Theor Appl Genet (2001) 102:463–470

Y. Yasui · S. Nasuda · Y. Matsuoka · T. Kawahara

The Au family, a novel short interspersed element (SINE) from *Aegilops umbellulata*

Received: 15 April 2000 / Accepted: 24 August 2000

Abstract A novel plant short interspersed nuclear element (SINE) was identified in the second intron of the acetyl CoA carboxylase gene of *Aegilops umbellulata* which has been designated "Au", for the host species in which it was discovered. Au elements have a tRNA-related region, direct flanking repeats, and a short stretch of T at the 3′ end, which are features common to Au and previously characterized SINEs. Au elements are detected in the genomes of several monocots and dicots by DNA dot hybridization and are also found in the tobacco genome by database searching. Au elements are present at an especially high copy number (approximately 10^4 copies per haploid genome) in wheat and *Ae. umbellulata*. This suggests a recent amplification of Au in the *Triticum* and *Aegilops* species. In situ hybridization revealed a dispersed distribution of Au elements on wheat chromosomes. Au elements were amplified by PCR from monocot and dicot species and the phylogenetic relationships among Au elements were inferred. This phylogenetic analysis suggests amplification of Au elements in a manner consistent with the retrotransposon model for SINE dispersion. The high copy number of Au elements and their dispersed distribution in wheat are desirable characteristics for a molecular marker system in this important species.

Keywords SINE · Repetitive DNA sequences · Retroposon · Wheat · *Aegilops umbellulata*

Communicated by J. Dvorak

Y. Yasui · T. Kawahara (✉)
Laboratory of Crop Evolution, Graduate School of Agriculture, Kyoto University, Nakajo 1, Mozume-cho, Muko, 617–0001, Japan
Tel: +81-75-921-0652, Fax: +81-75-932-8063

S. Nasuda
Laboratory of Plant Genetics, Graduate School of Agriculture, Kyoto University, Kitashirakawa-oiwake-cho, Sakyou-ku, Kyoto, 606–8502, Japan

Y. Matsuoka
Department of Bioscience, Fukui Prefectural University, 4–1-1 Ken-jyozima, Matsuoka-cho, Yoshida-gun, Fukui, 910–1142, Japan

Introduction

Short interspersed nuclear elements (SINEs) are repetitive DNA sequences abundant in eukaryotic genomes (reviewed in Okada 1991; Deininger and Batzer 1993). Their dispersion (called "retroposition") is mediated by reverse transcription (Rogers 1985). SINEs lack open reading frames and therefore do not code for any enzymes involved in retroposition. SINEs are thought to utilize the enzymatic machinery of retroposition expressed by non-LTR (Long Terminal Repeat) retroposable elements (Luan et al. 1993; Ohshima et al. 1996).

The structure of SINEs is well established. Typically, they include three sequence structures: (1) two internal promoters (boxes A and B) for the endogenous host RNA polymerase III, (2) direct repeats in the flanking host DNA at the site of SINE insertion, and (3) an AT-rich 3′ end (reviewed in Weiner et al. 1986; Deininger 1989). SINEs can be divided into two classes. SINEs in the first class are homologous to 7SL RNA (Ullu et al. 1982), an essential component of the signal recognition particle complex for protein translocation (Walter and Blobel 1982). The human Alu family and rodent B1 family are elements that belong to this class. It was proposed that Alu and B1 families may have been derived from 7SL RNA independently (Weiner et al. 1986). SINEs in the second class have a sequence homologous to tRNA in the 5′ region of the element. Most SINEs, except Alu and B1 families, belong to this class. This class of tRNA-derived SINEs was first identified in animal species (Daniels and Deininger 1985; Lawrence et al. 1985; Sakamoto and Okada 1985), and their molecular and evolutionary properties have been studied extensively. It was proposed that the tRNA-related region of SINEs might have originated from a primer tRNA used in reverse transcription of non-LTR type retrotransposons (Ohshima et al. 1993, 1996).

According to the accepted model for SINE dispersion, SINEs never excise from a site once they have inserted at that site, and their insertion sites are randomly distributed in the genome of their host. Thus, SINEs can be

powerful molecular markers for phylogenetic analysis. Recently, SINEs have been used as a molecular tool in phylogenetic studies (Kido et al. 1991; Murata et al. 1993; Nikaido et al. 1999). In addition to their utility for phylogenetic analyses, SINEs can be useful markers in molecular mapping. Waugh et al. (1997) showed that analysis of retrotransposon-based sequence-specific amplification polymorphisms (S-SAP) is an effective method for genome analysis in barley. SINE-based S-SAP analysis will be an effective technique in plants, if SINEs are abundant and dispersed in a target genome. However, only three SINE families have been reported in plant species; Ts in tobacco (Yoshioka et al. 1993), p-SINE in rice (Mochizuki et al. 1992), and S1 in oilseed rape (Deragon et al. 1994).

This report identifies a novel tRNA-derived SINE in the gene for acetyl-CoA carboxylase (ACCase: E.C. 6.4.1.2) in *Aegilops umbellulata*. We named this novel SINE "Au", for the host in which it was discovered, ***A****e.* ***u****mbellulata*. The aims of this study were to characterize Au at the molecular level and to determine the phylogenetic relationship of plant species that carry Au elements. Thus, database searches in the DDBJ/EMBLE/GENBANK, DNA dot-blot analysis, and *in situ* hybridization were carried out using target DNA of monocot and dicot plant species. Furthermore, Au elements were amplified by PCR from monocot and dicot species and their nucleotide sequences determined and compared. These data were used to infer the phylogenetic relationships and evolutionary processes relevant to Au elements and its plant host species.

