Theor Appl Genet (2006) 113:261–269
DOI 10.1007/s00122-006-0291-5

ORIGINAL PAPER

Phylogenetic analysis of the rDNA intergenic spacer subrepeats and its implication for the domestication history of foxtail millet, *Setaria italica*

Kenji Fukunaga · Katsuyuki Ichitani · Makoto Kawase

Received: 10 November 2005 / Accepted: 9 April 2006 / Published online: 6 May 2006
© Springer-Verlag 2006

Abstract We sequenced ribosomal DNA intergenic spacer subrepeats and their flanking regions of foxtail millet landraces from various regions in Europe and Asia and its wild ancestor to elucidate phylogenetic differentiation within each of types I–III found in our previous work and to elucidate relationships among these three types. Type I was classified into seven subtypes designated as Ia–Ig based on subrepeat sequences; C repeats downstream of those subrepeats are also polymorphic. Of these, subtypes Ia–Id and Ig were found in foxtail millet landraces. Subtypes Ia and Ib were distributed broadly throughout Asia and Europe. Subtype Ic was distributed in China, Korea and Japan. Subtype Id has a 20-bp deletion in subrepeat 3 and has a unique C repeat sequence. This subtype was found in a morphologically primitive landrace group from Afghanistan and northwestern Pakistan and differed greatly from other type I subtypes, implying that these landraces were domesticated independently. Subtypes Ig was found in a landrace from Pakistan and Ia and Ie–Ig were in six wild ancestor accessions. Type II was also highly polymorphic and four subtypes were found and designated as subtypes IIa–IId, but sequence analyses indicated type III as monomorphic. The present work indicates that type III should be classified as a subtype of type II (subtype IIe). Sequence polymorphism of subrepeats of types I–III indicated that subrepeats of subtype IIa are greatly divergent from others. Relationships among types I–III are much more complicated than anticipated based on previous RFLP work.

Communicated by O. Savolainen

K. Fukunaga (✉)
International Research Center for Japanese Studies,
3-2 Oeyama-cho, Goryo, Nishikyo-ku,
Kyoto, 610-1192, Japan
e-mail: kjfukunaga@hotmail.com

Research Institute for Humanity and Nature,
457-4, Motoyama, Kamigamo, Kita-ku,
Kyoto, 603-8047, Japan

K. Ichitani
Faculty of Agriculture, Kagoshima University,
1-21-24 Korimoto, Kagoshima, 890-0065, Japan

M. Kawase
National Institute of Agrobiological Sciences,
2-1-2, Kannondai, Tsukuba, Ibaraki, 305-8602, Japan

Introduction

Foxtail millet, *Setaria italica* (L.) P. Beauv. ssp. *italica* is one of the oldest cereals in Eurasia. Archaeological remains of this crop were found at Peiligang and Cishan sites near the Yellow River dating back to 5,000–6,000 BC (Li and Wu 1996), and at prehistoric sites in Europe (Küster 1984) and in the Transcaucasus (Lisitsina 1976). It has been used in various ways peculiar to different areas of Eurasia (Sakamoto 1987); it is thought to have played an important role in the Old World's early agriculture.

The geographical origin of foxtail millet remains a controversial issue. Cytological and genetic studies indicate that the wild ancestor of this crop is *S. italica* ssp. *viridis* (Kihara and Kishimoto 1942; Li et al. 1945; Le Thierry d'Ennequin et al. 2000; Wang et al. 1995). The geographical origin of domesticated ssp. *italica* cannot be determined by the distribution of ssp. *viridis* (Wang et al. 1995; Le Thierry d'Ennequin et al. 2000), which is commonly found in various areas in Europe and Asia

and colonizing in the New World. Various hypotheses have advanced monophyletic and polyphyletic origins. Vavilov (1926) inferred that the principal center of diversity of foxtail millet is East Asia, including China and Japan. Harlan (1975) suggested independent domestication in China and Europe based on archaeological evidence. Several researchers have supported this hypothesis based on archaeological, isozymes, 5S rDNA and morphological evidence (de Wet et al. 1979; Jusuf and Pernes 1985; Benabdelmouna et al. 2001; Li et al. 1995a, b, 1998). For example, Li et al. (1995b) concluded that foxtail millet was domesticated independently in China and Europe and landraces in Afghanistan and Lebanon had been domesticated separately in recent times because they had primitive characteristics such as numerous tillers with small panicles. In contrast to the hypotheses of multiple origins, Sakamoto (1987) suggested that foxtail millet originated somewhere in Central Asia, Afghanistan, Pakistan and northwestern India because strains with less restricted compatibility (Kawase and Sakamoto 1987) and with primitive morphological traits were found there. This hypothesis was unique in treating China as a secondary center of origin of foxtail millet. This phylogenetic mystery might be solved only by studies of genetic relationships among landraces of foxtail millet and their geographical differentiation patterns based on DNA polymorphism.

