

Complete chloroplast DNA sequences of Zanduri wheat (*Triticum* spp.)

M. Gogniashvili · P. Naskidashvili ·
D. Bedoshvili · A. Kotorashvili · N. Kotaria ·
T. Beridze

Received: 25 November 2014 / Accepted: 16 February 2015 / Published online: 11 March 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Georgia plays an important role in wheat formation. In the past, the Zanduri population of Georgia was a set of diploid—*Triticum monococcum* var. *hornemanii* ($2n = 14$) (Gvatsa Zanduri), tetraploid *Triticum timopheevii* ($2n = 28$) (Chelta Zanduri) and hexaploid *Triticum zhukovskyi* Men. et Er. ($2n = 42$). It is a Zanduri puzzle that wild *T. araraticum* was not found in Georgia, though cultivated *T. timopheevii* was only detected here. Next-generation sequencing technologies, which have been developed in recent years, enable the determination of complete nucleotide sequences of both chloroplast and mitochondrial DNA of many higher plants, including wheat. The genetic structure of Zanduri wheat is more accurately inferred by the complete sequences of chloroplast DNA. In the present investigation, the complete sequences of three Zanduri wheats (*T. timopheevii*, *T. zhukovskyi*, and *T. monococcum* var. *hornemanii*) and wild *T. araraticum*

are presented. Sequencing of chloroplast DNA was performed on an Illumina MiSeq platform. Chloroplast DNA molecules were assembled using the SOAPdenovo computer program. In comparison to *T. araraticum*, there are 12 SNPs, a 25 bp inversion in the *ccsA-ndhD* intergenic sequence, and a 38-bp inversion in the intergenic sequence *rbcL-rpl23* pseudogene identified in *T. timopheevii* and *T. zhukovskyi*. In addition, a 24 bp repeat of *trnG-trnI* intergenic sequence is present as a double copy in *T. araraticum*, whereas in *T. timopheevii* and *T. zhukovskyi*, it is present as a triple copy. Unlike *T. araraticum*, *T. timopheevii* and *T. zhukovskyi* have a 6 bp repeat in the gene *ndhH*, which results in a dipeptide duplication in the corresponding protein. Gvatsa Zanduri (*T. monococcum* var. *hornemanii*) chloroplast DNA slightly differs from other einkorn chloroplast DNA. In comparison to *T. monococcum*, four SNPs can be identified in *T. monococcum* (Gvatsa Zanduri), two in gene *matK* and one in gene *ndhD*. The sequenced chloroplast DNA molecules were compared to other *Triticum* and *Aegilops* species, and a phylogenetic tree was constructed. *T. araraticum*, *T. timopheevii* and *T. zhukovskyi* chloroplast DNA showed the closest phylogenetic relationship with the chloroplast DNA of *Ae. speltoides*. The most significant difference was in the 114-bp deletion within the gene *ndhH* in the Timopheevi species.

P. Naskidashvili: Deceased January 15, 2015.

M. Gogniashvili (✉) · T. Beridze
Institute of Molecular Genetics, Agricultural University of Georgia, #240 D. Agmashenebeli Alley, 0159 Tbilisi, Georgia
e-mail: m.gogniashvili@agruni.edu.ge

P. Naskidashvili · D. Bedoshvili
Agricultural University of Georgia, Tbilisi, Georgia

A. Kotorashvili · N. Kotaria
National Centre for Disease Control and Public Health, Tbilisi, Georgia

Keywords Chloroplast DNA · Illumina · Indels · Phylogeny · Sequencing · SNP · *Triticum* spp. · Wheat

Introduction

Wheat (*Triticum* L.) is the leading grain crop worldwide. It originated in the Fertile Crescent approximately 10,000 years ago and has since spread worldwide. There are four wild species, which grow in the Fertile Crescent of the Near East. The diploid wheat *Triticum monococcum* L. (einkorn) is among the first crops domesticated by humans in the Fertile Crescent. Using AFLPs, the site of domestication of a diploid wheat (einkorn) was identified (Heun et al. 1997). These results indicated that wild populations from the Karacadag Mountain of southeastern Turkey were more similar to domesticated einkorn than other wild populations. Wild einkorn underwent a process of natural genetic differentiation, most likely an incipient speciation, prior to domestication. Identified distinct wild einkorn races were designated as alpha, beta, and gamma. Only one of these natural races, beta, was exploited by humans for domestication (Kilian et al. 2007).

