Theor Appl Genet (2013) 126:1885–1896
DOI 10.1007/s00122-013-2103-z

ORIGINAL PAPER

Cloning and characterization of chromosomal markers in alfalfa (*Medicago sativa* L.)

Feng Yu · Yunting Lei · Yuan Li · Quanwen Dou · Haiqing Wang · Zhiguo Chen

Received: 26 November 2012 / Accepted: 10 April 2013 / Published online: 1 May 2013

Abstract Eleven tandemly repetitive sequences were identified from a Cot-1 library by FISH and sequence analysis of alfalfa (*Medicago sativa*). Five repetitive sequences (*Ms*CR-1, *Ms*CR-2, *Ms*CR-3, *Ms*CR-4, and *Ms*CR-5) were centromeric or pericentromeric, of which three were satellite DNAs and two were minisatellite DNAs. Monomers of 144, 148, and 168 bp were identified in *Ms*CR-1, *Ms*CR-2, and *Ms*CR-3, respectively, while 15 and 39 bp monomers were identified in *Ms*CR-4 and *Ms*CR-5, respectively. Three repetitive sequences were characterized as subtelomeric; one repetitive sequence, *Ms*TR-1, had a 184 bp monomer, and two repetitive sequences had fragments of 204 and 327 bp. Sequence analysis revealed homology (70–80 %) between *Ms*TR-1 and a highly repeated sequence (C300) isolated from *M.* ssp. *caerulea.* Three identified repetitive sequences produced hybridization signals at multiple sites in a few of the chromosomes; one repetitive sequence was identified as the E180 satellite DNA previously isolated from *M. sativa*, while the other 163 and 227 bp fragments had distinct sequences. Physical mapping of the repetitive sequences with double-target FISH revealed different patterns. Thus, nine novel tandemly repetitive sequences that can be adopted as distinct chromosome markers in alfalfa were identified in this study. Furthermore, the chromosome distribution of each sequence was well described. Though significant chromosome variations were detected within and between cultivars, a molecular karyotype of alfalfa was suggested with the chromosome markers we identified. Therefore, these novel chromosome markers will still be a powerful tool for genome composition analysis, phylogenetic studies, and breeding applications.

Communicated by B. Friebe.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-013-2103-z) contains supplementary material, which is available to authorized users.

F. Yu · Y. Lei · Y. Li · Q. Dou (✉) · H. Wang · Z. Chen
Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Plateau Institute of Biology, Chinese Academy of Sciences, Xining 810001, China
e-mail: douqw@nwipb.cas.cn

F. Yu · Y. Li
Graduate University of Chinese Academy of Sciences, Beijing 100049, China

Introduction

Cultivated alfalfa (*Medicago sativa* L.) is one of the most important perennial forage crops all over the world. Its well-recognized symbiosis with *Rhizobium* bacteria, an effective source of biological nitrogen fixation, provides an important source of protein for livestock. Cytogenetic research on alfalfa and closely related species has lagged behind other crops because its chromosomes are very small, ranging from 2–3 μm in length in root tip cells, morphologically very similar, numerous ($2n = 4x = 32$), and autotetraploid in nature with four nearly identical genomes (Bauchan and Hossain 1997). Initial identification of alfalfa chromosomes at the pachytene stage was conducted by Gillies and Ho (Gillies 1970; Ho and Kasha 1972). An advance was made by applying the technique of C-banding, which was used to characterize the somatic chromosomes of *M. sativa* CV CADL (cultivated alfalfa at the diploid level) (Masoud et al. 1991) and tetraploid

alfalfa (Falistocco et al. 1995; Bauchan and Hossain 2001). The C-banding pattern revealed that most of the chromosomes have centromeric bands, interstitial bands and a terminal band on the short arm, and a few chromosomes have interstitial bands on their long arms (Bauchan and Hossain 2001). Furthermore, fluorescence in situ hybridization (FISH) was applied to investigate the distribution of rDNA genes (Falistocco 2000). Although karyotyping in alfalfa based on conventional technology and FISH with rRNA genes has been performed, its dependence on technique or inadequate landmarks leads to a low degree of precision and limited utility.

FISH-based karyotyping, by physical mapping of repetitive sequences directly on metaphase chromosomes, allows identification of each chromosome, the study of chromosome aberrations and chromosome evolution, and helps with genome sequence assembly and breeding (Schmidt and Heslop-Harrison 1998; Jiang and Gill 2006). Repetitive sequences are indispensable as chromosome markers for molecular karyotyping. However, to date, besides commonly used chromosome markers such as rDNAs, only a limited number of chromosome markers, such as the repetitive sequence E180 (Xia and Erickson 1993; Rosato et al. 2012), are available in alfalfa. Therefore, a collection of chromosome markers is needed for a complete karyotyping system in alfalfa.