Ribosomal DNA (rDNA) in plants, as in animals, is arranged in tandem arrays and contains genes for 25 S, 18 S and 5.8 S rRNAs. These genes are highly conserved among species, but spacer regions such as internal transcribed spacers (ITS) and intergenic spacers (IGS) of the gene are less conserved than the genes themselves. Spacer regions are often used for constructing phylogenetic trees among species (Buckler and Holtsford 1996; Sallares and Brown 2004) and detecting intraspecific polymorphisms (e.g., Saghai-Maroof et al. 1984; Sano and Sano 1990). In particular, IGS has highly variable length, even within species or among cultivars, because it contains subrepeats which vary in repeat number (Rogers and Bendich 1987). We also analyzed rDNA RFLP in foxtail millet, cloned a repeat unit (Fukunaga et al. 1997), determined the sequence of the intergenic spacer, and identified the polymorphic region (Fukunaga et al. 2005). We found three major types designated as types I–III: type I was distributed broadly from East and Southeast Asia through Nepal, northern Pakistan, Afghanistan and Central Asia to Europe; type II was distributed predominantly in the Nansei Islands of Japan, Taiwan, the Philippines and rarely in continental parts of Eurasia; and type III was distributed frequently in India and sporadically in East Asia. Type I is ca. 300 bp shorter in a ribosomal DNA repeat unit than types II or III, whereas type III differs from type II in having an additional *Bam*HI site in the IGS (Fukunaga et al. 1997). Sequencing analysis showed unequivocally that differences between these three types are attributable to length differences caused by different numbers of subrepeats in the IGS and a single nucleotide substitution in a subrepeat (Fukunaga et al. 2005).

This study analyzed sequence polymorphism within and between RFLP types of rDNA to reveal genetic differentiation of foxtail millet from various regions. In particular, we examined the variation detected within type I. Based on that result, we demonstrate the relationships between types and between subtypes and discuss the domestication history of foxtail millet.

Materials and methods

Plant materials

We used 40 accessions of foxtail millet landraces that were originally collected from various regions (Table 1, Fig. 1). Foxtail millet is predominantly self-pollinating species and these accessions have been maintained by self-pollinating at Kyoto University and/or National Institute of Agrobiological Resources, Tsukuba, Japan. Most were used for RFLP analysis in our previous work and classified into three major types: types I–III (Fukunaga et al. 1997). They included 28 type I, 6 type II, and 6 type III accessions. Here, we used many type I accessions because they were distributed broadly in the Temperate Zone including hypothetical geographical origin(s). Two other types, types II and III, were respectively found predominantly in the Taiwan–Philippines Islands and South Asia. We also analyzed six ssp. *viridis* accessions, three from East Asia, two from Pakistan, and one from Turkey, to compare them with the foxtail millet sequence patterns.

DNA extraction, PCR and sequencing analysis

DNA was extracted from bulked seedlings from each accession as reported in our previous work (Fukunaga et al. 1997). The region containing subrepeats was chosen for sequence analyses (see Fig. 2). The region was amplified using a primer combination of IGS3 (5′-TAT GACTGAACGCCTCTAAG-3′) and IGS2 (5′-CACC TCATCACTTCCTC AT-3′) or IGS8 (5′-CACCATGT AACTCACAATGGAC-3′) and IGS4 (5′-CCCCCA CCCAGAAGTT-3′) (Fig. 2). The PCR conditions were 95°C for 5 min, 35 cycles of 94°C for 1 min, 60°C for 1 min and 72°C 2 min, followed by a final extension of

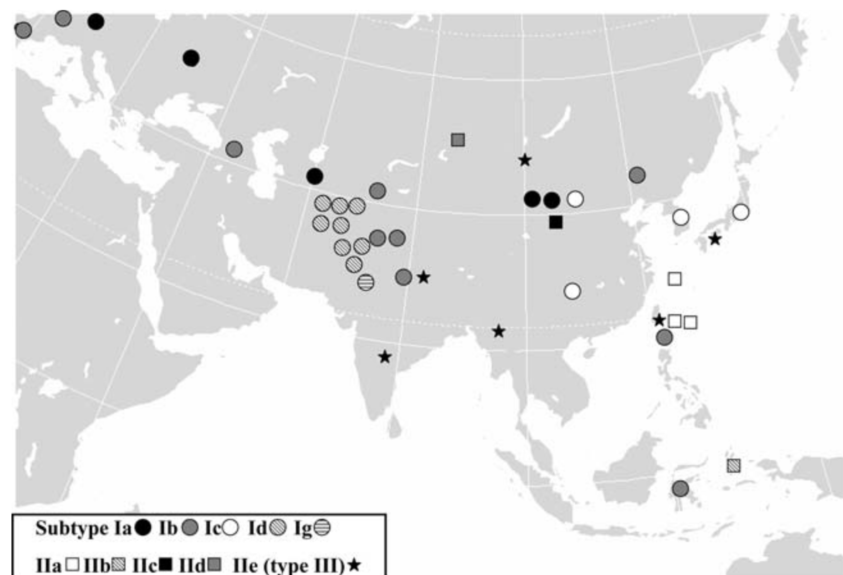

Fig. 1 Geographical distribution of the materials used in the present study and their rDNA type

72°C 5 min. The PCR mixture (50 μl) contained 50 ng template DNA, 200 μM of each d NTP, 0.2 μM of primers, 2 units Ex*Taq* DNA polymerase (Takara Bio Inc.), 1 mM $MgCl_2$ and 1 × buffer. The PCR products were checked using 1.2% agarose gel electrophoresis and purified simply using a Micro Spin S300 HR Column (Amersham Biosciences Corp.); alternatively, the target bands were cut out from the gel and purified by Wizard® SV Gel and PCR Clean-Up System (Promega Corp.). A sequencing reaction of the PCR product with either primer IGS8 or IGS4 for each PCR product was carried out using a BigDye Terminator kit (PE Applied Biosystems) according to the manufacturer's instructions with some modifications. Samples were examined using an ABI310 sequencer at Frontier Science Research Center of Kagoshima University.