The first step in the evolution of cultivated wheat was the formation of *Triticum dicoccoides* (Körn.) Schweinf., a tetraploid species with an A^uA^uBB genome (Schneider et al. 2008). The A and B chromosomes of the tetraploid were derived from an earlier hybridization between the wild A^uA^u diploid *Triticum urartu* Tumanian ex Gandilyan and wild diploid B genome donor: the ultimate source of this B genome is still discussed. To resolve the relationship of domesticated tetraploid wheat with its wild relatives, molecular fingerprinting of wild and domesticated tetraploid lines, based on AFLP marker data, was performed. These results clearly indicated that the domesticated tetraploid wheat is very closely related to wild populations sampled in southeastern Turkey. Northeast expansion of domesticated emmer cultivation resulted in sympatry with *Aegilops tauschii* Coss. (genome DD). Approximately 7000 years ago, the hexaploid bread wheat *T. aestivum* L. (BBA^uA^uDD) arose in the South Caucasus region by allopolyploidization of the cultivated Emmer wheat *Triticum dicoccum* Schrank with the Caucasian *Aegilops tauschii* subsp. *strangulata* (Eig) Tzvelev (Dvorak et al. 1998; Dubcovsky and Dvorak 2007).

Another line of *Triticum* species is the Timopheevi group, which contains the G genome. Hexaploid wheat *Triticum zhukovskyi* Menabde et Ericzyan (GGA^uA^u-A^mA^m) is the result of hybridization of *Triticum*

timopheevii (Zhuk.) Zhuk. (GGA^uA^u) with *T. monococcum* L. var. *hornemanii* (Clemente) Körn. (A^mA^m) (Menabde and Eritsian, 1960). *T. timopheevii* wheat was discovered in Western Georgia, where it was termed Zanduri wheat. In the past, the Zanduri population was a set of diploid—*T. monococcum* var. *hornemanii* (2n = 14) (Gvatsa Zanduri), tetraploid *T. timopheevii* (2n = 28) (Chelta Zanduri) and hexaploid *T. zhukovskyi* Men. et Er. (2n = 42). *T. timopheevii* was discovered and described in 1922 by P. Zhukovski, who named this species in honor of Prof. Stepan N. Timofeev, a major figure in the development of agronomic scientific thought in Georgia (Menabde 1948). Subsequently, Menabde and Eritsian found new species in the same fields where Zanduri wheat grew, and “in honor of Academician P.M. Zhukovski, a known connoisseur of cultivated flora who has performed extensive work in the field of cereals of Georgia,” they gave these species a new name called *Triticum zhukovskyi* (Menabde and Eritsian 1960). They described *T. zhukovskyi* as an allopolyploid of *T. timopheevii* and *T. monococcum*.

Georgia plays an important role in wheat formation. Georgian endemics include *Triticum karamyshevii* Nevski (*Triticum turgidum* L. subsp. *palaeocolchicum* Menabde Á. and D. Löve), *T. timopheevii*, *T. zhukovskyi*, *Triticum macha* Dekapr. et Menabde (*Triticum aestivum* L. subsp. *macha* (Dekapr. et Menabde) MacKey) and *Triticum carthlicum* Nevski (*Triticum turgidum* L. subsp. *carthlicum* (Nevski) Á. et D. Löve) (Menabde 1961; Hammer et al. 2011). Using phylogenetic research, a basic variety of wheat species has been revealed in the agriculture of Georgia. Some of these species bear an evolutionarily close affinity to wild wheat species or have retained some of their features. N. Vavilov focused great attention on Georgia in his studies. He has visited Georgia 16 times. On the basis of expeditions, which were held by his followers (Zhukovski, Menabde, Dekaprevich and others), the world became acquainted with several wheat species.

It is a Zanduri puzzle that the wild *T. araraticum* Jakubz. (*Triticum timopheevii* subsp. *armeniicum* (Jakubz.) Slageren) was not found in Georgia, though cultivated *T. timopheevii* is only detected here. Mori et al. (2009) analyzed the molecular variation at 23 microsatellite loci in the chloroplast genome of

Timopheevi wheats. Allelic diversity was evaluated using 94 accessions, which represented domesticated Timopheevi wheat (*T. timopheevii*), wild Timopheevi wheat (*T. araraticum*), and wild emmer wheat (*T. dicoccoides*). None of the *T. araraticum* plastotypes collected in South Caucasus were closely related to the *T. timopheevii* plastotype. However, the plastotypes found in northern Syria and southern Turkey showed closer relationships with *T. timopheevii*. These results suggested that the domestication of *T. timopheevii* wheat might have occurred in regions including southern Turkey and northern Syria (Mori et al. 2009). According to Mori et al. domesticated Timopheevi wheat might have been introduced to Georgia after domestication (Mori et al. 2009). There are also questions on how and why the domesticated Timopheevi wheat was transferred from Turkey or Syria to Georgia.

