Two major lineages of *Aegilops tauschii* Coss. revealed by nuclear DNA variation analysis

Takumi S[1], Mizuno N[1], Okumura Y[1], Kawahara T[2], Matsuoka Y[3]

[1] Graduate School of Agricultural Science, Kobe University, Nada-ku, Kobe 657-8501, Japan. [2] Graduate School of Agriculture, Kyoto University, Muko, Kyoto 617-0001, Japan. [3] Department of Bioscience, Fukui Prefectural University, Matsuoka, Eiheiji, Yoshida, Fukui 910-1195, Japan

INTRODUCTION

Aegilops tauschii Coss. (syn. *Ae. squarrossa* L.), a wild diploid self-pollinating goatgrass, is the D genome donor of common wheat[1,2]. The genome of *Aegilops tauschii* was brought into common wheat though a natural cross with tetraploid emmer wheat about 8,000 years ago. Habitats of *Ae. tauschii* are widely distributed from Syria and Turkey to China in Eurasia. The *Ae. tauschii* population carries large diversity at the molecular levels[3], considered to be a useful source for common wheat breeding. In fact, some studies have reported higher levels of genetic variability of glutenin subunits and gliadin in *Ae. tauschii* than in the D genome of common wheat[4].

The subspecific constitution in *Ae. tauschii* is not clear. Eig[5] categorized two subspecies, *tauschii* (syn. *eusquarrosa* Eig) and *strangulata* based on the spikelet morphology. It has been suggested that ssp. *strangulata* could have been the D genome donor[6,7] and that the birthplace of common wheat was most likely to lie within the region comprising Transcaucasia and the south coastal region of Caspian Sea that was the known distribution zone for ssp. *strangulata*[8]. Thus, the subspecific constitution is an important question to understand the intraspecific differentiation of *Ae. tauschii*, whereas the two typical forms of ssp. *tauschii* and *strangulata* are connected by a continuous range of intermediate forms[9]. Some recent reports also showed difficulty to distinguish the two subspecies based on molecular markers and suggest high gene flow between the subspecies[10,11].

In this study, we used 62 *Ae. tauschii* accessions covering the entire species range to study the population structure of *Ae. tauschii* based on polymorphisms of simple sequence repeat (SSR) and nucleotide variation of several loci. The implications of those findings for the intraspecific differentiation are discussed.

MATERIALS AND METHODS

A total of 62 *Ae. tauschii* accessions representing the entire natural habitat range were used in this study (Table 1). Their passport data including geographical coordinates were cited in the previous reports[12,13]. For each accession, seeds were propagated from a single plant by selfing and used for the DNA extraction.

The extracted total DNA was used for SSR analysis and PCR amplification of coding genes. A total of 17 microsatellite loci previously reported[14], Xdgm3, Xdgm14, Xdgm19, Xdgm33, Xdgm46, Xdgm61, Xdgm68, Xdgm88, Xdgm111, Xdgm125, Xdgm126, Xdgm127, Xdgm128, Xdgm129, XBARC110, Xcfd92 and XBARC126, were used for evaluation of the nuclear genome diversity. Thirty-five cycles of PCR were performed for amplification; 1 min at 94˚C, 1 min at 55˚C and 1 min at 72˚C. Amplified products were fractionated by electrophoresis through 8.0% denaturing polyacrylamide gels and visualized by silver-staining.

SSR bands were scored as present (1) or absent (0) for measuring the Nei's genetic distances. Based on all the pair-wise genetic distances, a dendrogram was constructed by the unweighted pair group method with arithmetic mean.

To study nucleotide diversities in protein-encoding nuclear loci, intragenic regions were amplified in 20 to 30 *Ae. tauschii* accessions by PCR, and the PCR products were directly sequenced by the automated fluorescent Dye Deoxy terminator cycle sequencing. Multiple sequence alignments were carried out using the ClustalW computer program, and phylogenetic trees were constructed by the Neighbor-Joining method. Moreover, Tajima's D and Fu and Li's D were estimated using DnaSP 4.0.