Analysis of sequence polymorphism in types I, II and III, and relationships between subrepeats

We analyzed sequence polymorphism between types I, II, and III, and within each length variant to clarify the genetic differentiation of foxtail millet landraces. The CLUSTAL W version 1.8 (Thompson et al. 1994, available at http://www.ebi.ac.uk/clustalw/#) were employed for nucleotide sequence alignment. We classified each type into subtypes based on the sequence of the subrepeats and detected polymorphism on a downstream region containing C repeats (Fig. 2). We compared the sequences from the first nucleotide of subrepeat 1 to the end of C repeats. We also divided each type sequence into subrepeats 1 and 2 and compared all subrepeats of types I–III to elucidate phylogenetic relationships of the subrepeats. We used Network software version 4.101 (available at http://www.fluxus engineering.com) for reduced median network construction.

Results

Sequencing polymorphism in type I and geographical distribution of its variants

Based on the multiple alignments of the subrepeats (Fig. 3), we classified type I into seven subtypes, Ia–Ig, of which Ia–Id and Ig were found in ssp. *italica* and Ia, Ie, If, and Ig in ssp. *viridis*. Relationships among subtypes Ia–Ig are shown in Fig. 4. Subtype Ib differs from subtype Ia in a single nucleotide substitution. Subtype Ic is different from type Ia in four nucleotide substitutions in subrepeats (Figs. 3, 4). Subtype Id is distinct from other subtypes. This subtype has at least three nucleotide substitutions from other type-I subtypes and has a 20-bp deletion in subrepeat 3 (Figs. 3, 4). Each single nucleotide difference was found between subtypes Ie and Ic, between subtypes If and Ia, and between subtypes Ig and Ia (Figs. 3, 4).

Polymorphism found in C repeats (see Figs. 2, 3) was associated to some extent with subrepeat sequences (Fig. 3, Table 1). Among the subtype Ia accessions, four ssp. *italica* accessions had a CCCCCCC (C7) sequence downstream of subrepeats (see Fig. 3); one ssp. *viridis* accession from China had a CCACCCC (C2AC4) sequence. All 10 subtype Ib, 4 Ic, and 8 Id accessions, respectively, revealed C7, CCCCCCT (C6T), and CCCCTCC (C4TC2) sequences. Subtypes Ie and If of ssp. *viridis* had C7 sequences. Type Ig, one ssp. *italica*

Table 1 Materials used in this study: rDNA sequence types and C repeat sequences

Country	Region	Code no.	rDNA type	C repeat
Japan	Kanagawa Pref.	035	Ic	C6T
	Kochi Pref.	012	III	C7
	Okinawa Pref.	021	IIa	C7
Korea	Kyongsang-Pukto	045	Ic	C6T
China	Heilungkiang Prov.	091	Ib	C7
	North China (110001[a])	049	Ia	C7
	North China (110002[a])	050	Ic	C6T
	North China (110007[a])	053	Ia	C7
	North China (110008[a])	054	IIc	C7
	Hunan Prov.	014	Ic	C6T
Mongolia	Uncertain[b]	095	III	C7
Taiwan	Taiwan Is. (Pington)	058	III	C7
	Lan-Hsü Is.	004	IIa	C7
	Lan-Hsü Is.	006	IIa	C7
The Philippines	Itbayat Is.	007	Ib	C7
Indonesia	Halmahera Is.	003	IIb	C7
	Sulawesi Is.	016	Ib	C7
Myanmar	Shan Prov.	024	III	C7
Nepal	Dhuche	017	Ib	C7
	Bhargu	019	III	C7
India	South India[c]	068	III	C7
Pakistan	NWFP (Dir Group)[d]	102	Ig	C6
	NWFP (Chitral Group)[d]	106	Id	C4TC2
	NWFP (Chitral Group)[d]	107	Id	C4TC2
	Gilgit Agency (Chitral Group)[d]	108	Id	C4TC2
	Gilgit Agnecy (Baltistan Group)[d]	110	Ib	C7
	Gilgit Agency (Baltistan Group)[d]	111	Ib	C7
Afghanistan	Jabalsalaj-Zenya	076	Id	C4TC2
	Takhar	077	Id	C4TC2
	Takhar	078	Id	C4TC2
	Badakhshan	079	Id	C4TC2
	Badakhshan	080	Id	C4TC2
Central Asia	Novosibirsk[b]	094	IId	C7
	Kirghizia[b]	081	Ib	C7
	Uzbekistan[b]	082	Ia	C7
	Georgia[b]	029	Ib	C7
	Ukraine[b]	093	Ia	C7
France	Cuon, Maine et Loire	088	Ib	C7
Germany	Uncertain[e]	085	Ia	C7
Spain	Canges de Narcea	084	Ib	C7
S. italica ssp. *viridis*				
Japan	Okayama Pref.	960	Ie	C7
China	Anhoi Prov.	953	If	C7
China	Beijing	954	Ia	C2AC4
Pakistan	Gilgit Agency	126	Ig	C6
Pakistan	Baltistan	127	Ig	C6
Turkey	Uncertain[f]	131	Ig	C6

NWFP North West Frontier Province

[a] All Provided by the National Institute of Agricultural Science, Japan

[b] All Provided by the N.I. Vavilov All-Union Institute of Plant Industry, St. Petersburg, Russian Federation