Repetitive sequences, tandem repeats, and dispersed repeats can be used for karyotyping by FISH. Repetitive sequences can be isolated by several strategies. Highly repeated DNA sequences are conventionally obtained by genomic DNA enzyme digestion, Southern hybridization, and sequence analysis (Xia and Erickson 1993). They can also be identified using bioinformatics tools when sequence information is available. Cot-1 DNA is enriched for highly and moderately repetitive DNA sequences. It has been used as "blocking" DNA to compete with labeled probe for repetitive sequences in FISH, and labeled as probes to localize heterochromatin in chromosomes (Zwick et al. 1997). Furthermore, repetitive sequences including microsatellite, minisatellite and satellite DNAs have been isolated and characterized in a Cot-1 library of *Dendrobium* (Begum et al. 2009) and similarly in sugar beet (*Beta vulgaris*) (Zakrzewski et al. 2010) and peanut (*Arachis hypogaea*) (Zhang et al. 2012).

In this study, we first constructed an alfalfa Cot-1 library and then conducted FISH screening with labeled clones on mitotic chromosomes. The clones showing strong hybridization signals were investigated and several distinct chromosome markers were revealed. They will be very useful in alfalfa chromosome identification. The applications of these chromosome markers in FISH-based karyotyping, genome organization, and comparative genomics are discussed.

Materials and methods

Plant materials

Four alfalfa cultivars and one *M.* ssp. *falcata* accession (PI 234815) were used in this study. These cultivars were 'Algonquin', 'Golden Empress', 'Aohan', and 'Zhongmu No. 1'. 'Aohan' was used to construct a Cot-1 library and cultivar 'Zhongmu No. 1' to prepare slides with mitotic chromosomes. Seeds were germinated on moist filter paper in petri dishes at room temperature. Root tips were excised 2–3 days after germination for chromosome spreading, and germinated plants were planted into pots and grown in an artificial climate chamber.

Genomic DNA extraction and Cot-1 DNA preparation

Total genomic DNA was extracted from fresh young leaves using a modified protocol originally developed by Allen et al. (2006). Cot-1 DNA preparation was performed following Zwick et al. (1997).

Cot-1 DNA library construction

Cot-1 DNA fragment 'A' bases were added to the Cot-1 DNA fragments by a PCR extension reaction. The reaction volumes were 10 μL, with 2.5 mmol/L dNTPs (Takara), 1 U of Taq DNA polymerase (Takara), 1 μL of the corresponding 10× PCR buffer, and 3 μg Cot-1 DNA fragment solution as a template. Reactions were directly conducted at 72 °C for 30 min in an ABI Veriti machine (Applied Biosystems). Subsequently, purified Cot-1 DNA fragments were ligated into the pGEM-T easy vector (Promega). The transformed *E. coli* cells (DH5α) were identified by a blue/white screen following the manufacturer's instructions (Promega). The white recombinant colonies were stored at −80 °C in LB medium containing 15 % v/v glycerol.

Probe preparation

The 5S rDNA was amplified by polymerase chain reaction (PCR) using genomic DNA of alfalfa as described by Fukui et al. (1994). pWrrn, which included fragments of wheat 18S–26S rDNA, was provided by Professor Tsujimoto (Tottori University, Japan). Inserts from candidate clones were amplified by PCR using T7 and SP6 primers. Purified PCR products were labeled with tetramethyl-rhodamine-5-dUTP (red) or fluorescein-12-dUTP (green) (Roche Diagnostics) by a random primer labeling method (Prime-it Fluor Fluorescence Labeling Kit, Agilent Technologies). pWrrn was labeled with tetramethyl-rhodamine-5-dUTP (red) by a nick-translation method.

Chromosome preparation

Root tips with a length of 1–2 cm were collected, pretreated in ice-cold water at 0 °C for 20–24 h, fixed in ethanol:glacial acetic acid (3:1, v/v) for 24 h at room temperature and then stored at 4 °C in the refrigerator. Each root tip was squashed in a drop of 45 % acetic acid.

FISH and microphotometry

FISH experiments were carried out as described by Dou et al. (2009). Chromosomes were counterstained with 4′,6-diamidino-2-phenylindole (DAPI). Images were captured with a cooled CCD camera (Photometrics CoolSNAP) under a fluorescence microscope (Leica) and processed with the meta imaging system (Universal Imaging Corporation). Finally, images were adjusted with Adobe Photoshop 6.0 for contrast and background optimization.

DNA sequencing and data analysis

Cot-1 DNA cloned products were sequenced by Sangon Biotech Co., Ltd. (Shanghai, China). Sequence data were analyzed using the DNAman software package (Lynnon Biosoft, Quebec, Canada). Sequences were queried against the NCBI nucleotide database using BLASTN to search for similar sequences.

Southern hybridization

Genomic DNA was isolated from young leaf tissue of cultivar 'Zhongmu No. 1'. DNA samples digested with restriction enzymes were electrophoresed in a 0.8 % agarose gel and transferred to a Hybond N^+ (Amersham) membrane. Probe labeling and hybridization was conducted using a Ready-To-Go kit (Dig High Prime DNA Labeling and Detection Starter Kit II, Roche). After hybridization, the membranes were washed at different stringencies by controlling the concentration of SSC and then the membranes were exposed to X-ray films.

