Primers A and F were not used for sequencing because it is better to use a different primer for sequencing reaction than was used for the PCR.

The microtitre plate wells were labelled as A reaction, C reaction, G reaction and T reaction. A reagent, C reagent, G reagent, T reagent and formamide loading dye were removed from the Thermosequenase fluorescent labelled primer cycle sequencing kit with

7-deaza-dGTP (Amersham Pharmacia Biotech) and stored on ice until required. The contents of the reagents are as follow:

A reagent: Tris-HCl (pH 9.5), magnesium chloride, TweenTM 20, NonidetTM P-40, 2-mercaptoethanol, dATP, dCTP, 7-deaza-dGTP, dTTP, ddATP, thermostable pyrophosphatase and Thermosequenase DNA polymerase.

C reagent: Tris-HCl (pH 9.5), magnesium chloride, TweenTM 20, NonidetTM P-40, 2-mercaptoethanol, dATP, dCTP, 7-deaza-dGTP, dTTP, ddCTP, thermostable pyrophosphatase and Thermosequenase DNA polymerase.

G reagent: Tris-HCl (pH 9.5), magnesium chloride, TweenTM 20, NonidetTM P-40, 2-mercaptoethanol, dATP, dCTP, 7-deaza-dGTP, dTTP, ddGTP, thermostable pyrophosphatase and Thermosequenase DNA polymerase.

T reagent: Tris-HCl (pH 9.5), magnesium chloride, TweenTM 20, NonidetTM P-40, 2-mercaptoethanol, dATP, dCTP, 7-deaza-dGTP, dTTP, ddTTP, thermostable pyrophosphatase and Thermosequenase DNA polymerase.

After mixing by briefly vortexing the tubes, 1.5 μ l of the A, C, G, T reagents from Thermosequenase fluorescent labelled primer cycle sequencing kit with 7-deaza-dGTP (Amersham Pharmacia Biotech) were mixed with 3.5 μ l of the DNA, primer and DMSO

premix in the microtitre plate wells labelled as A, C, G and T, respectively. The reaction mix were mixed throughly by pipetting up and down two or three times with the micropipette. Then each of the microtitre plate wells containing the reaction components was overlayed with one drop of light mineral oil to prevent evaporation. All of these steps were done on ice.

The cycling reaction was performed according to the following protocol: hot start at 94 °C for 3 min, followed by 25 cycles of denaturation at 94 °C for 20 sec, annealing at 50 °C for 40 sec with a ramping phase of 1°C/sec and elongation at 72 °C for 1 min, final elongation step at 72 °C for 3 min and subsequent cooling to 10 °C. 6 µl of the stop solution (80% formamide in 1 X TBE) was added and denatured for 3 min at 94 °C followed by immediate cooling on ice.

The sequences were analyzed on 4% Sequagel XR, 66 cm (MWG-Biotech Europe) by loading 1.3 µl of the sequencing reaction on the polyacrlamide gel in the order of G A T C. This was performed using LI-COR Model 4200 sequencer.

3.6 Calculation of Genetic Parameters

3.6.1. Genetic distance

The number of shared nucleotide between the two cultivars (Addisse T-54 and Key murrie T-71) and other 33 tef cultivars and *E. pilosa* lines is scored as follows:

$$S = \frac{M_{xy}}{M_x + M_y}$$

Where:

M_{xy} = shared sequence

M_x = number of sequence in one group

M_y = number of sequence in the other group

S = proportion of shared nucleotides.

$$D = \frac{-\ln S}{r}$$

Where:

r = total number of nucleotide

D = genetic distance.

3.6.2. The Number of Segregating Sites

The average number of pairwise nucleotide differences among DNA sequences is defined as:

$$d = \frac{2 \sum dij}{n(n-1)}$$

Where:

d_{ij} = the number of nucleotide differences between sequence i and j

n = the number of DNA sequences sampled from the two groups of *tef* and *E. pilosa*

d = the number of segregating sites, S , which is the number of sites occupied by at least two different nucleotides.

3.6.3. Heterozygosity

$$H = 1 - \sum P_{ij}^2$$

Where:

P_j = the relative frequency of nucleotide j ($j = 1, 2, 3, 4$)
corresponding to nucleotide A, T, C and G.

H = heterozygosity (nucleotide diversity)

Unbiased estimate:

$$H = \frac{n}{n-1} (1 - \sum P_{ij}^2)$$

Where:

H = unbiased estimate.

P_{ij} = the observed frequency of nucleotide j at site i.

n = the number of nucleotide sequence.

4. RESULTS

The 35 cultivars of *E.tef* and the two lines of *E. pilosa* on which the sequencing of cpDNA intergenic spacer region between *trnT*(UGU) and *trnL*(UAA) 5'exon was carried out are listed in Table 1 according to their alphabetical order. The PCR products of the cpDNA for the region between *trnT* (UGU) and *trnF* (GAA) genes that have been amplified by universal primers A and F of Taberlet *et al.* (1991) are presented in Fig. 3. The intergenic spacer region that has been sequenced is diagrammed in Fig. 2. To sequence the intergenic spacer between *trnT* (UGU) and the *trnL* (UAA)5' exon, the universal primer B was used. The portions of conserved tRNA gene sequences flanking this region are compared with other published sequences (Shinozaki *et al.*, 1986; Ohyama *et al.*, 1986) and are found to be identical. The alignment of the sequences of this intergenic spacer region for the 35 cultivars of *E. tef* and the two lines of *E. pilosa* are presented in Fig.4. The first nucleotide of each sequence is the 41st nucleotide of the *trnT* (UGU) gene upstream from the intergenic spacer region. The alignment ends with the 10th nucleotide of the *trnL*(UAA)5'exon gene downstream from the intergenic spacer. The total number of nucleotides between the conserved regions [*trnT* (UGU) and *trnL* (UAA)5'exon gene is 400. But for two cultivars of *E. tef* (Addisse T-54 and Key murrie T-71) and *E. pilosa* 30-4 the total number of nucleotides for this intergenic spacer region is 399.