[c] Provided by the University of Agricultural Science, Bangalore, India

[d] Morphological groups of Pakistani landraces classified by Ochiai et al. (1994)

[e] Provided by Karl Marx Universitat, Leipzig, DDR

[f] Provided by USDA, USA

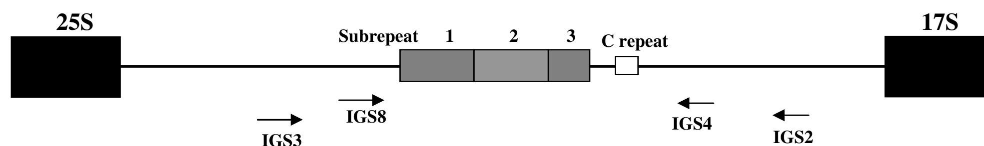

Fig. 2 Structure of the rDNA IGS and positions of primers used for amplification and/or sequencing. The *dark boxes* indicate rRNA gene and *gray boxes* and a *white box* indicate subrepeats and C repeats, respectively. Types II and III have two subrepeats 1–3, whereas type I lacks one complete subrepeat (see the text)

accession from Pakistan and three ssp. *viridis* accessions (two from Pakistan and one from Turkey) exhibited a CCCCCC (C6) sequence. The sequences of subrepeats and C repeats of representative accessions

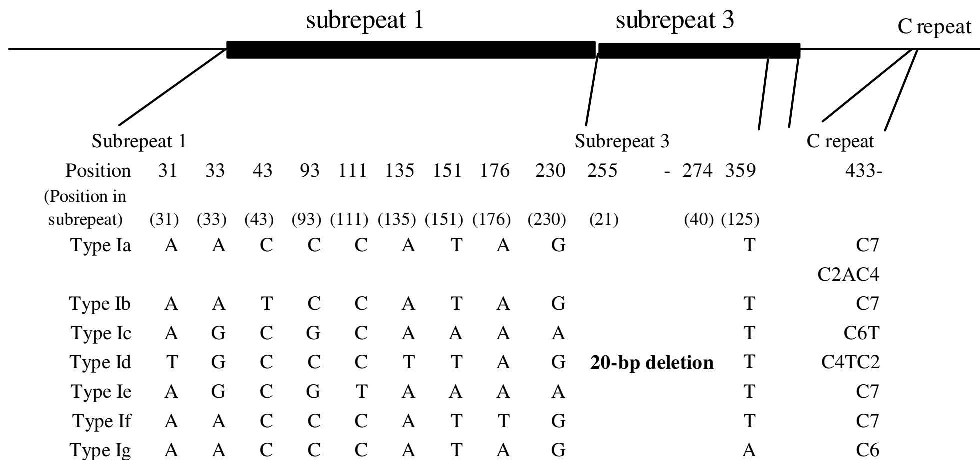

Fig. 3 Alignments of the IGS subrepeats between types Ia and Ig. Type I lacks one complete subrepeat. Here, we designated one subrepeat existing as subrepeat 1 and another missing subrepeat as subrepeat 2. Here, we designated the first nucleotides of subrepeat 1 as position 1. Length from the start of subrepeat 1 to the end of C repeats ranged from 419 to 439 bp. *Numbers in parentheses* indicate the position of sequence of each subrepeat

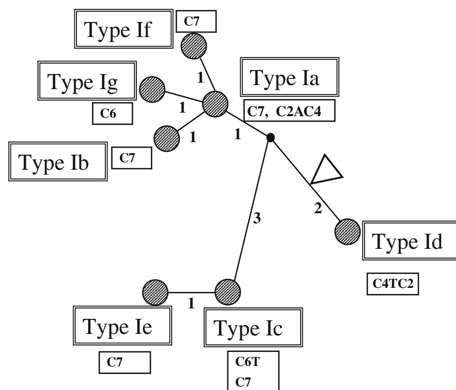

Fig. 4 Relationships of type I subtypes. *Numbers under branches* indicate nucleotide substitutions and a *triangle* indicates a 20-bp deletion. *Boxes* enclose C repeat sequences

were registered in the DDBJ (DDBJ No. AB247504-247510).

Geographical distribution of each subtype was shown in Fig. 1. Subtype Ia was distributed in a broad area ranging from East and Southeast Asia through Nepal and Central Asia to Europe. Subtype Ib is found sporadically in China, northeastern Pakistan, Ukraine, and Europe. Subtype Ic was localized in China, Korea, and Japan. Subtype Id was restricted in a group of landraces that were characterized by primitive morphology found in Afghanistan (Sakamoto 1987) and northwestern Pakistan (the Chitral group, according to Ochiai et al. 1994, see Table 1 and Fig. 1). Subtype Ig with C6 was found in a landrace from Pakistan (the Dir group according to Ochiai et al. 1994; see Table 1). In six accessions of ssp. *viridis*, one from Japan had type Ie with C7, one from China had type If with C7, one from China had type Ia with C2AC4, and two from Pakistan and one from Turkey had type Ig with C6 (Table 1).