Results

Isolation of Cot-1 DNA and construction of Cot-1 DNA library

Clean RNA-free genomic DNA of cultivar 'Aohan' was extracted and 60 μg of this was used to obtain Cot-1 DNA that contained fragments less than 1 kb (Fig. 1a). To check the quality of the Cot-1 DNA, a probe labeled with purified Cot-1 DNA was hybridized to alfalfa metaphase chromosomes by FISH. The hybridization pattern showed that most of the chromosomes had high-intensity signals in the pericentromeric and subtelomeric regions on their short arms (Fig. 1b). Comparison with the results reported by Bauchan and Hossain (2001) indicated that the high-intensity distribution of Cot-1 DNA was similar to the constitutive heterochromatic band patterns produced by C-banding. Usually, the heterochromatic bands in a chromosome reflect a large quantity of tandem repetitive sequences. Thus, it indicated that the Cot-1 DNA isolated in this study included a large proportion of tandem repeats. From this DNA, a Cot-1 DNA library with 115 clones was constructed.

Screening and characterization of Cot-1 clones using FISH

Cot-1 DNA is enriched for highly and moderately repetitive DNA sequences. However, only highly tandem repeated DNA sequences are ideal chromosome markers that always produce discernible "block" or "dot" signals on chromosomes when detected by FISH. To develop alfalfa chromosome markers, the Cot-1 clones were first characterized by FISH, and only those clones exhibiting significant FISH signals were selected as candidate chromosome markers for further analysis. Forty-four clones were randomly picked from the positive clones in the Cot-1 DNA library. Insert fragments were amplified by PCR and labeled. Since the Cot-1 DNA was prepared from genomic DNA of cultivar 'Aohan', another cultivar ('Zhongmu No. 1') was used for FISH screening to obtain universal chromosome markers across cultivars. Subsequently, the insert fragment fluorescence probes were hybridized to 'Zhongmu No. 1' interphase or metaphase chromosomes. The results showed that 38 clones, 86.4 % of those tested, produced distinct "dot" or "block" hybridization signals on the chromosomes. To characterize them, we categorized the positive clones into three types according to their chromosome distribution patterns (Table 1).

Type 1 showed hybridization signals around the centromere and included 15 clones (34.1 %). All clones, except clones 29, 11, and 38, showed centromeric signals on nearly all of the alfalfa chromosomes. Clones 29, 11, and 38 showed hybridization signals on some of the chromosomes. Type 2 showed hybridization signals on the subtelomeric or telomeric regions; 11 clones (25.0 %) belonged to this type. All clones, except clone 18, showed subtelomeric signals on one arm on most of the chromosomes. Clone 18 produced hybridization signals on both arm ends on all 32 alfalfa chromosomes. Type 3 showed hybridization signals not only in the pericentromeric and subtelomeric regions, but also in the interstitial regions. Type 3 included 12 clones, which represented 27.3 % of the total clones.

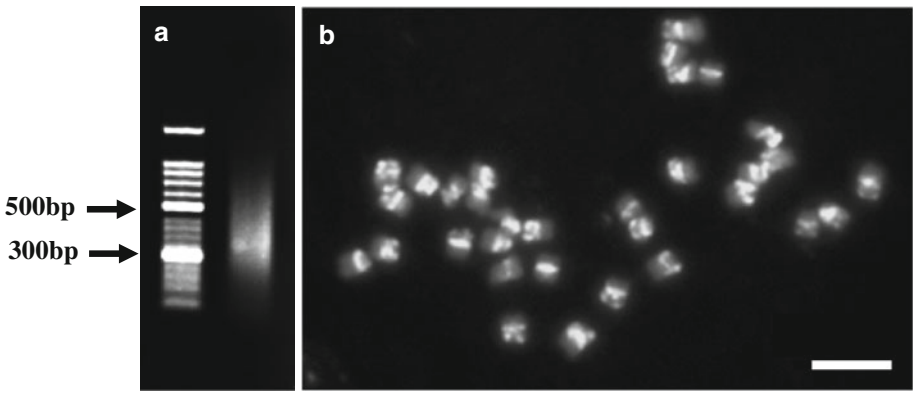

Fig. 1 **a** Electrophoresis pattern of Cot-1 DNA. M size markers, 1 Cot-1 DNA. **b** FISH pattern of mitotic chromosomes probed by Cot-1 DNA. *Scale bar* = 10 μm

Table 1 Characterization of clones from alfalfa (*M. sativa*) Cot-1 DNA library by FISH

FISH pattern	Number of clones (clone no.)	Percentage of the total
Type 1 (centromeric sites)	15 (3, 10, 11, 16, 19, 29, 38, 70, 76, 79, 81, 89, 91, 95, 123)	34.1
Type 2 (subtelomeric sites)	11 (14, 18, 39, 53, 65, 67, 74, 78, 80, 90, 94)	25.0
Type 3 (multiple sites)	12 (17, 26, 34, 36, 37, 54, 59, 68, 87, 102, 139, 161)	27.3
No distinct signals	6 (8, 25, 52, 72, 77, 88)	13.6

Sequencing and characterization of the selected clones

All 38 selected clones were sequenced, which showed that they contained inserts ranging from 46–470 bp in size. To characterize the sequences of the inserts, we performed homology searches against the nucleotide sequences in the NCBI database. Seventeen clones showed high homology to genomic sequences from *M. truncatula*, ten clones matched the previously reported tandem repetitive sequence E180 in *M. sativa* (Xia and Erickson 1993), seven clones had homologs to the tandem repetitive sequence C300 in *M. sativa* ssp. *caerulea* (Calderini et al. 1997), and the other three clones showed homology to a *M. sativa* retrotransposon polyprotein gene sequence, a *M. truncatula* chloroplast DNA, and a telomere-like sequence (Supplementary Table 2, see Supplementary Fig. 1).