4.1. Variation in *E.tef*

The intergenic spacer region between *trnT*(UGU) and *trnL*(UAA) 5'exon of the cpDNA of 35 cultivars of *E. tef* were sequenced to estimate the level of intraspecific variation. No differences were observed between the sequences of 33 cultivars. In this study the intergenic spacer region between *trnT*(UGU) 5'exon showed that cultivars, Addisse T-54 (cv2) and Key murrie T-71 (cv24) have shown a deletion of one nucleotide at position of nucleotide number 228. From nucleotide position number 220 to 228 the 33 cultivars posses a long strand of 8 nucleotides of G's while the two cultivars posses 7 G's (Fig. 5).

4.2. Variation in *E. pilosa*

Two lines of *E. pilosa*, *E. pilosa* (30-4) and *E. pilosa* (30-5) were also sequenced. *E. pilosa* (30-5) exhibited identical sequences to the 33 cultivars of *E. tef*. The other line, *E. pilosa* (30-4) showed a deletion of one nucleotide at nucleotide position number 228, i. e, at the same position where deletion was observed in the two cultivars of *E. tef* (Addisse T-54 and Key murrie T-71). Both results indicate the existence of intraspecific variation in *E. pilosa* and *E. tef*.

4.3. Interspecific Relationships Between *E. tef* and *E. pilosa*

As has been observed from the results of the sequences, the 33 cultivars of *E. tef* have shown identical sequences to one line of *E. pilosa* (*E. pilosa* 30-5). The remaining two cultivars of *E. tef* (addisse T-54 and Key murrie T-71) are identical in their sequences to the other line of *E. pilosa* (*E. pilosa* 30-4).

The genetic parameters estimation has also indicated the close relationship of these two species. The measure of genetic distance for these groups are indicated in Table 3. The estimation for the number of segregating sites is 0.000012531. In addition the heterozygosity (nucleotide diversity) for the 33 cultivars of *E. tef* and *E. pilosa* 30-5 is 0.6869625 while for the two cultivars of *E. tef* and *E. pilosa* 30-4 is 0.686302224. The unbiased calculated estimate is 0.68868421 and 0.688026601 for the 33 cultivars of *E. tef* and *E. pilosa* 30-5 and the other two cultivars of *E. tef* and *E. pilosa* 30-4 respectively.

Table 2. Sequences of the six universal primers for the amplification and sequencing of three non-coding regions of cpDNA. The code denotes the 3'-most base pair in the published tobacco cpDNA sequence (Shinozaki *et al.*, 1986). The B and A in the code refer to each strand of DNA (from Taberlet *et al.*, 1991).

Name	Code	Sequence 5'-3'
a	B48557	CATTACAAATGCGATGCTCT
b	A49291	TCTACCGATTTCGCCATATC
c	B49317	CGAAATCGGTAGACGCTACG
d	A49855	GGGATAGAGGGACTTGAAC
e	B49873	GGTCAAGTCCCTCTATCCC
f	A50272	ATTGAACTGGTGACACGAG

Table 3. The measure of genetic distance for two groups of *E. tef* and the other two groups of *E. pilosa*.

Groups	A	B	C	D
A	—	0.00069	0.00069	0.00066
B	0.00069	—	0.00069	0.00069
C	0.00069	0.00069	—	0.00069
D	0.00066	0.00069	0.00069	—

A = the 33 cultivars of *E. tef*

B = the 2 cultivars of *E. tef* (Addisse and Key murrie)

C = *E. pilosa* 30-4

D = *E. pilosa* 30-5

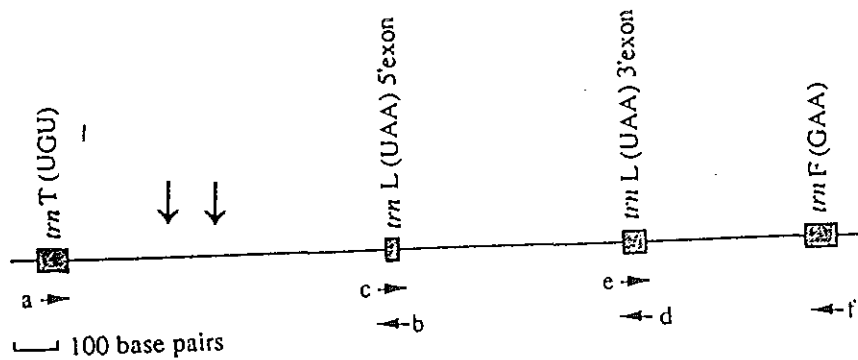


Fig. 2. Positions and directions of universal primers used to amplify three non-coding regions of cpDNA. Tips of arrows indicate the 3'ends of the primers. The expected size of PCR products of tobacco (Shinozaki *et al.*, 1986), rice (Hiratsuka *et al.*, 1989) and Marchantia (Ohyama *et al.*, 1986) respectively, are, in bp, 773, 833, 251 with primers a and b, 577, 614, 389 with c and d, 438, 324, 158 with e and f. The length of the non-coding regions in the figure corresponds to the tobacco sequence. (From Taberlet *et al.*, 1991).