RESULTS AND DISCUSSION

To study the population structure of *Ae. tauschii*, SSR analysis with 17 primer sets was conducted using total DNAs from 62 *Ae. tauschii* accessions representing the entire species range. An SSR phylogenetic tree was constructed based on the Nei's genetic distances, showing that there were two major lineages in *Ae. tauschii* (Fig. 1). Lineage 1 consisted of accessions from the eastern habitats mainly in Afghanistan and Pakistan. Lineage 2 included ssp. *tauschii* accessions of the western habitats and all accessions of ssp. *strangulata* accessions. So, both lineages included ssp. *tauschii*, whereas ssp. *strangulata* belonged only to the lineage 2. The phylogenetic tree was consistent with a previous report of the similar SSR analysis[15]. Two varieties, *meyeri* and *anathera*, were classified into the lineage 2 and 1, respectively.

In our previous study, the chloroplast DNA haplogroup network was constructed based on the biallelic base change and minisatellite sites in the flanking regions of four chloroplast microsatellite loci[12], and 18 haplogroups were identified[16]. Compared with the chloroplast DNA haplogroup network, accessions belonging to the largest chloroplast DNA haplogroup (HG7) were distributed in

the both two lineages. Other chloroplast DNA haplogroups including two large haplogroups HG9 and HG16 were classified into either the lineage 1 or the lineage 2. This fact supported a hypothesis that HG7 was ancestral to HG9, HG16 and many of the other minor haplogroups[16]. HG9 and HG16 were respectively classified into the lineage 2 and 1. HG16 was unique in that its distribution center is in the eastern habitats[16].

Table 1 The strain numbers and sources of the *Ae. tauschii* accessions used in this study

Afghanistan (AFG, n=10): KU-2012, KU-2028, KU-2032, KU-2042, KU-2058, KU-2627, KU-2633, KU-2636, KU-2639, PI47687

Armenia (ARM, n=3): IG126280, IG126293, IG48747

Azerbaijan (AZE, n=3): IG47182, IG47196, KU-2801

China (CHI, n=4): AT60, AT80, PI508262, PI508264

Dagestan (DAG, n=2): IG120866, KU-20-1

Georgia (GEO, n=3): AE454, KU-2826, KU-2834

India (IND, n=1): IG48042

Iran (IRN, n=18): KU-2068, KU-2069, KU-2078*, KU-20-8, KU-20-9*, KU-20-10, KU-2087, KU-2088*, KU-2090*, KU2096, KU-2100, KU-2110, KU-2111, KU-2118, KU-2122, KU-2126, KU-2155, KU-2160

Kazakhstan (KAZ, n=1): AE1090

Kyrgyzstan (KYR, n=1): IG131606

Pakistan (PAK, n=5): IG46682, KU-2001, KU-2003, KU-20-6, CGN10769

Syria (SYR, n=2): IG47259, IG46623

Tajikistan (TAJ, n=2): AE1038, IG48559

Turkey (TUR, n=3): KU-2132, PI486267, PI486277

Turkmenistan (TKM, n=2): IG126489, IG48508

Uzbekistan (UZE, n=2): IG48539, IG48567

KU, Plant Germ-Plasm Institute, Faculty of Agriculture, Kyoto University; PI, USDA-ARS; IG, ICARDA; CGN, Centre for Genetic Resources, The Netherlands; AE, IPK; AT, Kenji Kato, Okayama University; *, subspecies *strangulata*

Furthermore, we analyzed nucleotide sequences of 10 nuclear genes, namely, *Vrn-1*, *Wcor615*, *Ppd-1*, a *CONSTANS* homolog *TaHd1*, 2 MADS-box genes, *WPI1* and *WPI2*, and 4 abiotic stress-responsive genes encoding CBF/DREB-type transcription factors such as WCBF2, TaCBF2, WDREB2 and WDBF1 using 30 accessions of *Aegilops tauschii*. *Vrn-1*, *Ppd-1* and *TaHd1* are associated with determination of flowering time. *Wcor615* and CBF/DREB genes are cold and/or drought-responsive and related with development of abiotic stress tolerance. *WPI1* and *WPI2* function in flower development.