The search results showed that a few clones shared similarity with the same sequence in the database. Sequence comparison revealed that some clones had high homology to each other. This observation indicated that many of the Cot-1 clones contained partial or full sequences from the same repetitive sequence family. Thus, it was necessary to perform an assembly of the selected clones to generate contigs containing sequences belonging to the same repetitive sequence family. Subsequently, five contigs were established with a DNA sequence assembly program.

Contig assembly and monomer searching in the Type 1 clones

Three contigs were established in this group. Contig 1 was 271 bp and included nine clones (clones 3, 10, 16, 19, 76, 81, 89, 91, 95). Blast searching against the NCBI database showed that clones 3, 10, 16, 19, 76, 81, 91, and 95 had high homology to AC130807.30, while only clone 89 had high homology to AC166314.2 (Supplementary Table 2, see Supplementary Fig. 1). This confirmed that the fragments represented by these clones came from one integrated sequence. To determine the full length of this monomer, a 949 bp fragment was amplified from the genomic DNA by PCR with the oligonucleotide primers 5-gaatcaagttatcatcgtggaa-3 and 5-cgaacggtcacaaggtgt-3. The sequence analysis showed that the amplified fragment included six integrated monomer units and part of a seventh (Supplementary Fig. 2). We named this 144 bp repetitive monomer *Ms*CR-1.

Contig 2 included two clones, 70 and 79, which showed high homology to AC151462.30. Sequence assembly generated a sequence of 435 bp. Sequence analysis revealed that it contained one integrated monomer of 298 bp that was partially repeated. To confirm the full length of this monomer, a 988 bp fragment was amplified by PCR with the oligonucleotide primers 5-atttaatcgaatttcaaacc-3 and 5-gagattttagcctttatggt-3. Sequence analysis identified five shorter monomers of around 148 bp in this amplification. We named this 148 bp repetitive monomer *Ms*CR-2 (Supplementary Fig. 3).

Clones 29 and 123 showed high homology to AC157371 and had a high homology of 92.3 % to each other. This implied that they were derived from the same repeat family. A contig of 246 bp was generated from these two repeats. Monomer searching revealed that the contig contained one integrated monomer of 168 bp and a partial repeat. To confirm the full length of this monomer, a 660 bp fragment was amplified by PCR with the

oligonucleotide primers 5-tacttccccataataaagtt-3 and 5-ggtccataatcatcatacaa-3, which were designed according to the clone 29 sequence. The sequence analysis showed that nearly four fully integrated monomers were included in this fragment (Supplementary Fig. 4). Comparison with the available repetitive sequences *Mt*R1, *Mt*R2, and *Mt*R3 isolated from *M. truncatula* (Kulikova et al. 2004) showed that the sequence had 86.8 % homology to *Mt*R3. *Mt*R3 was characterized as a repeat located in the functional portion of all eight centromeres of *M. truncatula* (Kulikova et al. 2004). Since it was the first to be characterized in *M. sativa*, it was named *Ms*CR-3.

Clones 11 and 38 were two minisatellite DNAs. Sequence analysis revealed that clone 11 contained eight 15 bp monomers and clone 38 contained three 39 bp monomers. Their repeat arrays were organized in head-to-tail junctions. The alignments of each repetitive monomer are shown in Supplementary Figs. 5 and 6. The monomers in clones 11 and 38 were named *Ms*CR-4 and *Ms*CR-5, respectively.

Contig assembly and monomer searching in the Type 2 clones

One contig was established in this type. Contig 1 was 335 bp in size and included eight clones (clones 14, 39, 53, 67, 78, 80, 90, 94). Blast searches against the NCBI database showed that seven clones had 70–80 % homology to a tandem repetitive sequence, C300, in *M. sativa* ssp. *caerulea* (Calderini et al. 1997). Sequence analysis revealed that the contig contained one integrated monomer of 184 bp that was partially repeated. To confirm the full length of this monomer, a 1,334 bp fragment was amplified by PCR from the oligonucleotide primers 5-aaagggtcgaaaacgagtaa-3 and 5-attcagccaaaatggaggtt-3. The sequence analysis showed that nearly seven fully integrated 184 bp monomers were included in this fragment (Supplementary Fig. 7). We named the 184 bp repetitive monomer *Ms*TR-1.

Clones 65 and 74 had sizes of 204 and 327 bp (Supplementary Figs. 8, 9) and could not be grouped into any contigs. Sequence comparison showed that they shared very low similarity to each other and no integrated monomers were found in these sequences. Because their repetitive nature was inferred from the FISH patterns, clones 65 and 74 possibly carry partial fragments from a different repetitive family that contains long motifs.