↓ ↓ Indicates region that is sequenced.

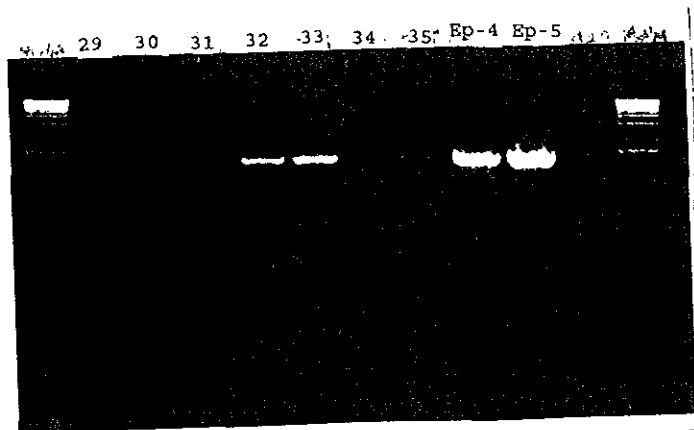
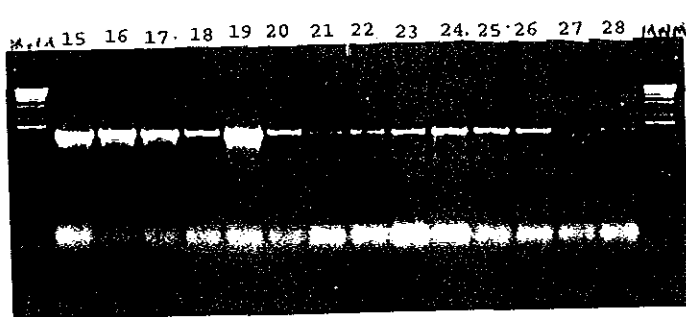
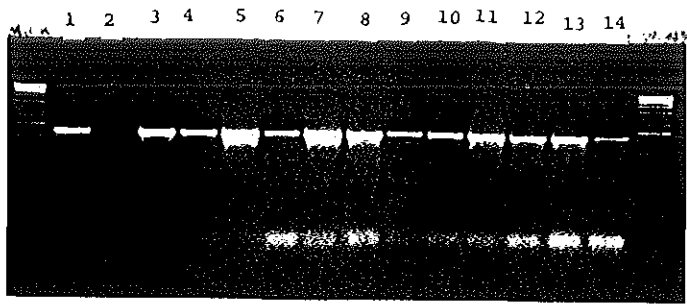


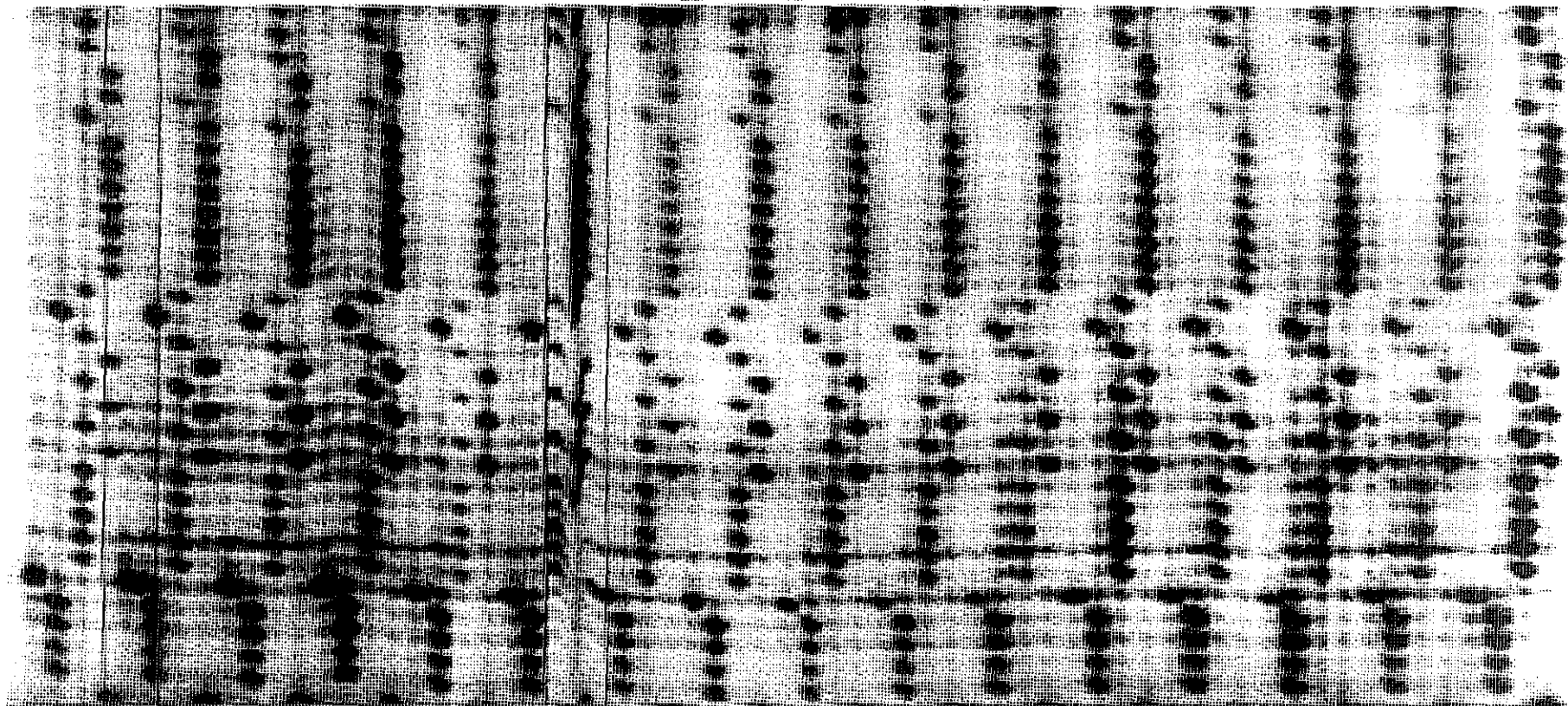
Fig. 3. The PCR product of the cpDNA for the region between *trnT* (UGU) and *trnF* (GAA) genes that has been amplified by the universal primers A and F.