Sequencing polymorphism in types II and III and geographical distribution of their variants

We found four subtypes in type II and designated as subtypes IIa–d whereas type III is monomorphic despite using six accessions from geographically remote regions. Figure 6 shows the complicated relationships among type II subtypes and type III. They are different from one another in 8–22 nucleotides in subrepeats. Polymorphism is apparent not only in subrepeats, but also in the regions between subrepeats and C repeats (positions from 593 to 656 bp, Fig. 5). Sequences of this region of subtype IId and type III are identical with those of type I. However, subtype IIa has four or five nucleotide differences from the other subtypes, whereas subtypes IIb and IIc, respectively, have two and one such nucleotide differences. Even within subtype IIa, a nucleotide polymorphism was discovered between subrepeats and C repeat. At position 631, one of three subtype IIa accessions, accession No. 6, has A, whereas all other landraces have G (see Fig. 5). All of these types have a C7 sequence in the C repeat (Fig. 5). Sequences including subrepeats and C repeats of representative accessions were also registered in the DDBJ (DDBJ No.AB247511-247516).

One accession from the Nansei Islands and two accessions from Taiwan were of type IIa. One from the Philippines was IIb, one from China type IIc, and one from Novosibirsk were IId. Subtype IIa is observed intensively in the Nansei Islands of Japan and Taiwan (Table 1, Fig. 1).

subrepeat 1 | subrepeat 2 | subrepeat 3 | C repeat

Subrepeat 1

Position	33	138	141	143	144	151	169	182	190	191	215	230	232
(Position in subrepeat)	(33)	(138)	(141)	(143)	(144)	(151)	(169)	(182)	(190)	(191)	(215)	(230)	(232)
Type IIa	G	A	G	T	G	T	A	T	T	A	T	G	C
Type IIb	G	G	C	C	T	T	A	T	C	G	T	G	A
Type IIc	A	G	G	G	G	T	A	A	C	G	C	G	C
Type IId	G	A	G	T	G	A	G	T	C	G	C	G	C
Type III	A	G	G	G	G	T	A	T	C	G	C	A	C

Subrepeat 2

Position	235	266	267	279	298	308	329	341	356	377	378	385	436	454	457
(Position in subrepeat)	(1)	(32)	(33)	(45)	(64)	(74)	(95)	(107)	(122)	(143)	(144)	(151)	(202)	(220)	(233)
Type IIa	T	C	G	C	T	C	C	C	A	G	G	T	T	G	T
Type IIb	C	C	G	T	C	C	C	T	G	C	T	T	T	G	C
Type IIc	C	T	A	T	C	C	T	T	G	G	G	T	G	G	C
Type IId	C	C	G	T	C	C	C	T	G	T	G	A	T	G	C
Type III	C	C	A	T	C	T	C	T	G	G	G	T	T	T	C

Subrepeat 3 | | C repeat

Position	480	501	515	520	547	563	575	628	629	630	631	640	641	645	667-
(Position in subrepeat)	(12)	(33)	(47)	(52)	(79)	(95)	(107)								
Type IIa	G	G	G	G	A	T	G	C	C	A	A/G*	C	G	T	C7
Type IIb	G	G	A	A	T	C	T	G	G	G	A	T	T	T	C7
Type IIc	A	A	A	G	A	C	T	G	G	G	A	C	C	C	C7
Type IId	G	G	A	G	A	C	T	G	G	G	A	C	C	T	C7
Type III	G	A	A	G	A	C	T	G	G	G	A	C	C	T	C7

Fig. 5 Alignments of the IGS subrepeats between subtypes IIa–d and type III. Each of subrepeats 1 and 2 is 234-bp long and subrepeat 3 is 125 bp. Nucleotide substitutions are also observed in the region from 593 to 656 bp which are not included in subrepeats or C repeats whereas the sequences of this region of subtype IId and type III are identical to those of type I accessions. *Asterisk* indicates one accession of type IIa (accession no. 6) has G at this position whereas all other accessions have A at this position. Length of the sequences from the start of subrepeat 1 to the end of C repeats of all of the variants is 673 bp. *Numbers in parentheses* indicate the position of sequence of each subrepeat

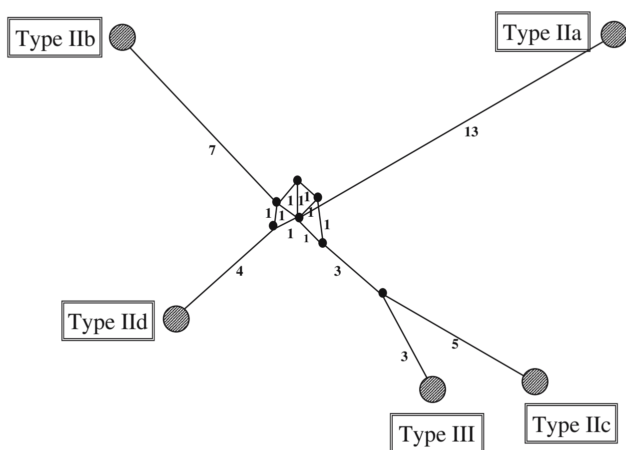

Fig. 6 A network tree of subtypes of types II and III. *Circles and dots* indicate subtypes and nodes, respectively. *Numbers under branches* indicate mutation steps

Relationships between subrepeats

Figure 7 shows that a network tree was drawn based on sequences of subrepeats. Subrepeat 1 of subtypes Ia and that of subtype Id and Ig are mutually identical, but they differ from that of other types in more than one nucleotide. Subrepeats differed in 1–4 nucleotides within type I, but are much more divergent within type II, as depicted in Figs. 3, 4, 5 and 6. Demonstrably, no subtypes of types II and III originated from a simple duplication of the type I subrepeat.