Although Timopheevi wheat is a minor crop worldwide, it is an important species for studying the evolution of polyploidy in tetraploid wheat, as well as for understanding the early history and spread of agriculture in the Near East (Mori et al. 2009). Genetic studies of plasmons in *Triticum* have revealed that the plasmon of Timopheevi wheat is distinct from that of emmer wheat (Wang et al. 1997). According to Wang et al. the ancestral female parent of both Emmer and Timopheevi wheat was *Aegilops speltoides* Tausch (Wang et al. 1997). Plasmons of *T. araraticum*, *T. timopheevii* and *T. zhukovskyi* belong to the G type; that of *T. aestivum* and *T. dicoccoides* belong to the B type; and *T. monococcum* belongs to the A² type.

Next-generation sequencing technologies, which have been developed in recent years, enable the determination of the complete nucleotide sequence of both chloroplast and mitochondrial DNAs of many higher plants, including wheat. The full sequence of Zanduri wheat is not yet known. The genetic structure of Zanduri wheat is more accurately inferred by the complete sequences of chloroplast DNA.

In the present investigation, the complete sequences of four Zanduri wheats are presented, including *T. araraticum*, *T. timopheevii*, *T. zhukovskyi*, and *T. monococcum* var. *hornemanii*, and these sequences were compared to other *Triticum* and *Aegilops* species, after which, we constructed phylogenetic tree. *T. araraticum*, *T. timopheevii* and *T. zhukovskyi* chloroplast DNA showed the closest phylogenetic relationship with the chloroplast DNA of *Ae. speltoides*.

Materials and methods

Plant material and DNA isolation

The seeds of wheat species were received from the seed bank of the Agricultural University of Georgia: *T. monococcum* var. *hornemanii* (Gvatsa Zanduri); *T. timopheevii*—corresponds to the domesticated Timopheevi subgroup G-2 (Mori et al. 2009) and *T. zhukovskyi*; *T. araraticum* corresponds to the wild Timopheevi subgroup G-3 (Mori et al. 2009).

The seeds were grown in water at room temperature. Total genomic DNA was extracted from young wheat leaves. The leaves were ground in liquid nitrogen, and DNA was isolated according to the modified Marmur's procedure (Beridze et al. 1967).

The PCR conditions were as follows: 1 min denaturing at 94 °C; 30 cycles of 94 °C denaturing (1 min), 55 °C annealing (1 min) and 72 °C extension (2 min); followed by a final extension step at 72 °C (5 min). Sigma-Aldrich chemicals were used for PCR reactions. PCR products were purified using GenElute PCR Clean-Up Kits (Sigma-Aldrich), dye-labeled using a Big Dye Terminator Kit (applied biosystems) and analyzed on either the Applied Biosystems 3100 or 3700 genetic analyzers (Laboratory Services Division of the University Guelph, ON, Canada). Sequences were manually aligned in Se-AL (Rambaut 2002).

Genomic DNA library preparation

Genomic DNA libraries were constructed using the TruSeq DNA Sample prep kit (Illumina, San Diego, CA, USA). Genomic DNAs were quantified using the Qbit BR reagents (Qbit 2.0 Fluorometer—Life Technologies). Briefly, 1 mcg of DNA was sheared into 300-bp fragments on a Covaris M220 focused ultrasonicator (Covaris Inc) using SonoLab™ 7.1 software for 250 cycles/burst, 20.0 Duty factor, 50.0 Peak power, in screw-cap microtubes. After shearing, the DNA was blunt-ended, 3'-end A-tailed and ligated to indexed adaptors. The adaptor-ligated genomic DNA was size selected with AMPure-beads using the gel-free protocol described in the TruSeq DNA Sample Prep manual. Size-selected DNA was amplified by PCR to selectively enrich for fragments that have adapters on both ends. Final amplified libraries were

run on an Agilent bioanalyzer DNA 2100 (Agilent, Santa Clara, CA, USA) to determine the average fragment size and to confirm the presence of DNA of the expected size range.