The bands on the left and right are the molecular weight marker.

The numbers refer to the cultivars of *E. tef* according to their alphabetical order.

Ep-4 and Ep-5 refer to *E. pilosa* 30-4 and *E. pilosa* 30-5 respectively.

Ep-4 Cv1 Cv2 Cv3 Cv4 Cv5 Cv6 Cv7 Cv8 Cv9 Cv10 Cv11 Cv12 Cv13 Cv14 Cv15



38

0909700.eps

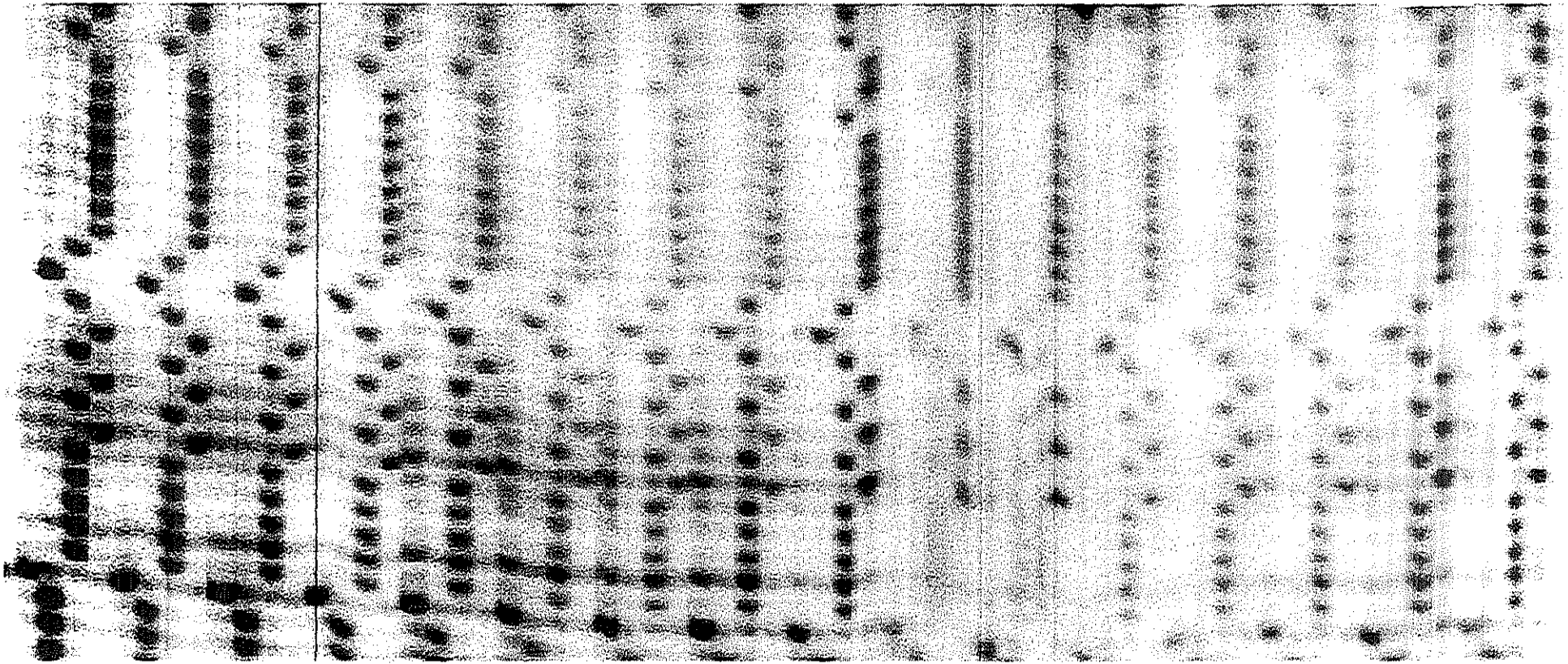
A.

Fig. 4. A portion of the sequencing results that show the long strands of C's where variations are observed. This region is the intergenic spacer between *trnT* (UGU) and the *trnL* (UAA) 5' exon. Since the universal primer B is used for sequencing of this region, the reading of the above sequence is from 5' to 3' direction of the complementary strand. The order of bases on the gel from left to right is G, A, T, C.