Materials and methods

Plant materials and total DNA extraction

Plant materials analyzed in this study are listed in Table 1. Total DNA was extracted from each plant using Plant DNAzol (Life Technologies).

Table 1 Plant materials used in this study

Species	Common name	Family
Monocot		
Triticum aestivum cv. Chinese Spring	Wheat	Gramineae
Aegilops umbellulata (KU 2777)[a]		Gramineae
Ae. umbellulata (KU 12180)[a]		Gramineae
Secale cereale	Rye	Gramineae
Hordeum vulgare	Barley	Gramineae
Avena sativa	Oat	Gramineae
Zea mays	Maize	Gramineae
Oryza sativa cv. Babawee	Rice	Gramineae
Dicot		
Nicotiana tabacum cv. Samsun	Tobacco	Solanaceae
Lycopersicon esculentum	Tomato	Solanaceae
Pisum sativum	Garden Pea	Leguminosae

[a] Accession number at the Laboratory of Crop Evolution, Graduate School of Agriculture, Kyoto University

PCR amplification and sequencing of the second intron of the ACCase gene of *Ae. umbellulata*

The second intron of the ACCase of two accessions of *Ae. umbellulata*, KU 2777 and KU 12180, was amplified by PCR using standard conditions (Saiki et al. 1988). The PCR primers were ACFW2065 (5′-CCGGAAGTGGATATTTCAC-3′) and ACRV2485 (5′-TGGGTACGGGTTTCAGATG-3′). These primers were designed from the DNA sequence of wheat ACCase (the DDBJ/EMBL/GenBank accession number AF029897) determined by Gornicki et al. (1997). PCR products were recovered from a 1.2% TAE agarose gel and purified with the GENE CLEAN II Kit (BIO 101). Purified PCR products were used in DNA sequencing reactions with the *Taq* Dye Deoxy Terminator Kit, FS (Applied Biosystems). The sequencing primers were the same as the PCR primers. Electrophoresis and data collection were carried out with a 373A DNA sequencer (Applied Biosystems).

Cloning the insertion in the second intron of the ACCase

A 200-bp insertion was re-amplified from the PCR products of the second intron of ACCase of *Ae. umbellulata* (KU 2777) using standard conditions. The amplification primers were AUFW7 (5′-GAGCCTTGGCGCAGTGGTA-3′) and AURV200 (5′-TCTGAAAATGATTCCTAAA-3′). These PCR products were used as a positive control sample (reference) for DNA dot hybridization (see below). The amplified fragment was cloned into the pUC18-*Sma*I-T Overhang Vector (Bayou Biolabs), and the clone was named pAU-1.

DNA-dot hybridization analysis

DNA-dot hybridization analysis was carried out with DNA samples from the plants listed in Table 1, except *Ae. umbellulata* (KU 12180). For dot hybridization, 2 µg, 400 ng, and 80 ng of total plant DNA was transferred to and fixed on a nylon membrane (Hybond N+, Amersham Pharmacia Biotech) using Bio-Dot SF (Bio-Rad Laboratories). As a positive control, 2 ng, 400 pg, and 80 pg of PCR products of the 200-bp Au element insertion were dotted on the same membrane. Hybridization was carried out at 63°C in a mixture containing 5×SSC, 0.1% sarcosyl, 0.02% SDS, and 0.5% Blocking Reagent (Boehringer Mannheim). Dig-dUTP-labeled PCR products from pAU-1 were used as a probe. Two universal primers (M13M4 and M13RV, Takara Shuzo) were used for PCR amplification. The nylon membrane was washed twice at room temperature in 2× SSC, 0.1% SDS for five min and then twice at 63°C in 0.1× SSC, 0.1% SDS for 15 min. DNA was detected by luminescence according to the manufacturer's protocol (Dig Detection Kit; Boehringer Mannheim). The signal intensity of each species was visually approximated by comparing with the positive controls.

PCR amplification, cloning, and sequencing of the Au element

Au elements were amplified by PCR from the plant materials listed in Table 1. The Au elements obtained from these reactions were compared with the Au element from the tobacco genome in the DDBJ/EMBLE/GENBANK database, and it was found that the 3′ region of Au was much more variable than the 5′ region (data not shown). Therefore, a degenerate primer was designed for the 3′ region of the Au element (AURV176; 5′-GGCAGCCCSAGTGCAYKWAGCTCCCG-3′). PCR amplification was performed in a 50-µl reaction containing *Taq* polymerase buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2 mM $MgCl_2$), 200 µM dNTP, 0.3 µM primers, 1.0 unit of *Taq* polymerase (Takara Shuzo), and 50 ng of total DNA for 25 cycles at 94°C for 60 s, 58°C for 30 s, and 72°C for 60 s. The PCR primers were AUFW7 and AURV176. The PCR product was cloned as described above. Cloned Au elements were sequenced with the *Taq* Dye Deoxy Terminator Kit, FS, using M13 universal primers (M13M4 and M13RV, Takara Shuzo).