Sequencing on an illumina MiSeq platform

The libraries were pooled in equimolar concentration and loaded onto a flow cell for cluster formation and sequenced on an Illumina MiSeq platform. The libraries were sequenced from both ends of the molecules to a total read length of 150 nt from each end. The raw.bcl files were converted into demultiplexed compressed fastq files using Casava 1.8.2 (Illumina). Sequencing of wheat DNA samples was performed at the facilities of the National Centre for Disease Control and Public Health, Tbilisi, Georgia.

Some regions of wheat chloroplast DNA were filled using Sanger sequencing.

Assembly of chloroplast DNA

FASTAQ files were trimmed using the computer program Sickle, a windowed adaptive trimming tool for FASTQ files using quality (<https://github.com/najoshi/sickle>). The reads were filtered by standard parameters (quality reads—20, cutoff length—20). The reads containing “N” were discarded. The filtered chloroplast reads were assembled using the SOAPdenovo2 software program (version 127mer) (Li et al. 2009). The reads were first *de novo* assembled into contigs with *k*-mers 83–93. All contigs were aligned to the reference chloroplast genome sequence using BLASTN (<http://www.ncbi.nlm.nih.gov>). If there was no overlap between two adjoining fragments, then the interval was filled using Sanger sequencing. For the assembly of *T. monococcum* var. *hornemanii*, einkorn chloroplast DNA was used (Middleton et al. 2014). For the three other wheat species, the chloroplast DNA of *T. aestivum* was used (Matsuoka et al. 2002). Merging of large overlapping contigs was performed according to EMBOSS 6.3.1: merger (Rice et al. 2000). Automatic annotation of chloroplast DNA was performed by Dogma (Wyman et al. 2004).

Complete chloroplast DNA sequences have been deposited into the DNA Data Bank of Japan (Table 2).

Results

Due to the high number of chloroplast DNA copies in plant genomic DNA sequencing, the number of reads for chloroplast DNA reaches 20,000 per base in some cases. However, for DNA resequencing, 40–50 reads are sufficient. Thus, the new methodology of chloroplast DNA resequencing was developed (Hancock-Hanser et al. 2013; Middleton et al. 2014; Tabidze et al. 2014). Briefly, this technique mixes and sequences the genomic DNA from many cultivars in one Illumina lane and barcodes each library to identify which reads are coming from which cultivar. Unlike the number of reads from chromosomal DNA, which are low and are thus ignored, the number of chloroplast and mitochondrial DNA reads are relatively high and sufficient to compute the chloroplast DNA sequence.

The sequence of Zanduri wheat DNA was performed on the Illumina MiSeq platform. The coverage of chloroplast DNA was app. 1000. Chloroplast DNA molecules were assembled using the SOAPdenovo2 software program (Li et al. 2009). Consequently, the chloroplast DNA sequence of four *Triticum* species was determined. The length of the chloroplast DNA is given in Table 1.

Using *T. araraticum* as an example, 9 SNPs can be identified in *T. timopheevii* and *T. zhukovskyi* (Table 3). The number of noncoding substitutions is 8, and the number of coding substitutions is 1. The coding substitution is synonymous, which does not alter the amino acid. One SNP was observed in 23 *rRNA* genes. In comparison with *T. araraticum*, a 38 bp inversion was observed in *T. timopheevii* and *T. zhukovskyi* at position 56312–56349 (intergenic sequence of *rbcL*—pseudogene *rpl23*). In addition, the 38-bp sequence is a palindrome with a 4 bp loop and 17 bp stem. Another inversion (25 bp long, also a palindrome) was observed at position 107013–107037 (intergenic *ccsA*—*ndhD*) with a 3 bp loop and 11 bp stem.

There were seven 1 bp indels detected in *T. timopheevii* and *T. zhukovskyi*. Six indels, which were longer than one bp, were observed (Table 4). Four of these indels were in intergenic spacers, and one indel was within an intron. The sixth is a 6 bp duplication in the *ndhH* gene (position 101,778–101,783). Consequently, dipeptide duplication occurs in the