A. The 7 C's are observed for *E. pilosa* 30-4 (Ep4) and Addisse T-54 (Cv2).

B. The 7 C's can be observed for Key murrie T-71 (Cv24).

Cv16 Cv17 Cv18 Cv19 Cv20 Cv21 Cv22 Cv23 Cv24 Cv25 Cv26 Cv27 Cv28 Cv31 Cv32 Ep-5



39

B

1509700.eps

	1		↓
EP-4	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
EP-5	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv1	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv2	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv3	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv4	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv5	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv6	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv7	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv8	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv9	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv10	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv11	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv12	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv13	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv14	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv15	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv16	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv17	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv18	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv19	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv20	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv21	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv22	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv23	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv24	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv25	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv26	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv27	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv28	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv29	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv30	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv31	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv32	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv33	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv34	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA
Cv35	<u>CATTACAAAT GCGATGCTCT AACCTCTGAG CTAAGTGGGC</u>		TTACATAACA

trnT (UGU)

Fig. 5. The alignment of the sequences of the intergenic spacer region between *trnT* (UGU) and *trnL* (UAA) 5' exon for the 35 cultivars of *E. tef* and the two lines of *E. pilosa*.

The underlined sequences on the left side of the arrow are a portion of *trnT* (UGU) gene.

The sequences on the right side of the arrow are the intergenic spacer region.

201

EP-4 TAGAATACTC TAAAAAAGGA GGGGGGG.AT AGAGAGAAAA ACTTTTGGAT
EP-5 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv1 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv2 TAGAATACTC TAAAAAAGGA GGGGGGG.AT AGAGAGAAAA ACTTTTGGAT
Cv3 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv4 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv5 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv6 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv7 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv8 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv9 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv10 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv11 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv12 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv13 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv14 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv15 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv16 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv17 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv18 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv19 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv20 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv21 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv22 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv23 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv24 TAGAATACTC TAAAAAAGGA GGGGGGG.AT AGAGAGAAAA ACTTTTGGAT
Cv25 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv26 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv27 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv28 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv29 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv30 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv31 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv32 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv33 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv34 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT
Cv35 TAGAATACTC TAAAAAAGGA GGGGGGGGAT AGAGAGAAAA ACTTTTGGAT

Fig. 5. continued

401

EP-4	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT	↓	<u>GGGGATATGG</u>
EP-5	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv1	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv2	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv3	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv4	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv5	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv6	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv7	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv8	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv9	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv10	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv11	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv12	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv13	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv14	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv15	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv16	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv17	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv18	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv19	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv20	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv21	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv22	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv23	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv24	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv25	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv26	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv27	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv28	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv29	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv30	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv31	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv32	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv33	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv34	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>
Cv35	TACACAAGGA	ATCTGGGTTT	CAAAGAAAAA	AAAGGAAAAT		<u>GGGGATATGG</u>

trnL (UAA)
5'-exon

Fig. 5. continued

The last nucleotide on the left side of the arrow is the end of the intergenic spacer region sequence. The underlined sequences on the right side of the arrow is a portion of *trnL* (UAA) 5' exon gene.

5. DISCUSSION

The rate of sequence divergence by base deletion of the non-coding region of cpDNA included in this study occurs only at a single and the same position for both *E. tef* and *E. pilosa*. To study intraspecific and interspecific variations among angiosperms, selected regions as in this study are the most important regions. The chloroplast genome has a very high AT content in the non-coding region than the coding region (Morton and Clegg, 1993). From cluster analysis of chromatographic data, Endashaw Bekele and Lester (1981) arranged most of the *E. tef* accessions in a cluster which included *E. pilosa* in some of these subgroups and other wild species in other subgroups. Though a great deal of sequence homology has been observed both among the cultivars of *E. tef* and *E. pilosa*, it was possible to group these species into two based on these sequence results.

The first group consisting of the 33 cultivars of *E. tef* and *E. pilosa* 30-5. The second group containing the remaining two cultivars of *E. tef* (Addisse T-54 and Key Murrie T-71) and *E. pilosa* 30-4. Based on several morphological and biochemical data, Endashaw Bekele (1978) suggested polyphyletic origin of tef where different groups of tef have originated from related but different groups of wild *Eragrostis* species. The present cpDNA data where the 33 cultivars of tef show complete resemblance to each other on one hand and with one of the *E. pilosa* line, and the two tef cultivars show complete matching with the other *E. pilosa* line goes partly in conformity with Endashaw Bekele's (1978) suggestion, that these species are polyphyletic in origin.

This study has shown that, though cpDNA variation provides some information for screening and selecting genetically variable cultivars, the conservative nature of this molecule prevents to get more polymorphisms at intraspecific level. However, to study genetic variations at interspecific levels this molecule, especially the non-coding regions play an important role. Therefore, the usual assumptions that there is little or no intraspecific variability in the chloroplast genome (e.g. Harris and Ingram, 1991) goes in parallel with the results of this study on *E. tef* and *E. pilosa*. This agrees with the *trnL*(UAA) 5'exon, *trnL*(UAA) 3'exon and *trnF*(GAA) results of Endashaw Bekele *et al* (in press). These authors sequenced these two regions for six cultivars of *E. tef* and did not get any variation.