In situ hybridization

The PCR-amplified Au element was hybridized to metaphase chromosomes, interphase and prophase nuclei of Chinese Spring wheat. The Au element in the ACCase of *Ae. umbellulata* was PCR-labeled using Biotin-16-dUTP (Boehringer Mannheim). As a positive control of hybridization, a cereal-centromere specific clone pGP7 (Presting et al. 1998) was also PCR-labeled by incorporatin g Digoxigenin-11-dUTP (Boehringer Mannheim). Conditions for hybridization, post-hybridization wash, and signal detection were described elsewhere (Schubert et al. 1998). FISH signals were observed by an epifluorescent microscope (Olympus BX60) and captured by a Photometrics SenSys CCD camera interfaced with a Macintosh computer. Images were captured and processed using computer programs IPLab version 3.1 and Photoshop version 4.0.

Phylogenetic analyses

The Au element was amplified by PCR and sequenced from several isolates of wheat (8), rye (9), barley (7), oats (5), maize (8), tobacco (6), and tomato (7). (These DNA sequences are deposited in the DDBJ/EMBLE/GENBANK database under the accession numbers from AB046121 to AB046170.) These 50 sequences were aligned to each other and to the sequence of the Au element in the second intron of the ACCase from *Ae. umbellulata* and the sequences of four tobacco Au elements found in the DDBJ/EMBLE/GENBANK database. The alignment was carried out using CLUSTAL W, version 1.5 (Thompson et al. 1994). After eliminating the common primer sequences, a data matrix was obtained for the 55 sequences ×148 characters. (This data matrix is available from the authors.) Distance estimation with the one parameter method (Jukes and Cauntor 1969) and neighbor-joining analysis (Saitou and Nei 1987) were performed with PAUP, version 4.0b3a (Swofford 1998). Bootstrap analyses (Felsenstein 1985) with a 1000 re-sampling data set were also performed by PAUP. The nucleotide diversity within species (Nei 1987) was estimated with the DnaSP program (Rozas and Rozas 1999).

Results

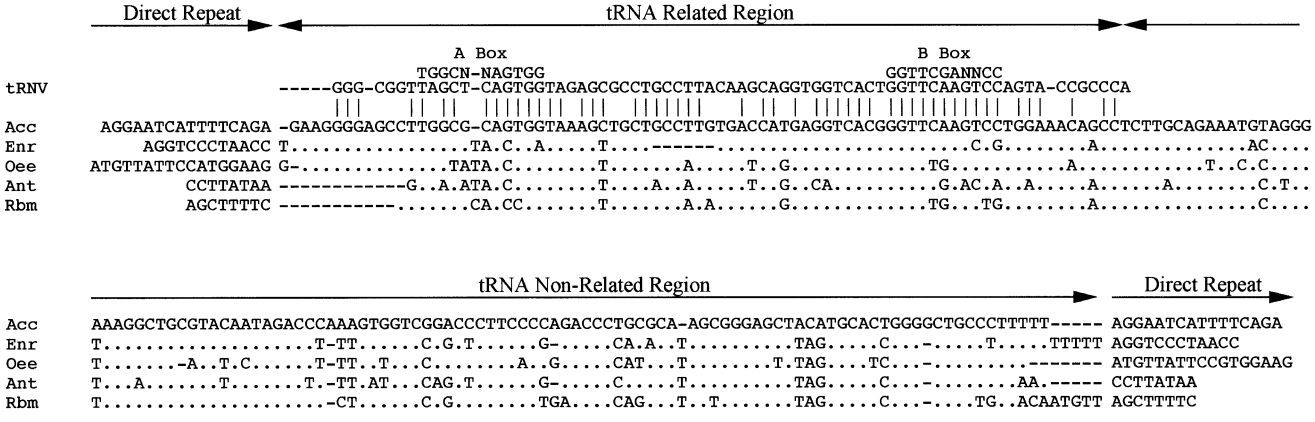
Characterization of the Au element

The nucleotide sequence of the second intron of the ACCase gene was determined from two accessions of *Ae. umbellulata*. One of these two sequences included a 200-bp insertion flanked by direct repeats of a 16-bp sequence (Fig. 1), which we named "Au" (***A**egilops **u**mbellulata*). Database searching in the DDBJ/EMBLE/GENBANK revealed that the 5′ region of Au is homologous to several tRNA genes. The extent of similarity varied between 70% and 80% for segments up to 70 nucleotides long (tRNA-Val, *Porphyra purpurea* chloroplast DNA: 74% similarity in 70 nucleotides; tRNA-Thr, *Campylobacter jejuni*: 80% in 56 nucleotides; and tRNA-Cys, *Nicotiana rustica*: 70% in 69 nucleotides). The nucleotide sequence of the tRNA-Val gene of *P. purpurea* is given in Fig. 1. In the tRNA-related region of Au, two internal promoters of RNA polymerase III (boxes A and B; Galli et al. 1981) are conserved (Fig. 1). Similar to SINEs from rat (ID) and rice (p-SINE1), Au has a short stretch of repeated T residues at the end of the unique sequence region (i.e., the tRNA non-related region; Fig. 1). Database searching also revealed that sequences related to Au were located in non-coding regions of four genes from tobacco: enr-T1, oee2, TA-29, and rbcmtT. These sequences were flanked by direct repeats and have conserved the A and B boxes, though they have a one base pair insertion in their A box (Fig. 1). The unique sequence regions of Au and the four tobacco SINEs are not similar to the sequences of previously reported SINE elements; thus, these sequences can be considered as members of a novel SINE family named the Au family.