Table 1 Regions sequenced using capillary electrophoresis

	Forward	Reverse
Primer 1 for filling gaps		
<i>T. timopheevii</i> (56641–56643)	AGTCTTCGTTTATTGGGAAAGAATC	TGACTCACGATTACGCCTGT
<i>T. zhukovskiyi</i> (56637–56647)		
<i>T. araraticum</i> (56491–56631)		
<i>T. aestivum</i> (56391–56494)		
<i>T. monococcum</i> (56575–56690)		
Primer 2 for filling gaps		
<i>T. timopheevii</i> (82515–82517, 133897–133899)	GTGTGCTCGGGATAGGTGTT	GGAACTCAGAGCATAGGAGGG
<i>T. zhukovskiyi</i> (82511–82519, 133895–133903)		
<i>T. araraticum</i> (82455–82509, 133832–133886)		
<i>T. aestivum</i> (82183–82288, 133616–133721)		
<i>T. monococcum</i> (83156–83271, 134602–134717)		

Table 2 Accession numbers and the length of Zanduri chloroplast DNA in the Japan Data Bank

	Accession number	Length in bp
<i>T. monococcum</i> var. <i>hornemaniai</i>	LC005977	136,861
<i>T. araraticum</i>		136,118
<i>T. timopheevii</i>	AB976560	136,157
<i>T. zhukovskiyi</i>		136,157
<i>T. aestivum</i> subsp. <i>macha</i>	NC_025955.1	135,899

Table 3 SNPs in Zanduri chloroplast DNA

Nucleotide position (according to <i>T. araraticum</i>)	Locus	<i>T. araraticum</i>	<i>T. timopheevii</i>	<i>T. zhukovskiyi</i>
2607	Gene <i>matK</i>	A	T	A
3315	Intron <i>trnK-UUU</i>	T	T	C
11862	Intergenic <i>psbZ-trnG</i>	T	G	G
46235	Intergenic <i>trnT-trnL</i>	G	T	T
48971	Intergenic <i>ndhJ-ndhK</i>	G	C	C
60624	Intergenic <i>petA-psbJ</i>	A	C	C
70091	Intergenic <i>psbB-psbT</i>	G	T	T
76810	Intergenic <i>rps8-rpl14</i>	C	G	G
112661	23S <i>rRNA RGC160</i>	C	T	T

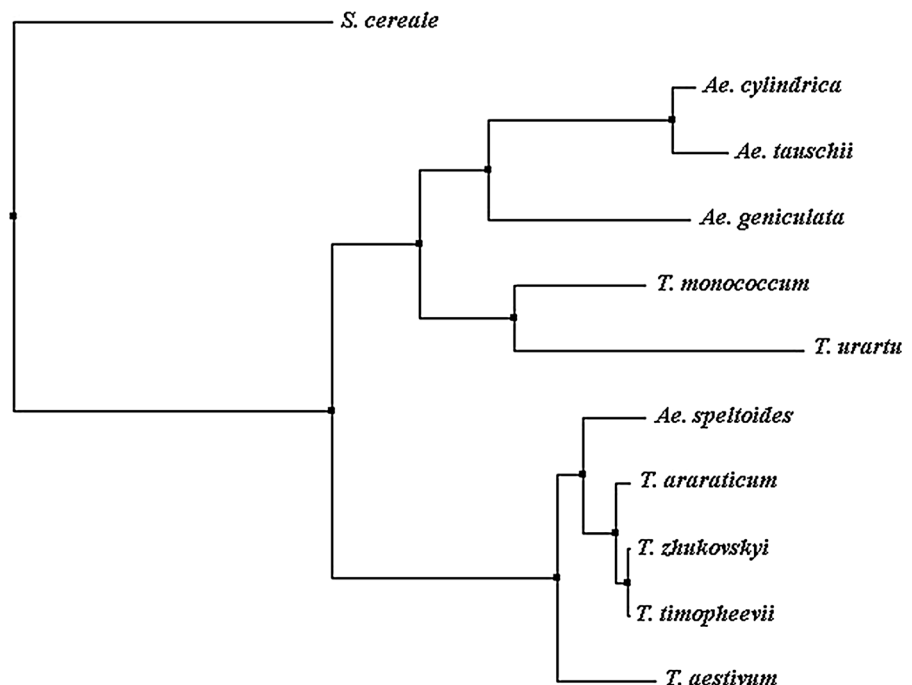
corresponding protein (Ile-Arg; Position 69–70 in the *ndhH* protein). A 24 bp sequence of the *trnG-trnI* intergenic region, which is observed as a single sequence in *T. aestivum*, was present as a double copy in *T. araraticum* and as a triple copy in *T. timopheevii* and *T. zhukovskiyi* (Table 4).