Ogihara *et al.* (1991) pointed out that the intergenic spacer regions have not been widely used for sequence comparisons because these regions show a great deal of variations among divergent plants such that the processes of sequence changes can not be precisely traced. Therefore, this region is quite informative to study the evolutionary constraints operating on both the genes and the intergenic spacer regions at inter and intraspecific levels, because the sequences at this level have not greatly diverged.

The different mode of inheritance of nuclear and organelle genomes leads the organelle genomes to have lower genetic diversity than nuclear genome even if the same mutation rates and selection pressures operate on both genomes (Thompson, 1977). The maternal inheritance of chloroplast genes, and their rapid segregation during vegetative growth made

the heteroplasmic individuals to be rare (Banks and Birky, 1985). This results in the effective number of chloroplast genes in a population of hermaphroditic plants equal to the number of individual plants, while it is twice the number of individual plants for nuclear genes; thus the expected heterozygosity for neutral alleles is half as great in organelle genes as in nuclear genes in the same population (Birky *et al.*, 1983).

For the analysis of chloroplast gene functions in higher plants cpDNA transformations are suitable (Kofer *et al.*, 1998). These authors have shown that mutagenesis of *ndh* genes affects photosynthetic electron flow and carbon metabolism.

From this study it could be suggested that the variations in *E. tef* and *E. pilosa* that was observed through the study of other molecular patterns such as rDNA, flavonoids, protein analysis and even morphological characteristics are indicative of polymorphism rather than cpDNA analysis. Conclusions from comparative analyses are based on the principle that conservation among homologous sequences is a reflection of biological importance (Thompson *et al.*, 1997). Since the cpDNA sequences of all the cultivars of *E. tef* have shown insignificant sequence variations, the complete cpDNA sequence for a single cultivar may represent all the members of *E. tef* species and even its closest wild *Eragrostis* species such as *E. pilosa*. Here it should be emphasized that the complete cpDNA sequence for this species can provide a significant information concerning cpDNA genes because these genes play a vital role in plant growth and metabolism, particularly in the process of photosynthesis (Kofer *et al.*, 1998). Moreover, it provides some

information about nuclear genes since some genes might have transferred from cpDNA to the nucleus during evolution of these plants (Martin and Herrmann, 1998).

6. CONCLUSION

The results of sequencing of the intergenic spacer region between *trnT* (UGU) and *trnL* (UAA) 5' exon of the 35 *E. tef* cultivars and the two lines of *E. pilosa* was made aligned. From the results a deletion of single base was observed for two cultivars of *E. tef* (Addisse T-54 and Key murrie T-71) and one line of *E. pilosa* (*E. pilosa* 30-4) on one hand and the identical sequence of the other 33 cultivars of *E. tef* and the other line of *E. pilosa* (*E. pilosa* 30-5) on the other hand. One group consists of the 33 cultivars of *E. tef* and *E. pilosa* 30-5 and the other group contains the two cultivars of *E. tef* (Addisse T-54 and Key murrie T-71) and *E. pilosa* 30-4. From this data it could be concluded that the origin of *E. tef* as well as *E. pilosa* is of species complex in nature and possibly polyphyletic. This confirms the suggestion of Endashaw Bekele (1978) that different groups of *E. tef* have originated from related but different groups of wild *Eragrostis* species. This sequencing result has also shown the close relationship of *E. tef* with *E. pilosa*.

This intergenic spacer region has been expected to show high rate of mutation as compared to the other region of chloroplast genome. Since the other region of this genome is relatively conserved, further work on these conserved regions to detect intraspecific variation has no importance. This seems to be true also for any other species. But working on the other wild *Eragrostis* species on sequencing of this intergenic spacer region as well as the other non-coding regions such as the *trnL* (UAA) intron, the intergenic spacer region between *trnL* (UAA) 3' exon and *trnF* (GAA) is very important. This may give information

on the evolutionary position of *E. tef*, *E. pilosa* and the other wild *Eragrostis* species.

7. REFERENCES

- Ayliffe, M. A. and Timmis, J. N. (1992). Tobacco nuclear DNA contains long tracts of homology to chloroplast DNA. *Theor. Appl. Genet.* **85**: 229-238.
- Banks, J.A., and Birky, C. W. (1985). Chloroplast DNA diversity is low in a wild plant, *Lupinus texensis*. *Proc. Natl. Acad. Sci. USA* **82**: 6950-6954.
- Bedbrook, J. R. and Bogorad, L. (1976). Endonuclease recognition sites mapped on *Zea mays* chloroplast DNA. *Proc. Natl. Acad. Sci. USA* **73**: 4309-4319.
- Bedbrook, J. R., Kolodner, R. and Bogorad, L. (1977). *Zea mays* chloroplast ribosomal RNA genes are part of a 22,000 bp inverted repeat. *Cell* **11**: 739-749.
- Birky, C. W., Maruyama, T. and Fuerst, P. (1983). An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. *Genetics* **103**: 513-527.
- Bonnard, G., Michel, F., Weil, J. H. and Steinmetz A. (1984). Nucleotide sequence of the split tRNA Leu (UAA) gene from *Vicia faba* chloroplasts: evidence for structural homologies of the chloroplast tRNA Leu intron with the intron from the autospliable *Tetrahymena* ribosomal RNA precursor. *Mol. Gen. Genet.* **194**: 330-336.
- Bonnard, G., Weil, J. and Steinmetz, A. (1985). The intergenic region between the *Vicia faba* chloroplast tRNA Leu (CAA) and tRNA Leu (UAA) genes contains a partial copy of the split tRNA Leu (UAA) gene. *Curr. Genet.* **9**: 417-422.