Distribution and copy number of Au

DNA dot-blot hybridization showed that Au elements are present at high copy number in the genomes of *Ae. umbellulata* and wheat (Fig. 2a). Assuming 1.6×10^{10} bp in the haploid genome of wheat (Arumuganathan and Earle 1991), wheat contains approximately 10^4 copies of

Fig. 1 The nucleotide sequences of Au elements. *Acc* indicates the Au element in the second intron of the ACCase gene of *Ae. umbellulata*. *Enr*, *Oee*, *Ant*, and *Rbm* indicate Au elements found in the enr-T1, oee2, TA29, and rbcmtT genes of tobacco, respectively. *tRNV* indicates the sequence of $tRNA^{Val}$ from the chloroplast genome of *P. purpurea*. Identical nucleotides between RNV and Acc are connected by *vertical lines*. *Dots* indicate nucleotides identical to the Acc sequence. *Dashes* indicate gaps. The DDBJ/EMBLE/GenBank accession numbers are: Acc, AB046134; enr-T1, Y13862; oee2, X58910; TA29, X52283; rbcmtT, U35619; tRNA-Val, U38804



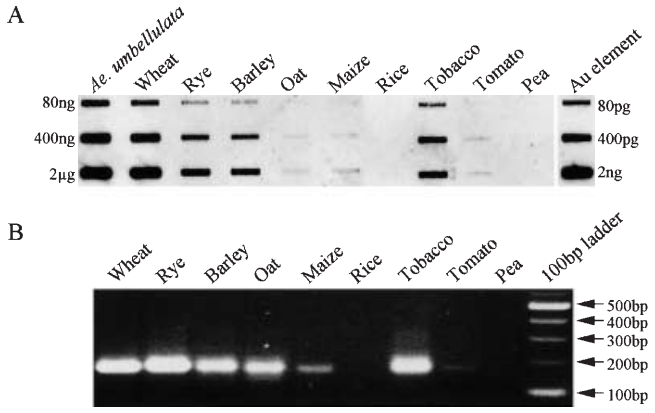


Fig. 2 DNA dot hybridization and PCR amplification of the Au element. **A** DNA dot hybridization profile for the Au element. Total DNA (2 µg, 400ng, 80 ng) was hybridized with the pAU-1 insert labeled by DIG-dUTP. PCR products of Au elements in the second intron of ACCase (2 ng, 400 pg, 80 pg) were used as positive controls (the right column). **B** PCR amplification for Au elements from nine cultivated species. A 100-bp ladder is loaded in the lane on the right

Au in its haploid genome. Au elements are present at a moderate copy number (about 10^2 per haploid genome) in the genomes of rye, barley, and tobacco. It should be noted, taking into account the high divergence of Au elements between *Ae. umbellulata* and the dicot plants, that the copy numbers of Au elements in the dicot plants might be higher than estimated. In oats, maize, and tomato, faint signals are detected (copy numbers are less than 10), but no signal is detected in rice and pea. As shown in Fig. 2b, PCR products of the expected size (approximately 170 bp) are detected in all plant materials except rice and pea. These data indicate that Au elements are not present in the genomes of rice and pea.

Chromosomal distribution of Au elements in wheat

The distribution of Au elements on wheat chromosomes was examined using FISH. FISH signals were faint but detectable on every metaphase chromosome of wheat (Fig. 3). No major clustering of FISH signals was apparent on wheat chromosomes. The distribution of FISH signals did not appear to be associated with a particular genome of the allohexaploid, nor with the heterochromatic regions of wheat chromosomes (Gill et al. 1991). These results strongly suggest that Au elements are randomly dispersed in the wheat genome. However, it is noted that NORs (nucleolar organizing regions) and centromeric regions are less populated by Au elements. This hybridization profile contrasts with the previously reported distribution of $S1_{Bn}$ elements in *Brassica napus*, which are predominantly located in the pericentromeric