Within type II, subtype IIa differs greatly from other type-II subtypes, as shown in Figs. 5 and 6; subrepeats 1 and 2 of this subtype are also very different from others in at least four and six nucleotides. Two subrepeats also differ from each other in ten nucleotides. Subrepeats 1 and 2 in subtype IIb are divergent from others, but these two repeats are on the same branch of the network tree (Fig. 7). Subrepeat 1 of subtype IIc and type I subrepeats are different in 1–4 nucleotides, and subrepeats 2 are different from them by at least three nucleotides. Subrepeat 1 of subtype IIc is different from any of type I subrepeats 1 in 1–4 nucleotides, whereas its subrepeat 2 is different from them in at least three nucleotides. Subtype IId is also divergent from others in at least two nucleotides, but its subrepeats 1 and 2 are mutually similar. Type III subrepeats 1 and 2 are mutually similar; they are divergent from Type I subrepeats in 1–4 nucleotides. Types IIb–d and III could have originated by duplication of subrepeats from type I followed by differentiation of two subrepeats, but it is also possible to presume that type I arose through deletion of subrepeats from these longer types.

Discussion

Utility of the rDNA IGS sequence in detection of intraspecific polymorphisms

As ribosomal DNA (the 25 S-5.8 S–17 S rRNA gene) was mapped on a single locus using FISH in foxtail millet (Benabdelmouna et al. 2001) but it consists of tandemly arrayed sequences from several hundreds to thousands of copies (Rogers and Bendich 1987), it would be logically possible to presume some different sequences admixed in a tandem array. However, the sequences obtained by direct sequencing were not ambiguous, which indicated that one homologous sequence dominates in each type even though there might be paralogous copies in a very low frequency *in S. italica*.

Many studies of rDNA spacer length polymorphisms have been carried out to detect intraspecific polymorphism (e.g., Saghai Maroof et al. 1984; Sano and Sano 1990) and infer phylogenetic relationships. The present

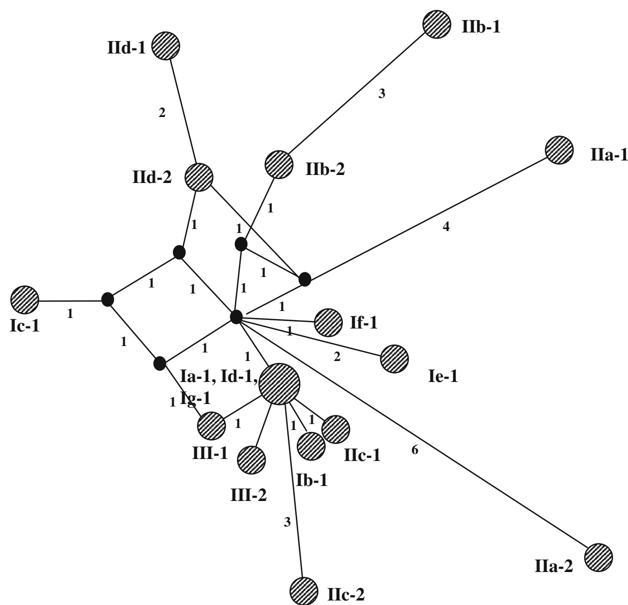

Fig. 7 A network tree of subrepeats. *Circles and black dots* indicate subrepeats and nodes, respectively. *Numbers under branches* indicate mutation steps

work indicates that not only length, but also the sequences of subrepeats and their flanking regions in the rDNA IGS are highly variable. To date, few studies have examined the IGS sequence polymorphism (Chou et al. 1999). Nevertheless, the present study demonstrates that these sequences are useful for detecting polymorphism between landraces or individuals, and therefore for inferences of intraspecific evolution.

Genetic differentiation within type I and multiple origins of foxtail millet in the Temperate Zone

This study was carried out to address the question whether foxtail millet was domesticated monophyletically or polyphyletically. We specifically examine type I landraces, which are broadly distributed throughout the Temperate Zone. In particular, we devote attention to the morphologically primitive landrace group in Afghanistan and northwestern Pakistan, which might be either a relict of an old cultivated form or an independent domesticate from other landraces.

This study revealed seven different sequence subtypes for type I. Subtypes Ia and Ib are distributed broadly: they differ in only one nucleotide. Type Ic is distributed in China, Korea, and Japan. This distribution pattern is similar to those of type A determined by intraspecific hybrid pollen sterility (Kawase and Sakamoto 1987), esterase isozymes (Kawase and Sakamoto 1984), prolamine alleles (Nakayama et al. 1999), and RFLP markers (Fukunaga et al. 2002), which also indicate that landraces of this area have close genetic relationships. This subtype differs from other subtypes of type I in subrepeat sequences and C repeats (Figs. 3, 4). These differences are unexplainable by a simple mutation step from other subtypes such as subtypes Ia, Ib, Id, If, and Ig. Subtype Ie, found in a ssp. *viridis* accession from Japan, is different from subtype Ic in a single nucleotide substitution both in subrepeats and in C repeat. Considering the network tree in Fig. 4, it is most reasonable to conclude that the differentiation between subtypes Ia and Ic occurred before domestication and that subtype Ic landraces were domesticated independently from other subtypes. As archaeological studies have indicated (Li and Wu 1996), genetic analyses also support that East Asia is a center of origin of foxtail millet.