Clone 18 produced FISH signals at both arm ends in all chromosomes. The clone sequence showed that it contained multiple repeats of a microsatellite TTTAGGG in a head-to-tail junction. Most plant species have this *Arabidopsis*-type microsatellite (TTTAGGG) array at the ends of each chromosome (Richards and Ausubel 1988). Thus, the sequence revealed in clone 18 is a typical telomere sequence.

Contig assembly and monomer searching in the Type 3 clones

Twelve clones were characterized as Type 3. From these, one contig of 417 bp was established, which involved ten clones (clones 17, 26, 34, 36, 37, 54, 59, 102, 139, 161). Blast searches against available sequence data revealed that the contig sequence had high homology to the published repeat E180, which was isolated in alfalfa (Xia and Erickson 1993). Sequence comparison between each clone and E180 showed homology ranging from 88.8 to 95.9 %. Comparison between different copies of E180 revealed that they were 87.8–92.6 % homologous to each other (Xia and Erickson 1993). This indicated that the generated alignment belonged to the E180 repeat family. The clones included in this alignment could be characterized as E180-like clones.

Clones 68 and 87 had sizes of 163 bp and 227 bp (Supplementary Figs. 10, 11). Blast searches against the database showed that clones 68 and 87 were homologous to different *M. truncatula* genomic sequences. Sequence comparison revealed low similarity to each other and no monomers were found in either sequence. This implied that clones 68 and 87 included partial fragments of two different repeat families that produce FISH signals at multiple sites on the alfalfa chromosomes.

Southern hybridization

Southern hybridization was conducted with all of the newly identified repeated sequences as probes for alfalfa genomic DNA digested with various restriction enzymes (Fig. 2). The results showed that all of the repeats, except clone 65, showed ladder-like hybridization patterns with the genomic

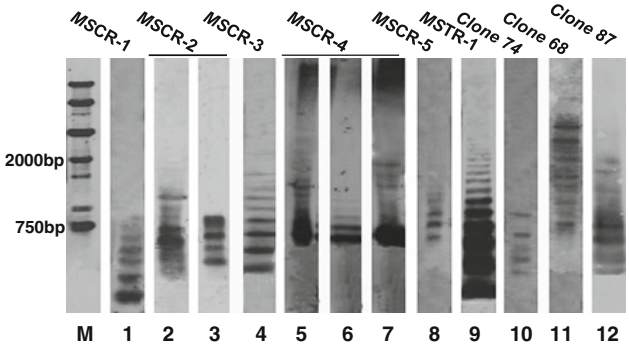

Fig. 2 Southern hybridization of the repetitive sequences to enzyme-digested genomic DNA of *M. sativa.* M size markers; *1* Taq I, *2* Nde I, *3* Mse I, *4* Hinf I, *5* Hinf I, *6* Alu I, *7* Hind III, *8* Hinf I, *9* Alu I, *10* Hinf I, *11* Taq I, *12* Taq I

DNA digested by one or more restriction enzymes. This indicated that these sequences were tandemly repeated in the alfalfa genome. The exceptional result of clone 65 was probably due to the limited restriction enzymes used in this study. A similar ladder-like result for clone 65 should be obtained using other restriction enzymes.

Proportional composition of the new repeats in Cot-1 DNA

FISH screening showed that 39.4 % of the selected clones were characterized as centromeric, 29.0 % were characterized as telomeric or subtelomeric, and 31.6 % showed centromeric telomeric, subtelomeric and interstitial bands simultaneously. To confirm the distribution of each sequence in the Cot-1 library, the sequence information of all available Cot-1 DNA clones (115 in total) was analyzed. The results showed that *Ms*CR-1 appeared at the highest frequency (44.3 %), followed by E180 at 15.7 % and *Ms*TR-1 at 10.4 %, while *Ms*CR-4, *Ms*CR-5, clone 65, and clone 74 appeared as the lowest frequency (0.8 % each) and the others had frequencies between and 6.9 %. In total, centromeric and telomeric or subtelomeric repeats occupied about 56.7 and 12.1 % of the Cot-1 DNA, while 20.0 % of the Cot-1 DNA comprised subtelomeric and interstitial bands simultaneously.

Physical mapping of the newly identified tandem repeats on mitotic chromosomes using FISH

Five novel tandemly repetitive sequences were characterized as Type 1 clones, which produced FISH signals around the centromere. The *Ms*CR-1 contig sequence included nine clones. FISH results showed that all nine clones shared the same hybridization pattern: high-density FISH signals around the centromere on 30 of the 32 chromosomes (Fig. 3a). FISH screening showed that the clones belonging to *Ms*CR-2, *Ms*CR-3, *Ms*CR-4, and *Ms*CR-5 each produced signals on a portion of the chromosomes. To determine whether the signals produced by different repeats overlapped, we conducted double-target FISH with different probe combinations. The FISH pattern detected by the combination of *Ms*CR-3 (clone 29) and *Ms*CR-2 (clone 70) showed that 6 chromosomes had only *Ms*CR-3 hybridization signals, 4 had only *Ms*CR-2, 18 had both, and 4 had no signals (Fig. 3b). *Ms*CR-3 (clone 29) and *Ms*CR-4 (clone 11) produced a different distribution pattern. They showed that 9 chromosomes had only *Ms*CR-3 hybridization signals, 5 had only *Ms*CR-4, 16 had both, and 2 had none (Fig. 3c). The FISH pattern detected by the combination of *Ms*CR-3 (clone 29) and *Ms*CR-5 (clone 38) showed that 10 chromosomes had only *Ms*CR-3 hybridization signals, 6 had only *Ms*CR-5, 14 had both, and 2 had none (Fig. 3d). These results showed that *Ms*CR-2, *Ms*CR-3, *Ms*CR-4, and *Ms*CR-5 were distributed on 22, 24, 20, and 21 chromosomes, respectively, and were co-localized on a few of the chromosomes.