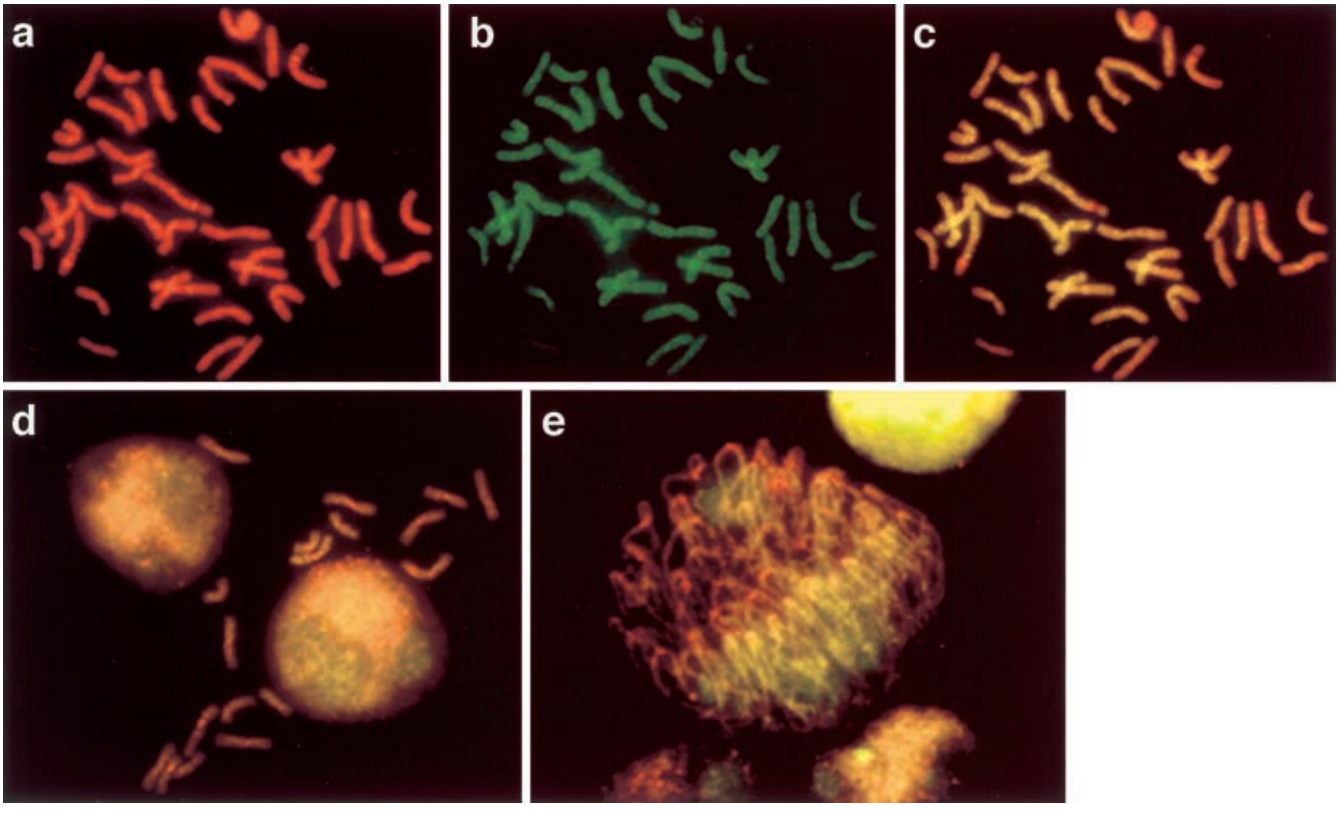


Fig. 3 Dispersed distribution of Au elements in wheat genome revealed by FISH analysis. PI-stained metaphase chromosomes (**a**) were simultaneously hybridized with the Au element and centromere probes. Hybridization sites of the Au element probe were detected by streptavidin-FITC (**b**). Weak green signals were observed along every chromosome. An overlay of the images in (**a**) and (**b**) indicates that Au elements are under-represented in the NORs (**c**). Neither interphase (**d**) nor prophase (**e**) nuclei show intense FISH signals of Au elements in contrast to the localized hybridization sites of the centromere probe (*red spots*)

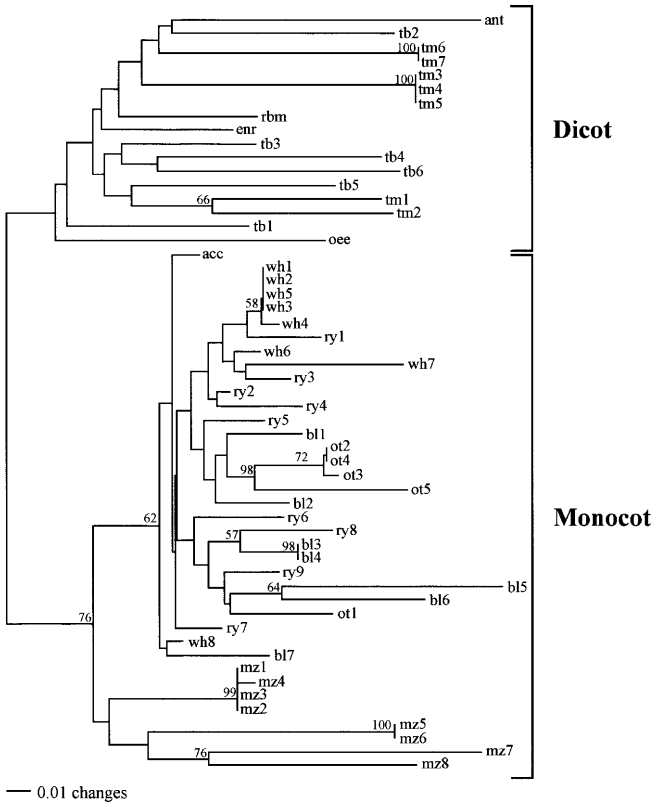


Fig. 4 Neighbor-joining tree for 55 Au elements. *Numbers* above the nodes indicate bootstrap values for 1000 replicate analyses. Bootstrap values below 50% are not shown. The root was assumed at the midpoint of the tree. Abbreviations; *Acc* ACCase, *Enr* enr-T1, *Oee* oee2, *Ant* TA29, *Rbm* rbcmtT, *wh* wheat, *ry* rye, *bl* barley, *ot* oats, *mz* maize, *tb* tobacco, *tm* tomato

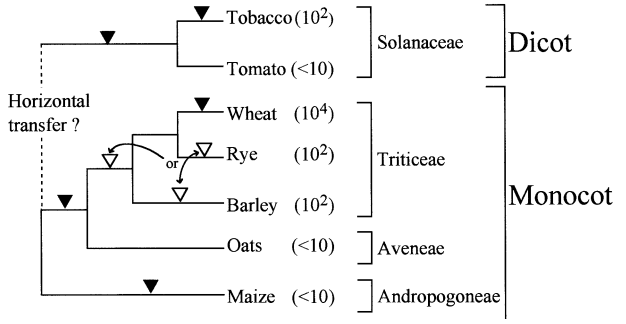


Fig. 5 Evolutionary history of the Au element. *Triangles* indicate an amplification of the Au element suggested from DNA dot hybridization and phylogenetic analyses. We could not determine the positions of amplification events in rye and barley (see text). These ambiguous positions are denoted by *open triangles*. The phylogenetic relationships among monocot species were based on Clayton and Renvoize (1986)

Table 2 The nucleotide diversity (π) within species

Species	N[a]	π
Wheat	9	0.041 (0.013)[b]
Rye	9	0.066 (0.007)
Barley	7	0.113 (0.016)
Oat	5	0.070 (0.021)
Maize	8	0.137 (0.026)
Tobacco	10	0.175 (0.014)
Tomato	7	0.163 (0.026)

[a] Sample size
[b] Standard deviation in parentheses

and NOR regions (Goubely et al. 1999). The Au element probe also hybridized to interphase and prophase nuclei (Fig. 3d, e). Hybridization signals were scattered on the extended chromosome structure, while the positive control probe pGP7 hybridized to the centromeric regions.