Type Id differs from other subtypes of type I in subrepeat sequences, lacking 20 bp in subrepeat 3, and a specific nucleotide in C repeat (Figs. 3, 4). The landraces with subtype Id are found in Afghanistan (Sakamoto 1987) and northwestern Pakistan (the Chitral group, according to Ochiai et al. 1994, see Table 1), and are characterized by primitive morphology such as many tillers with small panicles. Those authors concluded that distribution of the primitive morphology should represent an old relict of this foxtail millet providing evidence of monophyletic origin of this cereal in this region. However, a 20-bp deletion of subrepeat 3 and multiple mutation steps that differ from other type I subtypes do not support that this type is an ancestral form of other landraces, as suggested by Sakamoto (1987), Ochiai et al. (1994), and Ochiai (1996). Rather, those facts support that these landraces were domesticated independently, as asserted by Li et al. (1995b). In the RFLP analysis of nuclear DNA (Fukunaga et al. 2002), Afghan landraces were grouped into a single cluster and differentiated from other Central Asian landraces. It is also congruent with a clear genetic differentiation pattern observed in this work and supports the probable independent origin of these morphologically primitive landraces. Although only limited accessions of ssp. *viridis* were used in this work and subtype Id was not found among them, a wild ancestor with this subtype might be found through further analyses.

Another interesting point is that type Ig with C6 was found in one Pakistan landrace belonging to the Dir group (Ochiai et al. 1994) and in three ssp. *viridis* accessions. This landrace might have originated independently from other landraces, but other possibilities exist such as introgression of the wild subspecies to ssp. *italica*. Further analyses should be made to prove which is more likely.

Genetic differentiation in types II and III and origin of foxtail millet in subtropical and tropical zones

In this work, we specifically examined the sequence polymorphism in type I because type I is distributed broadly in the temperate zone, which might include geographical origin(s) of this millet, whereas types II and III are distributed mainly in subtropical and tropical regions. In spite of the small number of type II accessions that were investigated, this type is highly polymorphic, whereas type III is monomorphic. We classified types II and III as different types in the previous papers because they have the same length, but differ only in their absence or presence of a *Bam*HI recognition site in subrepeat 2 (Fukunaga et al. 1997, 2005). However, they proved to have the same length, differ in a *Bam*HI recognition site in subrepeat 2 and several substitutions (Fukunaga et al. 2005). The present study showed that type II is highly polymorphic at the sequence level and type III is paraphyletic to subtypes of type II. Therefore, type III should be treated as subtype of type II (Table 1, Fig. 5) and hereafter we designate type III as subtype IIe.

Figures 5 and 6 show that five subtypes of type II (including subtype IIe) and differ greatly from one another. In Fig. 7, we also tried to clarify evolutionary relationships among subrepeats. These analyses indicate that relationships between RFLP variants of rDNA are much more complicated than we expected in the previous analysis (Fukunaga et al. 1997) because many nucleotide substitutions were detected among subrepeats. Many different events might have been involved in differentiation of the rDNA gene. We have analyzed sequence polymorphism by tentatively dividing the subrepeats as subrepeats 1 and 2. Moreover, we should consider the possibility of much more complicated evolution of rDNA in which recombination could have occurred in different ways because rDNA length polymorphism arises in complicated ways, e.g., by gene conversion and unequal crossover. We also analyzed the relationships of subrepeats by dividing a subrepeat further into two parts because subrepeats might have arisen in duplication of shorter subrepeats (ca. 130 bp) followed by deletion, as suggested in our previous paper (Fukunaga et al. 2005). Thereby, we obtained similar results and arrived at the same conclusion as that mentioned previously (precise data not shown).

Among subtypes of type II, it is difficult to attribute the origins of type IIa landraces of the Nansei Islands of Japan, and Taiwan, to a simple duplication of known subrepeats because both subrepeats 1 and 2 differ markedly from other subrepeats. Even these two subrepeats differ greatly from one another. Most likely, these landraces originated independently. We have identified some markers that are specific to this region (Jusuf and Pernes 1985; Fukunaga et al. 2002). These results indicate that landraces in this region are genetically distinct from those from other regions, although *GBSS 1* gene types (Kawase et al. 2005), positive phenol color reaction (Kawase and Sakamoto 1982) and a prolamin 2 allele (Nakayama et al. 1999) are common with some landraces in continental Asia. The origin of tropical landraces, including Taiwan and the Philippines, remains unresolved. We require more analyses of Asian landraces because foxtail millet is an important and old staple-food crop with significance in Eurasian history, particularly in Asian history.

Acknowledgements This work has been partially funded by a Research Institute For Humanity and Nature (RIHN) project 2-5PR(Project Leader Prof. Y.I.Sato). KF also thank Prof. Y. Yasuda, IRCJS, for his encouragement and supports.

References

Benabdelmouna A, Abirached-Darmency M, Darmency H (2001) Phylogenetic and genomic relationships in *Setaria italica* and its close relatives based on the molecular diversity and chromosomal organization of 5S and 18S-5.8S-25S rDNA genes. Theor Appl Genet 103:668–677

Buckler IV ES, Holtsford TP (1996) Zea systematics: ribosomal ITS evidence. Mol Biol Evol 13:612–622

Chou CH, Chiang Y-H,Chiang T-Y (1999) Within- and between-individual length heterogeneity of rDNA-IGS in *Miscanthus sinensis* var. glaber (Poaceae): Phylogenetic analyses. Genome 42:1088–1093

De Wet JMJ, Oestry-Stidd LL, Cubero JI (1979) Origins and evolution of foxtail millet (*Setaria italica*), Journ d'Agric et de Bot 26:53–64