Based on the nucleotide sequence variations, a phylogenetic tree was constructed for each gene and compared with the SSR phylogenetic tree. The *Vrn-1*,

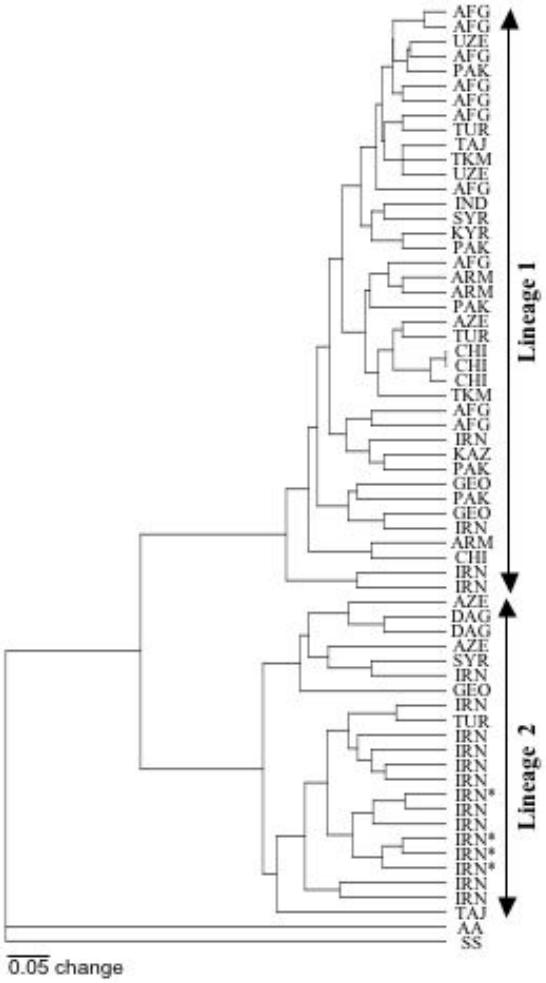

Fig. 1. A phylogenetic tree based on data from the SSR analysis. *Triticum urartu* (AA) and *Ae. speltoides* accessions were used as outgroups. *, subspecies *strangulata*

TaHd1, *WDREB2* and *WDBF1* sequences were clearly divided into two lineages, consistent with the pattern seen in the SSR-based phylogenetic tree (Fig. 2A). The *WPI1* sequences were also classified into two lineages, but the grouping pattern was completely different from that of the SSR-based tree (Fig. 2C). The two-lineage structure was not observed for the rest five genes (Fig. 2B). Tajima and Fu & Li tests on the 10 loci were conducted to examine the neutral mutation theory. Tajima's D and Fu and Li's D were significantly negative ($P<0.05$) only in *TaCBF2*. Tajima's D in *WPI2* showed significantly a positive result ($P<0.05$). No significant results were observed in other examined loci.

These results indicated that *Ae. tauschii* population was generally diverged into two major lineages, and that spp. *strangulata* was derived only from one of the two lineages. Overall, the two-lineage structure seemed to reflect the west-to-east dispersal and diversification history of *Ae. tauschii*.