- Clegg, M. T. (1993). Chloroplast gene sequences and the study of plant evolution. *Proc. Natl. Acad. Sci. USA* **90**: 363-367.
- Costanza, S. H., de Wet, J. M. J. and Harlan, J. R. (1979). Literature review and numerical taxonomy of *Eragrostis tef* (T'ef). *Eco. Bot.* **33**: 413-424.
- Doyle, J. J. and Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus* **12**: 13-15.
- Endashaw Bekele (1978). Biochemical and morphological studies of the relationships *Eragrostis tef* and some other *Eragrostis* species. M.Sc thesis, University of Birmingham.
- Endashaw Bekele (1986). The relationship between wild and cultivated species of *Eragrostis* (Poaceae) using flavonoid patterns, caryopsis and pollen morphology as markers. *SINET: Ethiop. J. Sci.* **9**: 189-213.
- Endashaw Bekele (1996). Morphological analysis of *Eragrostis tef*: detection for regional patterns of variation. *SINET: Ethiop. J. Sci.* **19**: 117-140.
- Endashaw Bekele, Fido, R. J., Tatham, A. S. and Shwery, P. R. (1995). Heterogeneity and polymorphism of seed proteins in *tef* (*Eragrostis tef*). *Hereditas* **122**: 65-72.
- Endashaw Bekele and Lester, R. N. (1981). Biochemical assessment of the relationships of *Eragrostis tef* (Zucc.) Trotter with some wild *Eragrostis* species (Gramineae). *Ann. Bot.* **48**: 717-725.
- Endashaw Bekele, Marie, S., Holst, A., Jakobsen, K. S. and Nordal, I. (in press). The non-coding regions of cpDNA, 18S rRNA and the transcription factor, VPI sequence variation in *Eragrostis tef* (Zucc.) Trotter.

- Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C., Meng, B., Li, Y., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K. and Sugiura, M. (1989). The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.* **217**: 185-194.
- Jones, B. M. G., Ponti, J., Tavasoli, A. and Dixon, P. A. (1978). Relationships of the Ethiopian cereal *Eragrostis tef* (Zucc.) Trotter: Evidences from morphology and chromosome number. *Ann. Bot.* **4**: 1369-1373.
- Kebebew Assefa, Seyfu Ketema, Hailu Tefera, Ngugen, H. T., Blum, A., Mulu Ayele, Bai, G., Belay Simane and Tiruneh Kefyalew (1998). Diversity among germplasm lines of the Ethiopian cereal tef [*Eragrostis tef* (Zucc.) Trotter]. (in press).
- Kofer, W., Koop, H. U., Wanner, G. and Steinmueller, K. (1998). Mutagenesis of the genes encoding subunits A, C, H, J and K of the plastid NAD(P)H plastoquinone-oxidoreductase in tobacco by polyethylene glycol-mediated plastome transformation. *Mol. Gen. Genet.* **258**: 166-173.
- Lester, R. N. and Endashaw Bekele (1981). Amino acid composition of the cereal tef and related species of *Eragrostis* (Graminae). *Cereal Chem.* **58**: 113-115.
- Maier, R. M., Neckermann, K., Igloi, G. L. and Koessel, H. (1995). Complete sequence of the maize chloroplast genome: Gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *J. Mol. Biol.* **251**: 614-628.

- Martin, W. and Herrmann, R. G. (1998). Gene transfer from organelles to the nucleus: How much, What happens, and Why? *Plant physiol.* 118: 9-17.
- Melakehail Mengesha (1966). Chemical composition of Teff *Eragrostis tef* compared with that of wheat, barley and grain sorghum. *Eco. Bot.* 20: 268-273.
- Melakehail Mengesha, Picket, R. C. and Davis, R. L. (1965). Genetic variability and interrelationship of characters in teff, *Eragrostis tef* (Zucc.) Trotter. *Crop Science* 5: 155-157.
- Morton, B. R. and Clegg, M. T. (1993). A chloroplast DNA mutational hotspot and gene conversion in a noncoding region near *rbcL* in the grass family (Poaceae). *Curr. Genet.* 24: 357-365.
- Mullet, J. E. (1988). Chloroplast development and gene expression. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 39: 475-502.
- Mulu Ayele (1994-95). Comparison of optimum moisture environments against stress environments for developing draught resistant tef (*Eragrostis tef* (Zucc.) Trotter). *Acta Agronomica Hungarica* 43: 223-228.
- Nickrent, D. L. and Patrick, J. A. (1998). The nuclear ribosomal DNA intergenic spacers of wild and cultivated soybean have low variation and cryptic subrepeats. *Genome* 41: 183-192.
- Ogihara, Y., Terachi, T. and Sasakuma, T. (1988). Intramolecular recombination of chloroplast genome mediated by short direct-repeat sequences in wheat species. *Proc. Natl. Acad. Sci. USA* 85: 8573-8577.