Fukunaga K, Domon E, Kawase M (1997) Ribosomal DNA variation in foxtail millet, *Setaria italica* (L.) P.Beauv. and a survey of variation from Europe and Asia. Theor Appl Genet 97:751–756

Fukunaga K, Wang ZM, Kato K, Kawase M (2002) Geographical variation of nuclear genome RFLPs and genetic differentiation in foxtail millet, *Setaria italica* (L.) P. Beauv. Genet Res Crop Evol 49:95–101

Fukunaga K, Ichitani K, Taura S, Sato M, Kawase M (2005) Ribosomal DNA intergenic spacer sequence in foxtail millet, *Setaria italica* (L.) P. Beauv. and its characterization and application to typing of foxtail millet landraces. Hereditas 142:38–44

Harlan JR (1975) Crops and Man. Soc Amer Agron Crop Sci Soc America, Madison, WI

Jusuf M, Pernes J (1985) Genetic variability of foxtail millet *(Setaria italica* (L.) P. Beauv): Electrophoretic study of five isoenzyme systems. Theor Appl Genet 71:385–391

Kawase M, Sakamoto S (1982) Geographical distribution and genetic analysis of phenol color reaction in foxtail millet, *Setaria italica* (L.) P. Beauv. Theor Appl Genet 63:117–119

Kawase M, Sakamoto S (1984) Variation, geographical distribution and genetic analysis of esterase isozymes in foxtail

millet, *Setaria italica* (L.) P. Beauv. Theor Appl Genet 67:529–533

Kawase M, Sakamoto S (1987) Geographical distribution of landrace groups classified intraspecific pollen sterility in foxtail millet, *Setaria italica* (L.) P.Beauv. Jpn J Breed 37:1–9

Kawase M, Fukunaga K, Kato K (2005) Diverse origins of waxy foxtail millet crops in East and Southeast Asia mediated by multiple transposable element insertions. Mol Genet Genomics 274:131–140

Kihara H, Kishimoto E (1942) Bastarde zwischen *Setaria italica* und *S. viridis* (in Japanese with German summary). Bot Mag 20:63–67

Küster H (1984) Neolithic plant remains from Eberdingenhochdorf, southern Germany. In: Van Zeist WV, Casparoe WA (eds) Plants and ancient man (studies in palaeoethnobotany). AA Bakame, Rotterdam, Boston, pp 307–311

Le Thierry d'Ennequin M, Panaud O, Toupance B, Sarr A (2000) Assessment of genetic relationships between *Setaria italica* and its wild relatives *S. viridis* using AFLP marker. Theor Appl Genet 100:1061–1066

Li Y, Wu SZ (1996) Traditional maintenance and multiplication of foxtail millet (*Setaria italica* (L.) P. Beauv) landraces in China. Euphytica 87:33–38

Li HW, Li CH, Pao WK (1945) Cytological and genetic studies of interspecific cross of the cultivated foxtail millet, *Setaria italica* P. Beauv. and the green foxtail millet, S. *viridis* L. J Amer Soc Agron 37:32–54

Li Y, Cao YS, Wu SZ, Zhang XZ (1995a) A diversity analysis of foxtail millet (*Setaria italica* (L.) P.Beauv.) landraces of Chinese origin. Genet Res Crop Evol 45:279–285

Li Y, Wu SZ, Cao YS (1995b) Cluster analysis of an international collection of foxtail millet (*Setaria italica* (L.) P. Beauv). Genet Res Crop Evol 45:279–285

Li Y, Jia J, Wang W, Wu S (1998) Intraspecific and interspecific variation in *Setaria* revealed by RAPD analysis. Genet Res Crop Evol 45:279–285

Lisitsina A (1976) Main types of ancient farming on the Caucasus on the basis of palaeo-ethnobotanical research. Ber Deut Bot Ges 91:47–57

Nakayama H, Namai H, Okuno K (1999) Geographical variation of the alleles at the two prolamin loci, *Pro1* and *Pro2*, in foxtail millet, *Setaria italica* (L.) P. Beauv. Genes Genet Syst 74:293–297

Ochiai Y (1996) Variation in tillering and geographical distribution of foxtail millet (*Setaria italica* P. Beauv.) Breed Sci 46:143–146

Ochiai Y, Kawase M, Sakamoto S (1994) Variation and distribution of foxtail millet (*Setaria italica* P. Beauv.) in the mountainous areas of northern Pakistan. Breed Sci 44:413–418

Rogers BO, Bendich AJ (1987) Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. Plant Mol Biol 9:509–520

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

Sallares R, Brown TA (2004) Phylogenetic analysis of complete 5 external transcribed spacers of the 18S ribosomal RNA genes of diploid *Aegilops* and related species (Triticeae, Poaceae). Genet Res Crop Evol 51:701–712

Sakamoto S (1987) Origin and dispersal of common millet and foxtail millet. Japan Agric Res Quart 21:84–89

Sano Y, Sano R (1990) Variation of the intergenic spacer region of ribosomal DNA in cultivated and wild rice species. Genome 33:209–218

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

Vavilov NI (1926) Studies on the origin of cultivated plants. Inst Appl Bot Plant Breed, Leningrad

Wang RL, Wendell J,Dekker J (1995) Weedy adaptation in *Setaria* spp.: I. Isozyme analysis of the genetic diversity and population genetic structure in *S. viridis*. Am J Bot 82:308–317