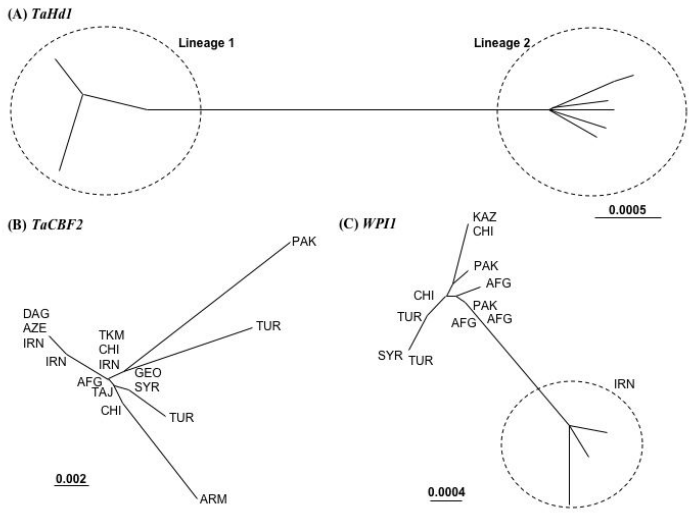

Fig. 2. Phylogenetic trees based on the nucleotide variation of protein-encoding genes. (A) *TaHd1*. This type of the phylogenetic trees corresponds to the SSR-based tree. (B) *TaCBF2*. (C) *WPI1*. The trees on (B) and (C) are inconsistent with the SSR-based tree.

ACKNOWLEDGEMENTS

We thank Drs. J. Valkoun, J. Konopka and L. Visser (IPK, USDA) for the *Ae. tauschii* accessions. This work was supported by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan to TK (no. 17201045) and by grants from Hyogo Science and Technology Association and Elizabeth Arnold Fuji Foundation to ST.

REFERENCES

1. Kihara H (1944) Discovery of the DD-analyser, one of the ancestors of *Triticum vulgare*. Agric Hortic 19: 889-890 (in Japanese)
2. McFadden ES, Sears ER (1944) The artificial synthesis of *Triticum spelta*. Rec Genet Soc Am 13: 26-27
3. Cadwell KS, Dvorak J, Lagudah ES, Akhunov E, Luo MC, Wolters P, Powell W (2004) Sequence polymorphism in polyploid wheat and their D-genome diploid ancestor. Genetics 167: 941-94
4. Giles RJ, Brown TA (2006) *GluDy* allele variations in *Aegilops tauschii* and *Triticum aestivum*: implications for the origin of hexaploid wheats. Theor Appl Genet 112: 1563-1572
5. Eig A (1929) Monographisch-kritische Übersicht der Gatteung *Aegilops*. Repertorium Specierum Novarum Rgni Vegetabilis. Beihefte 55: 1-228
6. Tsunewaki K (1966) Comparative gene analysis of common wheat and its ancestral species. II. Waxiness, growth habit and awnedness. Jpn J Bot 19: 175-229
7. Nishikawa K, Furuta Y, Wada T (1980) Genetic studies on alpha-amylase isozymes in wheat. III. Intraspecific variation in *Aegilops squarrosa* and birthplace of hexaploid wheat. Jpn J Genet 55: 325-336
8. Nakai Y (1979) Isozyme variation in *Aegilops* and *Triticum*. IV. The origin of the common wheats revealed from the study on esterase isozymes in synthesized hexaploid wheats. Jpn J Genet 54: 175-189
9. Dudnikov AJ (1998) Allozyme variation in Transcaucasian population of *Aegilops squarrosa*. Heredity 80: 248-258
10. 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
11. Saeidi H, Rahiminejad MR, Vallian S, Heslop-Harrison JS (2006) Biodiversity of diploid D-genome *Aegilops tauschii* Coss. In Iran measured using microsatellites. Genet Resource Crop Evol 53: 1477-1484
12. Matsuoka Y, Mori N, Kawahara T (2005) Genealogical use for chloroplast DNA variation for intraspecific studies of *Aegilops tauschii* Coss. Theor Appl Genet 111: 265-271
13. Matsuoka Y, Takumi S, Kawahara T (2007) Natural variation for fertile triploid F_1 formation in allohexaploid wheat speciation. Theor Appl Genet 115: 509-518
14. Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier MH, Leroy P, Ganal MW (1998) A microsatellite map of wheat. Genetics 149: 2007-2023
15. Pestsova E, Korzun V, Goncharov NP, Hammer K, Ganal MW, Röder MS (2000) Microsatellite analysis of *Aegilops tauschii* germplasm. Theor Appl Genet 101: 100-106
16. Y. Matsuoka, S. Takumi and T. Kawahara. Flowering time diversification and dispersal in central Eurasian wild wheat *Aegilops tauschii* Coss.: genealogical and ecological framework. (submitted)