Phylogenetic relationships among Au elements

A sample of 55 Au sequences from seven plant species was determined and analyzed for their evolutionary interrelationships. A neighbor-joining tree was deduced and is shown in Fig. 4. The neighbor-joining tree indicates the existence of two major groups; one is a monocot cluster and the other is a dicot cluster (Fig. 4). These two clusters are separated by a high bootstrap value of 76%. In the monocot group, Au elements from maize form a monophyletic group (bootstrap value of 40%), and wheat, rye, barley and oat sequences are clustered into a large group (bootstrap value of 62%), though the bootstrap values were low. This indicates independent amplification of the Au element in maize and in a common ancestor of wheat, rye, barley and oats. In the dicot group, Au elements from tobacco and tomato were not completely separated. This indicates amplifications of the Au element in a common ancestor of tobacco and tomato.

Table 2 shows the intra-specific diversity of the Au element for wheat, rye, barley, oats, maize, tobacco and tomato. The intra-specific diversity is much lower within wheat than within the other species. This indicates recent amplifications of Au elements within the wheat genome. Intra-specific divergence was higher within tobacco and tomato than within the other species. These high intra-specific divergences within two dicot species probably indicate ancient amplification of the Au element in dicots.

Discussion

Molecular structure and distribution of the Au element

Three SINE-like retroposons were previously reported in plant genomes, namely Ts (Yoshioka et al. 1993), p-SINE (Mochizuki et al. 1992), and S1 (Deragon et al. 1994) elements. The novel element reported here, Au, shares three structural characteristics with these previously identified SINEs from plants. First, the Au element has a tRNA-related sequence in its 5′ region (Fig. 1). Two internal promoters for RNA polymerase III (the A and B boxes) are also conserved (Fig. 1). Within this

tRNA-related region, the B box (GGTTCGANNCC) is highly conserved for almost all Au sequences investigated in this study (data not shown). Unfortunately, the A-box sequence was in the amplifying primer (AUFW 7); therefore, the extent of its conservation could not be determined at this time. Second, the Au elements from *Ae. umbellulata* were flanked by direct repeats; this is a general property of SINEs, and was also observed for Au elements present in the database, including one in the wheat ACCase gene and four Au elements in the tobacco genome (Fig 1). The third structural characteristic common to SINEs and the Au element is a short stretch of repeated T nucleotides at the 3′ end: repeated T residues were detected on Au elements in the wheat ACCase gene and in tobacco enrT1 and oee2 (Fig. 1). A short stretch of repeated T is also observed at the 3′ ends of rat ID and rice p-SINE elements. Such repeated T residues can function as a termination signal for RNA polymerase III (DeChiara and Brosius 1987). Database searching did not reveal any similarity in the unique sequence regions of previously reported SINEs and the members of the Au element family; thus, the Au family is considered to be a novel SINE family found in plant genomes.

The Au element family is widely represented in plant genomes, and two significant features of its distribution are noted here. Au elements were detected in monocot and dicot plants by dot hybridization (Fig. 2a, b). In contrast, the three previously reported plant SINEs are endemic; i.e., p-SINE1 was found only in the genus *Oryza* (Umeda et al. 1991; Mochizuki et al. 1992; Mochizuki et al. 1993), Ts are found only in the order Tubiflorae (Yoshioka et al. 1993), and S1 is found only in the family Cruciferae (Lenoir et al. 1997). In addition, Au elements are abundant in the wheat genome; importantly, the Au element is the first SINE identified in wheat. Thus, the Au element has potential as a new molecular marker in wheat and wheat-related species.

High potential of the Au element as a wheat genetic marker

Au elements are present at a higher copy number in *Ae. umbellulata* and wheat genomes than in other plant species (Fig. 2). This strongly indicates that the Au element underwent an explosive increase in its copy number in *Triticum* and *Aegilops*. Consistent with this hypothesis, there is a low level of nucleotide diversity of the Au element in wheat (Table 2). Furthermore, the Au element is present in only one of the two alleles of the ACCase gene in *Ae. umbellulata*. This allelic variation suggests a recent insertion of an Au element, and implies that the Au element has a high capacity for retroposition in the *Triticum* and *Aegilops* genomes. Thus, the Au element has two advantageous qualities for use as a wheat molecular marker: a high copy number in wheat genomes and recent transposition, which can lead to easy detection of insertional polymorphism between closely related taxa. Moreover, Au elements are dispersed throughout the wheat genome, as shown by FISH analysis (Fig. 3), which is another desirable trait for a molecular genetic marker.

In future studies, we propose to apply S-SAP analysis in *Triticum* and *Aegilops* using the Au element. A recent study demonstrated that S-SAP analysis employing the BARE-1-like retrotransposon is an effective method to reveal genetic divergence in barley (Waugh et al. 1997). Therefore, S-SAP analysis with Au could facilitate the construction of a fine genetic map of wheat and inference of the phylogenetic relationships between *Triticum* and *Aegilops* (Zhang et al., in preparation).