The physical distribution of the three identified repeats in the Type 2 clones was also investigated by double-target FISH. *Ms*TR-1, clone 65 and clone 74 produced major signals on the subtelomeric regions of 24–26, 26, and 26 chromosomes, respectively. The FISH signals of the three repeats overlapped on most of the chromosomes. However, clone 65 produced more signals in the subtelomeric regions of three chromosomes than *Ms*TR-1 (Fig. 3e), and clone 74 produced more weak signals in the interstitial regions of four chromosomes than *Ms*TR-1 (Fig. 3f).

Three different repetitive sequences were identified among the Type 3 clones. Most of the clones belonged to E180, an AT-rich satellite DNA sequence (Xia and Erickson 1993). FISH localization of the E180 repeat in a few *Medicago* species showed stronger hybridization signals in the subtelomeric and interstitial loci than in the pericentromeric loci (Rosato et al. 2012). Our results showed that the E180 repeat produced hybridization signals in 30–32 chromosomes and was preferentially distributed in the interstitial loci. Over half of the chromosomes showed more than one E180 hybridization site. Double-target FISH revealed that the chromosome distribution patterns of clones 68 and 87 were significantly different from those of E180. Clones 68 and 87 produced more signals in the pericentromeric and subtelomeric regions than in the interstitial regions (Fig. 3g, h). Comparison of the FISH patterns between clones 68 and 87 showed that clone 68 produced stronger hybridization signals in the pericentromeric loci than in the subtelomeric loci, whereas clone 87 produced stronger hybridization signals in the subtelomeric loci than in the pericentromeric loci.

Distribution of each sequence

The 18S–26S rDNA and 5S rDNA are positioned in two and four pairs of chromosomes, respectively, in alfalfa (Falistocco 2000). In our study, E180 produced multiple hybridization signals in 30–32 of the chromosomes. Using 18S–26S rDNA and 5S rDNA as references, combined with the FISH pattern of E180 (Fig. 3k, l) and the chromosome arm ratio, the chromosomes of cultivar 'Zhongmu No. 1' could be characterized into 16 pairs (Fig. 5a, b). The FISH signals for each of the sequences identified in this study were determined (Figs. 3m–p, 4a–d). They are as follows, *Ms*CR-1: present on all chromosomes except one chromosome of pairs 5, 7, and 11 (Fig. 5c); *Ms*CR-2: 4, 5, 6, 8, 9, 15, and one chromosome of pairs 2, 10, and 14 (Fig. 5d); *Ms*CR-3: 3, 5, 6, 8, 9, 10, 13, and 14 (Fig. 5e); *Ms*CR-4: 2, 3, 4, 5, 8, 13, 14, and one chromosome of pairs