To illustrate the evolutionary relationship among the studied cultivars, a phylogenetic tree was constructed based on multiple alignments using Jalview version 2 (Waterhouse et al. 2009). Figure 1 depicts the resulting phylogenetic tree. The tree was drawn using some previously published results (Middleton

Table 4 Long indels in Zanduri chloroplast DNA

Nucleotide position (according to <i>T. araraticum</i>)	Locus	<i>T. araraticum</i>	<i>T. timopheevii</i>	<i>T. zhukovskyi</i>	Type of indel	Amino acid duplication
4944	Intron <i>rps16</i>	–	+AAAA	+AAAA	Duplication	
14242–14290	Intergenic <i>trnG-trnI</i>	24 bp, 2 copies	24 bp, 3 copies	24 bp, 3 copies	Duplication	
14547	Intergenic <i>trnG-trnI</i>	–	TATTATTT	TATTATTT	Insertion	
18900	Intergenic <i>trnC-rpoB</i>	–	AGCAT	AGCAT	Duplication	
60981	Intergenic <i>petA-psbJ</i>	16 T	-TTTT	-TTTT	Deletion	
101778	Gene <i>ndhH</i>	–	AAGGAT	AAGGAT	Duplication	Arg-ileu

Fig. 1 Complete chloroplast genome phylogeny of *Triticum* and *Aegilops* species, including Zanduri wheat and *Triticum macha*. Neighbour joining tree using PID (Waterhouse et al. 2009). The GenBank accessions used for the analyses are *Aegilops cylindrica* Host NC_023096.1; *Aegilops geniculata* Roth NC_023097; *Aegilops speltoides* NC_022135; *Aegilops tauschii* NC_022133; *Triticum urartu* NC_021762; *Secale cereale* L. NC_021761



et al. 2014). *T. timopheevii* and *T. zhukovskyi* chloroplast DNA showed the closest phylogenetic relationship with the chloroplast DNA of *Ae. speltoides* (Fig. 1). Their sequences were more than 99 % identical.

If we take the chloroplast DNA of *T. monococcum* (Middleton et al. 2014) as a reference, four SNPs can be identified in *T. monococcum* var. *hornemaniai* (Gvatsa Zanduri), two in the gene *matK*, one in the gene *ndhD*, and one in the intron of *atpF* (Table 5). Middleton et al. observed a 1076 bp-long insertion at position 86250–87326 of *T. monococcum* (2014). This region is absent in all chloroplast DNA of higher plants, but according to GenBank, it is a component of

mitochondrial DNA. In our opinion, it is a mitochondrial DNA sequence that, during assembly, was erroneously included into chloroplast DNA by the authors.

Discussion

The term Zanduri is a Georgian word and unites three species of the genus *Triticum*: *T. monococcum* var. *hornemaniai* (Gvatsa Zanduri), *T. timopheevii* (Chelta Zanduri), and *T. zhukovskyi*. Apparently, these three species were grown together for centuries. Co-cultivation of *T. monococcum* (Gvatsa Zanduri) and *T.*

Table 5 SNPs in *T. monococcum* var. *hornemanii* (Gvatsa Zanduri)

Nucleotide position, according to (<i>T. monococcum</i>) (Middleton et al. 2014)	Locus	<i>T. monococcum</i> subsp. <i>monococcum</i> as a reference	Gvatsa Zanduri
2592	Gene <i>matK</i>	T	G
2593	Gene <i>matK</i>	G	A
33649	Intron <i>atpF</i>	G	A
109159	Gene <i>ndhD</i>	G	T

Table 6 *rpl23* Gene and *rpl23* pseudogene in *Aegilops* and *Triticum* species

Species	Plasmon	Gene <i>rpl23</i>	Pseudogene <i>rpl23</i>
<i>Ae. speltoides</i> NC_022135	G	–	–
<i>T. araraticum</i>	G	+	+
<i>T. timopheevii</i> AB976560	G	+	+
<i>T. zhukovskyi</i>	G	+	+
<i>T. monococcum</i> NC_021760	A ²	+	+
<i>T. monococcum</i> var. <i>hornemanii</i> , Gvatsa Zanduri LC005977	A ²	+	+
<i>T. aestivum</i> NC_025955.1	B	+	+

timopheevii (Chelta Zanduri) contributed to the formation of a new species, called *T. zhukovskyi*, on the basis of allopolyploidy. Wild *T. araraticum* can also be included in this group. According to Wang et al., the ancestral, female parent of Timopheevi wheat is *Aegilops speltoides* (Wang et al. 1997). Plasmons of *T. araraticum*, *T. timopheevii*, and *T. zhukovskyi* belong to the G type; those of *T. aestivum* and *T. dicoccoides* belong to the B type; and *T. monococcum* belongs to the A² type (Table 6).