- Ogihara, Y., Terachi, T. and Sasakuma, T. (1991). Molecular analysis of the hot spot region related to length mutations in wheat chloroplast DNAs. I. Nucleotide divergence of genes and intergenic spacer regions located in the hot spot region. *Genetics* **129**: 873-884.
- Ogihara, Y., Terachi, T. and Sasakuma, T. (1992). Structural analysis of length mutations in a hot-spot region of wheat chloroplast DNAs. *Curr. Genet.* **22**: 251-258.
- Ohyama, K., Fukuzawa, H., Kohchi, J., Shirai, H., Sano, T., Sano S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Auto S., Inokuchi, H. and Ozeki, H. (1986a). Complete nucleotide sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Plant Mol. Biol. Rep.* **4**: 148-175.
- Ohyama, K., Fukuzawa, H., Kohchi, J., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Auto S., Inokuchi, H. and Ozeki, H. (1986b). Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* **322**: 572-574.
- Palmer, J. D. (1986). Isolation and structural analysis of chloroplast DNA. *Methods in Enzymology* **118**: 167-187.
- Palmer J. D., Jorgensen, R. A. and Thompson, W. F. (1985). Chloroplast DNA variation and evolution in *Pisum*: patterns of change and phylogenetic analysis. *Genetics* **109**: 195-213.
- Palmer, J. D. and Thompson, W. F. (1982). Chloroplast DNA rearrangements are more frequent when a large inverted sequence is lost. *Cell* **29**: 537-550.

- Pillay, M. (1997). Variation of nuclear ribosomal RNA genes in *Eragrostis tef* (Zucc.) Trotter. *Genome* **40**: 815-821.
- Seyfu Ketema (1997). *Tef. Eragrostis tef* (Zucc.) Trotter. *Promoting the conservation and use of underutilized and neglected crops*. 12. Institute of Plant Genetics and Crop Plant Research, Gaterslaben/ International Plant Genetic Resources Institute, Rome, Italy.
- Shah, F. H., Rashid, O., Simons, A. J. and Dunsdon, A. (1994). The utility of RAPD markers for the determination of genetic variation in oil palm (*Elaeis guineensis*). *Theor. Appl. Genet.* **89**: 713-718.
- Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Sinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Dino, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H. and Sugiura, M. (1986). The complete nucleotide sequence of the tobacco chloroplast genome. *Plant Mol. Biol. Rep.* **4**: 110-147.
- Shih, C., Linial, M., Goodenow, M. M. and Hayward, W. S. (1984). Nucleotide sequence 5' of the chicken c-myc coding region: Localization of a noncoding exon that is absent from myc transcripts in most avian leukosis virus-induced lymphomas. *Proc. Natl. Acad. Sci. USA* **81**: 4697-4701.
- Shimada, H. and Sugiura, M. (1989). Pseudogenes and short repeated sequences in the rice chloroplast genome. *Curr. Genet.* **16**: 293-301.

- Soltis, D. E., Soltis, P. S., Ranker, T. A. and Ness B. D. (1989). Chloroplast DNA variation in a Wild plant, *Tolmiea menziesii*. *Genetics* **121**: 819-826.
- Sugiura, M. (1992). The chloroplast genome. *Plant Molecular Biology* **19**: 149-168.
- Sugiura, M. (1996). Structure and replication of chloroplast DNA. In: *Molecular genetics of photosynthesis*. Andersson, B., Salter, A. H. and Barber, J. (Eds.). Oxford University press. Oxford.
- Taberlet, P., Gielly, L., Pautou, G. and Bouvet, J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol. Biol.* **17**: 1105-1109.
- Tadesse Ebba (1969). T'ef (*Eragrostis tef*). The cultivation usage, and some of the known diseases and insect pests. Haile Selassie I University, College of Agriculture, Dire Dawa. *Exp. Sta. Bull.* 60.
- Tadesse Ebba (1975). T'ef (*Eragrostis tef*) cultivars: morphology and classification, part II. *Experiment Station Bulletin* 66. College of Agriculture, Addis Ababa University, Dire Dawa, Ethiopia.
- Tareke Berhe, Nelson, L. A., Morris, M. R. and Schmidt, J. W. (1989). Inheritance of phenotypic traits in tef: II seed colour. *J. Heredity* **80**: 65-67.
- Thompson, M. D., Zhang, L., Hong, L. and Hallik, R. B. (1997). Extensive structural conservation exists among several homologs of two *Euglena* chloroplast group II introns. *Mol. Gen. Genet.* **257**: 45-54.
- Thompson, G. (1977). The effect of a selected locus on linked neutral loci. *Genetics* **85**: 753-788.

- Vogel, J. C., Russell, S. J., Rumsey, F. J., Barrett, J. A. and Gibby, M. (1998). Evidence for maternal transmission of chloroplast DNA in the genus *Asplenium* (Aspleniaceae, Pteridophyta). *Bot. Acta* **111**: 247-249
- Vosman, B., Arens, P., Rus-KkorteKaas, W. and Smulders, M. J. M. (1992). Identification of highly polymorphic DNA regions in tomato. *Theor. Appl. Genet.* **85**: 239-244.
- Wang, Z. Y., Second, G. and Tanksley, S. D. (1992). Polymorphism and phylogenetic relationships among species in the genus *Oryza* as determined by analysis of nuclear RFLPs. *Theor. Appl. Genet.* **83**: 565-581.
- Wolfe, K. H., Li, W. and Sharp, P. M. (1987). Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad. Sci. USA* **84**: 9054-9058.
- Zurawski, G., Clegg, M. T. and Brown, A. H. D. (1984). The nature of nucleotide sequence divergence between barley and maize chloroplast DNA. *Genetics* **106**: 735-749.