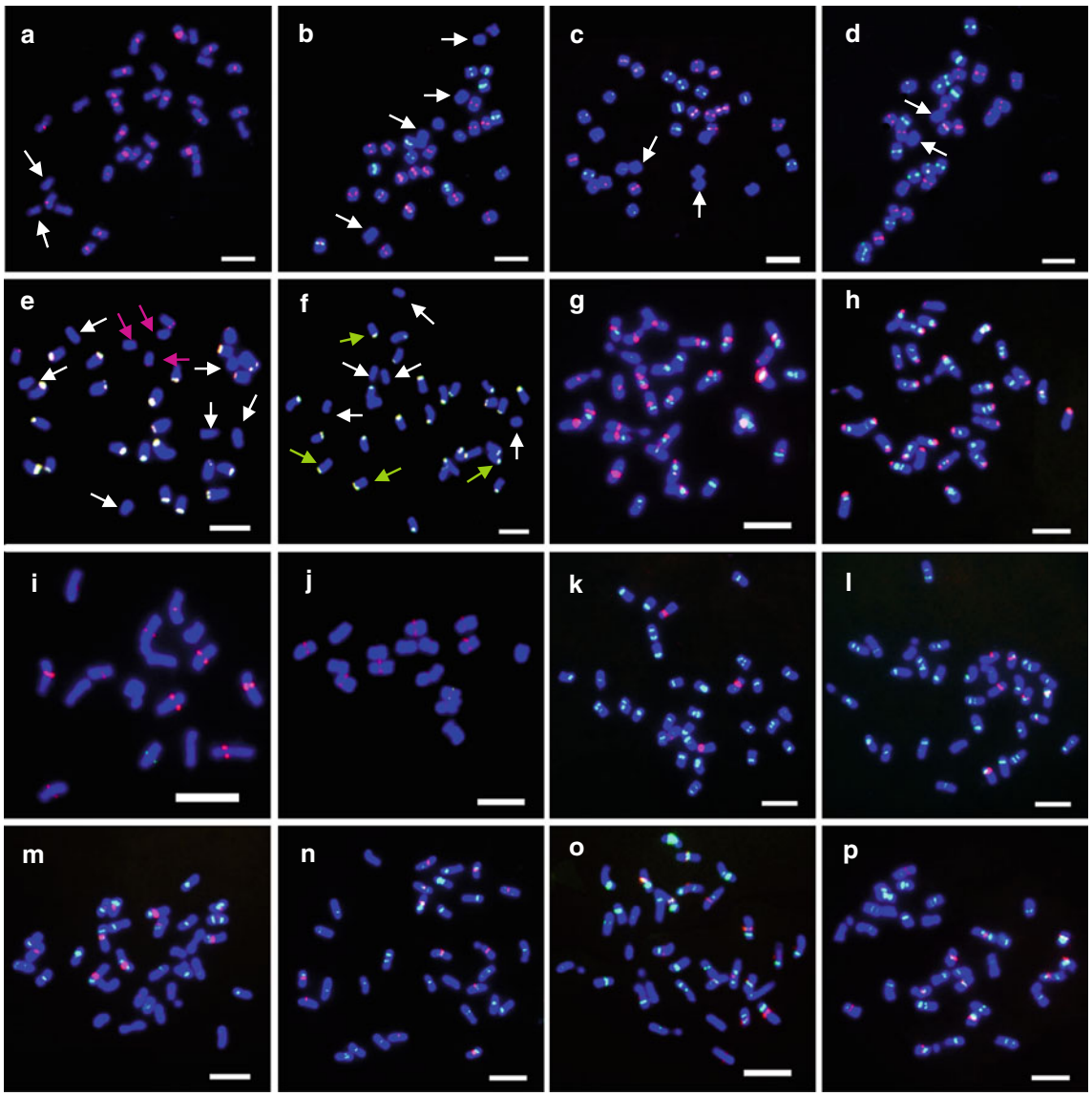

Fig. 3 Physical mapping of identified repeats on mitotic chromosomes of alfalfa (*M. sativa*) cultivar 'Zhongmu No. 1' using FISH. **a** Probed by *Ms*CR-1 (clone 76). **b** Probed by *Ms*CR-3 (clone 29) (*green*) and *Ms*CR-2 (clone 70) (*red*). **c** Probed by *Ms*CR-3 (clone 29) (*green*) and *Ms*CR-4 (clone 11) (*red*). **d** Probed by *Ms*CR-3 (clone 29) (*green*) and *Ms*CR-5 (clone 38) (*red*). **e** Probed by *Ms*TR-1 (clone 78) (*green*) and clone 65 (*red*). **f** Probed by *Ms*TR-1 (clone 78) (*green*) and clone 74 (*red*). **g** Probed by E180 (clone 34) (*green*) and clone 68 (*red*). **h** Probed by E180 (clone 34) (*green*) and clone 87 (*red*). *Red*, *yellow*, and *white arrows* indicate chromosomes carrying only *red*, interstitial *red*, and no signals, respectively. **i–j** FISH signals in *M. falcata*. **i** Probed by *Ms*CR-3 (clone 29) (*green*) and *Ms*CR-2 (clone 70) (*red*). **j** Probed by *Ms*CR-3 (clone 29) (*green*) and *Ms*CR-5 (clone 38) (*red*). **k–p** FISH patterns of mitotic chromosomes of alfalfa (*M. sativa*) cv. 'Zhongmu No. 1' probed with E180 (*green*) combined with: **k** 18S–26S rDNA; **l** 5S rDNA; **m** *Ms*CR-1; **n** *Ms*CR-2; **o** *Ms*CR-3; **p** *Ms*CR-4. *Scale bar* = 10 μm

1, 7, 10 (Fig. 5f); *Ms*CR-5: 1, 2, 3, 4, 5, 6, 11, 12, 14, and one chromosome of pair 15 (Fig. 5g); *Ms*TR-1: all except 1, 7, and one chromosome of pair 15 (Fig. 5h); clone 65: all except 14, and one of pairs 1, 9, 12 (Fig. 5i); clone 74: all except 1, 7, and one of pair 12 (Fig. 5j); clone 68: all except 7 and 14; clone 87: all.

To test the conservation or variability of the identified repeats among different alfalfa varieties, four distinct cultivars, the Chinese cultivars 'Zhongmu No. 1' and 'Aohan', USA cultivar 'Golden Empress', and Canadian cultivar 'Algonquin', were examined with different FISH combinations. Three repeats, E180, *Ms*CR-3, and *Ms*TR-1, were used as representatives of the identified repeat sequences. Two FISH combinations, one consisting of 18S–26S rDNA, 5S rDNA, and E180 (Fig. 4e–h) and another comprising E180, *Ms*CR-3, and *Ms*TR-1 (Fig. 4i–l), were applied to these four cultivars. Though significant variations in the presence and size of the repeated sequences were observed among the cultivars, relatively conserved karyotypes were still obtained with these chromosome markers (Fig. 6).