Possibility of horizontal transfer of the Au element

Although the Au element is widely distributed in monocot and dicot plants, it is not found in the genomes of rice and pea by either DNA dot-blot analysis or PCR amplification. This result suggests the possibility of a horizontal transfer of the Au element between an ancestor of wheat-rye-barley-oats-maize and an ancestor of tobacco-tomato (Fig. 5). However, there are two alternative explanations for the absence of Au in rice and pea. First, the Au element could have accumulated a large number of mutations, if they have not been amplified in rice and pea for a long time; these mutations might interfere with detection of the element by hybridization or PCR amplification. A second possibility is that the Au element was eliminated in these genomes by stochastic loss. As is the case of the R1 and R2 retrotransposon in insects (Jakubczak et al. 1991), a low copy number of retrotransposons can be eliminated from the genome by stochastic loss. The process of horizontal transfer of the Au element will be better understood after analysis of additional information on the distribution of Au elements in the plant kingdom.

Evolution of the Au element

The above discussion suggests several events involving amplifications and horizontal transfers of the Au element within the plant kingdom. These hypothetical events are presented graphically in Fig. 5. An amplification of Au elements in wheat is indicated by its high copy number (Fig. 2) and low intra-specific divergence (Table 2), and this is supported by the observed allelic variation in the second intron of ACCase in *Ae. umbellulata*. Two clusters composed of maize and of wheat-rye-barley-oats (Fig. 4) suggest independent amplifications within the maize lineage and in a common ancestor of wheat, rye, barley, and oats. Au is present at a higher copy number in rye and barley than in oats, which indicates an amplification of Au. However, it is not possible to determine whether there was a single amplification event in an ancestor of rye and barley or if there were two independent amplification events, one in each species. Within the di-

cot plants tomato and tobacco, Au elements were not in separate groups. This implies that the Au element was amplified in a common ancestor of tobacco and tomato. Furthermore, the high copy number of Au in tobacco indicates that it is likely to have undergone amplification in this species.

Lenoir et al. (1997) reported that S1 elements in Cruciferae can be divided into 19 clusters with high (>70%) bootstrap values. They also indicate that most of these clusters were composed of species-specific elements and correspond to recent species-specific bursts. In our analyses, only maize sequences were clustered in one group (Fig. 4), whereas sequences from the other species were not clustered into species-specific groups. This difference may indicate that Au and S1 elements experience amplification in a different manner or by a different mechanism.

Deininger et al. (1992) proposed two models for the amplification of SINE elements, the master gene model and the transposon model. In the master gene model, it is assumed that only a few master copies have the ability to retropose, and most elements newly inserted in the host genome are not able to retropose. According to this model, a mutation that occurs in a master copy will be shared by all subsequent copies; thus, it is expected that SINE sequences will form a species-specific cluster if amplification proceeds according to this model (Deininger et al. 1992; Deininger and Batzer 1993). As suggested by Lenoir et al. (1997) this may be the case with the S1 element in the Cruciferae. On the other hand, in the transposon model, it is assumed that newly inserted progenitors possess the ability to retropose. Under the transposon model, numerous types of sequences will be increased within each species, and species-specific mutations (i.e., diagnostic mutations) would not be detectable. This would interfere with the ability to detect species-specific clusters (Bains 1986; Deininger et al. 1992; Deininger and Batzer 1993).

The Au element demonstrates a small number of species specific-clusters and a low bootstrap value in the phylogenetic tree (Fig. 4), and thus would seem to fit the transposon model more closely than the master gene model. This possibility can be tested and confirmed by studies on the in vitro and in vivo transcription activity of Au, including determination of the sequences of Au transcripts. Such studies will help define the manner of amplification of the Au element in the plant kingdom.

Acknowledgements Special thanks are due to Dr. N. Mori for providing us with the total DNA of rice. We are also grateful for numerous suggestions received from Drs. T. Ohsako and K. Tsuji. This study was supported by the Sumitomo foundation and Grants-in-Aid for Scientific Research (No. 10836011) from the Ministry of Education, Science, Sports, and Culture of Japan. Contribution number 103 from the Laboratory of Crop Evolution, Graduate School of Agriculture, Kyoto University.

References

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:208–218

Bains W (1986) The multiple origins of human sequences. J Mol Evol 23:189–199

Clayton WD, Renvoiz SA (1986) GENERA GRAMINUM. Grasses of the World. Her Majesty's Stationery Office, London

Daniels GR, Deininger PL (1985) Repeat sequence families derived from mammalian tRNA genes. Nature 317:819–822

DeChihara TM, Brosius J (1987) Neral BC1 RNA: cDNA clones revea l nonrepetitive sequence content. Proc Natl Acad Sci USA 84: 2624–2628

Deininger PL (1989) SINEs: short interspersed repeated DNA elements in higher eucaryotes. In: Berg DH, Howe MM (eds) Mobile DNA. American Society of Microbiology, Washington, D.C., pp 619–636

Deininger PL, Batzer MA (1993) Evolution of retroposons. Evol Biol 27:157–196

Deininger PL, Batzer MA, Hutchinson CA, Edgell MH (1992) Master genes in mammalian repetitive DNA amplification. Trends Genet 8:307–311

Deragon JM, Landry BS, Pélissier T, Tutois S, Tourmente S, Picard G (1994) An analysis of retroposition in plants based on a family of SINEs from *Brassica napus*. J Mol Evol 39: 378–386