According to Gill and Friebe (2002), two *Ae. speltoides* (nuclear genome SS) with different plasmons (I and II) contributed to the formation of *Triticum* lines. *Ae. speltoides* (plasmon I) participated in the formation of the *T. turgidum*-*T. aestivum* line and *Ae. speltoides* (plasmon II) in the *T. timopheevii*–*T. zhukovskyi* line. Middleton et al. (2014) sequenced the chloroplast DNA of *Ae. speltoides* var. *ligustica* (Savign.) from Turkey. The phylogenetic tree (Fig. 1) revealed that *Ae. speltoides* branches from the Timopheevi group. Thus, it can be concluded that this species belongs to *Ae. speltoides* with plasmon II.

Structural alterations of the chloroplast genome in the genera *Triticum* and *Aegilops* tend to occur at “hot spots” on the physical map. Deletions/insertions were found in a region between the *rbcl* (coding for the

large subunit of ribulose-1,5-bisphosphate carboxylase) and *petA* (cytochrome f) genes, indicating that it is a hot spot for variation (Ogihara et al. 1988). Guo and Terachi (2005) determined the nucleotide sequence in the hotspot, between a stop codon of *rbcl* and a *HindIII* site in *cemA*. These results revealed that a considerable number of polymorphisms were present in the loop, with a stem-loop structure at the 3' end of the gene *rbcl*. In contrast to the loop, a pair of 17 bp inverted repeats (i.e., a stem) was monomorphic and highly conserved in all of the accessions.

Differences between the chloroplast DNA of *Ae. speltoides* and Timopheevi line species are shown in Table 6. Unlike the Timopheevi line, *rpl23* pseudogene (300 bp), *rpl23* gene (286 bp), and *psbE-petL* intergenic (452 bp) sequences were missing in *Ae. speltoides*. However, in the Timopheevi line, compared to *Ae. speltoides*, the *ndhH* gene lacks 114 bp and the intergenic sequence of *ndhH-ndhF* is missing 27 bp.

The Zanduri puzzle is not yet solved. Menabde distinguishes between primary and secondary centers of wheat formation (Menabde 1948). According to Menabde main types of wheat variety are concentrated in the primary center of speciation. The presence of wild wheats and primary cultivated species

convincingly suggest the primary center of speciation and their absence—a secondary center. The primary species genetically exhibit the greatest affinity to wild relatives, and the secondary greatest distance. This criterion allows asserting the primacy of the Georgian hearth in the process of wheat speciation. In the category of primary cultured species Menabde include: *T. macha*, *T. timopheevii*, *T. monococcum* (Gvatsa Zanduri), *T. karamyshevii*. The whole range of diverse species of cultivated wheat is present only in the Georgian hearth, and a significant portion of the primary species is not moved beyond this region.

It is not known how the Georgian tribes came to the South Caucasus and where they lived at the period of wheat domestication. Proto-Georgian (Proto-Kartvelian) language is included among seven language families of Eurasia postulated to form a linguistic superfamily that evolved from a common ancestor around 15,000 years ago (Pagel et al. 2013).

Acknowledgments The authors wish to acknowledge the constant interest and support of Mr. K. Bendukidze who untimely passed away on 13th November, 2014. This research was funded by the Knowledge Fund. The Knowledge Fund is a funding organization of the Free University of Tbilisi and Agricultural University of Georgia. Correction of the manuscript in terms of English was funded through the University Research Program by the U.S. Embassy in Georgia, grant No S-GE800-13-GR-122.

References

- Beridze TG, Odintsova MS, Sissakian NM (1967) Distribution of bean leaf DNA components in the cell organell fractions. *Molek Biol USSR* 1:142–153
- Dubcovsky J, Dvorak J (2007) Genome plasticity a key factor in the success of polyploidy wheat under domestication. *Science* 316:1862–1866
- Dvorak J, Luo MC, Yang ZL, Zhang HB (1998) The structure of the *Aegilops tauschii* genepool and the evolution of hexaploid wheat. *Theor Appl Genet* 97:657–670
- Gill BS, Friebe B (2002) Cytogenetics, phylogeny and evolution of cultivated wheats (2002) In: Bread wheat; FAO Plant Production and Protection Series (FAO), no. 30 Curtis, B.C., Rajaram, S., Gomez Macpherson, H. (eds.)/FAO, Rome (Italy). Plant Prod Protect Div
- Guo CH, Terachi T (2005) Variations in a hotspot region of chloroplast DNAs among common wheat and *Aegilops*. *Genes Genet Syst* 80(4):277–285
- Hammer K, Filatenko AA, Pistrick K (2011) Taxonomic remarks on *Triticum* L. and *×Triticosecale* Wittm. *Genet Resour Crop Evol* 58:3–10
- Hancock-Hanser BL, Frey A, Leslie MS, Dutton H, Archer FI, Morin PA (2013) Targeted multiplex next-generation sequencing: advances in techniques of mitochondrial and nuclear DNA sequencing for population genomics. *Mol Ecol Res* 13(2):254–268
- Heun M, Schaefer-Pregl R, Klawan D, Castagna R, Accerbi M, Borghi B, Salamini F (1997) Site of Einkorn wheat domestication identified by DNA fingerprinting. *Science* 278:1312–1314
- Kilian B, Ozkan H, Walther A, Kohl J, Dagan T, Salamini F, Martin W (2007) Molecular diversity at 18 loci in 321 wild and 92 domesticate lines reveal no reduction of nucleotide diversity during *Triticum monococcum* (Einkorn) domestication: implications for the origin of agriculture. *Mol Biol Evol* 24:2657–2668
- Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, Wang J (2009) SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25:1966–1967
- Matsuoka Y, Yamazaki Y, Ogihara Y, Tsunewaki K (2002) Whole chloroplast genome comparison of rice, maize, and wheat: implications for chloroplast gene diversification and phylogeny of cereals. *Mol Biol Evol* 19(12):2084–2091
- Menabde VL (1948) Wheats of Georgia. Edition of Academy of Science of Georgian SSR, Tbilisi, 272 pp. (in Russian)
- Menabde VL (1961) Cultivated flora of Georgia. In: Sakhokia MF (ed) Botanical excursions over Georgia. Publishing House of the Academy of Sciences of Georgian SSR, Tbilisi, pp 69–76 (in Russian)
- Menabde VL, Eritsian AA (1960) Investigation of Georgian wheat Zanduri. *Soobsch Acad Sci GSSR* 25:731–736
- Middleton CP, Senerchia N, Stein N, Akhunov ED, Keller B, Wicker T, Kilian B (2014) Sequencing of chloroplast genomes from wheat, barley, rye and their relatives provides a detailed insight into the evolution of the Triticeae tribe. *PLoS One* 9:e85761
- Mori N, Kondo Y, Ishii T, Kawahara T, Valkoun J, Nakamura C (2009) Genetic diversity and origin of timopheevi wheat inferred by chloroplast DNA fingerprinting. *Breeding Sci* 59:571–578
- Ogihara Y, Terachi T, 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
- Pagel M, Atkinson QD, Calude AS, Meade A (2013) Ultra-conserved words point to deep language ancestry across Eurasia. *Proc Natl Acad Sci USA* 110:8471–8476
- Rambaut A (2002) SE-AL Sequence alignment program. Department of Zoology, University of Oxford, UK
- Rice P, Longden I, Bleasby A (2000) EMBOSS: the European molecular biology open software suite. *Trends Genet* 16(6):276–277
- Schneider A, Molnar I, Molnar-Lang M (2008) Utilisation of *Aegilops* (goatgrass) species to widen the genetic diversity of cultivated wheat. *Euphytica* 163:1–19
- Tabidze V, Baramidze G, Pipia I, Gogniashvili M, Ujmajuridze L, Beridze T, Hernandez AG, Schaal B (2014) The complete chloroplast DNA sequence of eleven grape cultivars. Simultaneous resequencing methodology. *Journal International des Sciences de la Vigne et du Vin J Int Sci Vigne Vin* 48(2):99–109

- Wang G-Z, Miyashita NT, Tsunewaki K (1997) Plasmon analysis of *Triticum* (wheat) and *Aegilops*: PCR-single-strand conformational polymorphism (PCR-SSCP) analyses of organellar DNAs. *Proc Natl Acad Sci USA* 94:14570–14577
- Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ (2009) Jalview version 2: a multiple sequence alignment and analysis workbench. *Bioinformatics* 25(9):1189–1191
- Wyman SK, Jansen RK, Boore JL (2004) Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* 20(17):3252–3255