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783-791

Galli G, Hofstetter H, Birrnstiel ML (1981) Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. Nature 294:626–631

Gill BS, Friebe B, Endo TR (1991) Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). Genome 34:830–839

Gornicki P, Faris J, King I, Podkowinski J, Gill B, Haselkorn R (1997) Plastid-localized acetyl-CoA carboxylase of bread wheat is encoded by a single gene on each of the three ancestral chromosome sets. Proc Natl Acad Sci USA 94: 14179–14184

Goubely C, Arnaud P, Tatout C, Heslop-Harrison JS, Deragon JM (1999) S1 SINE retroposons are methylated at symmetrical and non-symmetrical positions in *Brassica napus*: identification of a preferred target site for asymmetrical methylation. Plant Mol Biol 39:243–255

Jakubczak JL, Burke WD, Eickbush TH (1991) Retrotransposable elements R1 and R2 interrupt the rRNA genes of most insects. Proc Natl Acad Sci USA 88:3295–3299

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In Munro HN (ed) Mammalian protein metabolism. Academic Press, New York, pp 21–132

Kido Y, Aono M, Yamaki T, Matsumoto K, Murata S, Saneyoshi M, Okada N (1991) Shaping and reshaping of salmonid genomes by amplification of tRNA-derived retroposons during evolution. Proc Natl Acad Sci USA 88:2326–2330

Lawrence CB, McDonnnell DP, Ramsey WJ (1985) Analysis of repetitive sequence elements containing tRNA-like sequences. Nucleic Acids Res 13:4239–4252

Lenoir A, Cournoyer B, Warwick S, Picard G, Deragon JM (1997) Evolution of SINE S1 retroposons in Cruciferae plant species. Mol Biol Evol 14:934–941

Luan DD, Korman MH, Jakubczak JL, Eickbush TH (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR transcription. Cell 72:595–605

Mochizuki K, Umeda M, Ohtsubo H, Ohtsubo E (1992) Characterization of a plant SINE, p-SINE-1, in rice genomes. Jpn J Genet 57:155–166

Mochizuki K, Ohtsubo H, Hirano HY, Sano Y, Ohtsubo E (1993) Classification and relationships of rice strains with the AA ge-

nome by identification of transposable elements at nine loci. Jpn J Genet 68:205–217

Murata S, Takasaki N, Saitou M, Okada N (1993) Determination of the phylogenetic relationships among Pacific salmonids by using short interspersed elements (SINEs) as temporal landmarks of evolution. Proc Natl Acad Sci USA 90:6995–6999

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nikaido M, Rooney AP, Okada N (1999) Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: hippopotamuses are the closest extant relatives of whales. Proc Natl Acad Sci USA 96: 10261–10266

Ohshima K, Koishi R, Matsuo M, Okada N (1993) The 3′ ends of tRNA-derived short interspersed repetitive elements. Mol Cell Biol 16:3756–3764

Ohshima K, Hamada M, Terai Y, Okada N (1996) Several short interspersed repetitive elements (SINE) in distant species may have originated from a common ancestral retrovirus: characterization of a squid SINE and a possible mechanism for the generation of tRNA derived retroposons. Proc Natl Acad Sci USA 90:6260–6264

Okada N (1991) SINEs: short interspersed elements of the eukaryotic genome. Trends Ecol Evol 6:358–361

Presting GG, Malysheva L, Fuchs J, Schubert I (1998) A *TY3/GYPSY* retrotransposon-like sequence localized to the centromeric regions of cereal chromosomes. Plant J 16:721–728

Rozas J, Rozas R (1999) DnaSP version 3: an integrated program for molecular population genetics and molecular evolutionary analysis. Bioinformatics 15:174–175

Rogers JH (1985) The origin and evolution of retroposons. Int Rev Cytol 93:187–279

Saiki RK, Celfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sakamoto K and Okada N (1985) Rodent type-2 Alu family, rat identifier sequence, rabbit C family, and a bovine or goat 73-bp repeat may have evolved from tRNA genes. J Mol Evol 22:134–140

Schubert I, Shi F, Fuchs J, Endo, TR (1998) An efficient screening for terminal deletions and translocations of barley chromosomes added to common wheat. Plant J 14:489–495

Swofford DL (1998) PAUP*4.0: phylogenetic analysis using parsimony, beta test version 4.03b. Sinauer, Sunderland, Massachusetts

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680

Ullu E, Murphy S, Melli M (1982) Human 7S RNA consists of a 140 nucleotide middle repetitive sequence inserted in an Alu sequence. Cell 29:195–202

Umeda M, Ohtsubo H, Ohtsubo E (1991) Diversification of the rice *Waxy* gene by insertion of mobile DNA elements into introns. Jpn J Genet 66:569–586

Walter P, Blobel G (1982) The signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. Nature 299:691–698

Waugh R, Mclean k, Flavell AJ, Pearce SR, Kumar A, Thomas BBT, Powell W (1997) Genetic distribution of Bare-1-like retroposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). Mol Gen Genet 253:687–694

Weiner AM, Deininger PM, Efstratiadis A (1986) Non viral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annu Rev Biochem 55:631–661

Yoshioka Y, Matsumoto S, Kojima S, Ohshima K, Okada N, Machida Y (1993) Molecular characterization of a short interspersed repetitive element from tobacco that exhibits sequence homology to specific tRNAs. Proc Natl Acad Sci USA 90:6562–6566