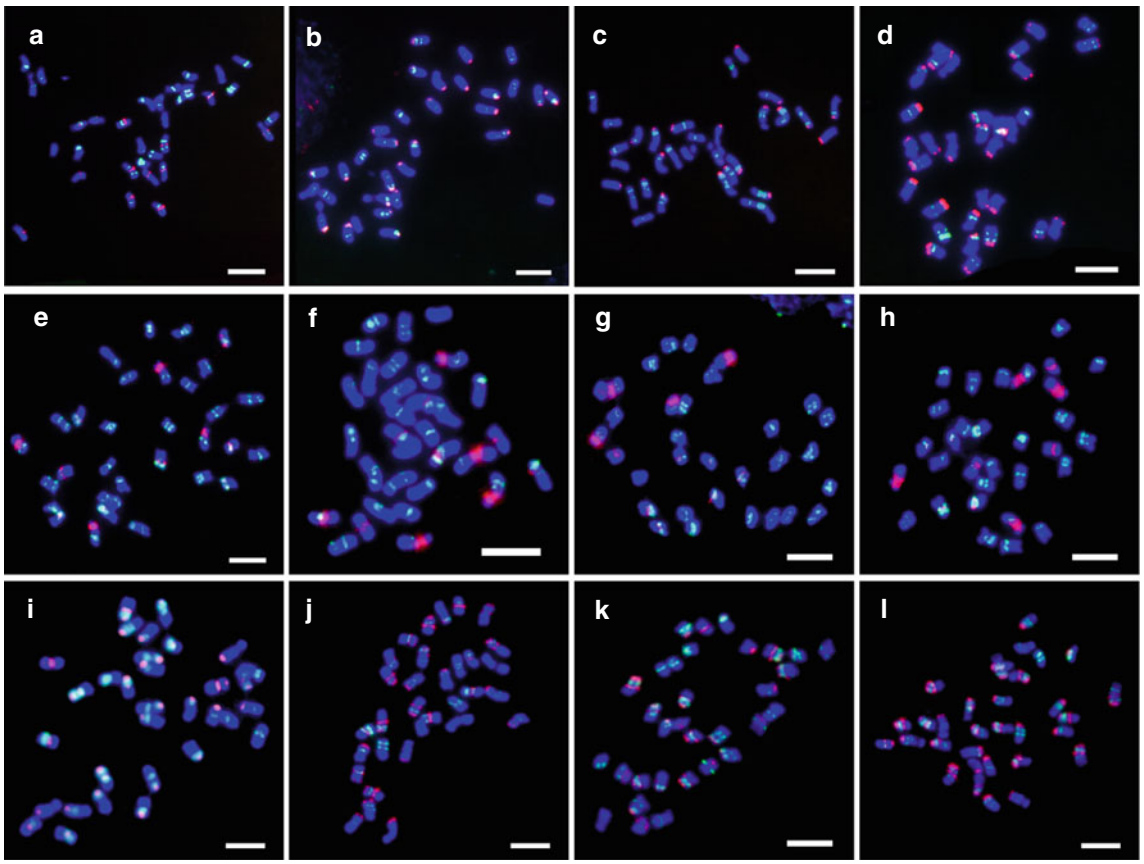

Fig. 4 FISH patterns of mitotic chromosomes of alfalfa (*M. sativa*) cv. 'Zhongmu No. 1' probed by E180 (*green*) combined with: **a** *Ms*CR-5; **b** *Ms*TR-1; **c** clone 65; **d** clone 74. **e–h** FISH patterns when probed by E180 (*green*) combined with 18S–26S rDNA (*red*) and 5S rDNA (*red*) in alfalfa cultivars: **e** 'Zhongmu No. 1'; **f** 'Aohan'; **g** 'Golden Empress'; **h** 'Algonquin'. **i–l** FISH patterns when probed by E180 (*green*) combined with *Ms*TR-1 (*red*) and *Ms*CR-3 (*red*) in alfalfa cultivars: **i** 'Zhongmu No. 1'; **j** 'Aohan'; **k** 'Golden Empress'; **l** 'Algonquin'. *Scale bar* = 10 μm

Alfalfa is regarded as an autotetraploid. However, our results suggest that diploid-like karyotyping is possible by applying multiple repeat sequences. Each chromosome can be identified by the following features:

Chromosomes 1 and 2

Strong hybridization signals of 18S rDNA and 5S rDNA are found on the middle of the long arms. Both chromosome 1 and 2 frequently carry weak E180 signals on the short arms with some variations. The different distribution patterns of E180 discriminate chromosome 1 from 2.

Chromosomes 3 and 4

E180 signals are present in the interstitial regions of the short and long arms, and *Ms*CR-3 is frequently present between the E180 signals.

Chromosomes 5 and 6

Both have similar E180 distribution patterns to chromosomes 3 and 4. However, the E180 signals in the long arms that are co-localized with *Ms*CR-2 and *Ms*CR-3 are much closer to the centromere than those in chromosomes 3 and 4.

Chromosomes 7 and 8

Both show E180 hybridizations on their short arms and the pericentric regions of their long arms. Chromosome 7 is clearly identified by the absence of terminal repetitive sequences such as *MS*TR-1 and clone 74. Weaker hybridization at the pericentric region of the long arm in chromosome 8 also distinguishes it from chromosome 7.

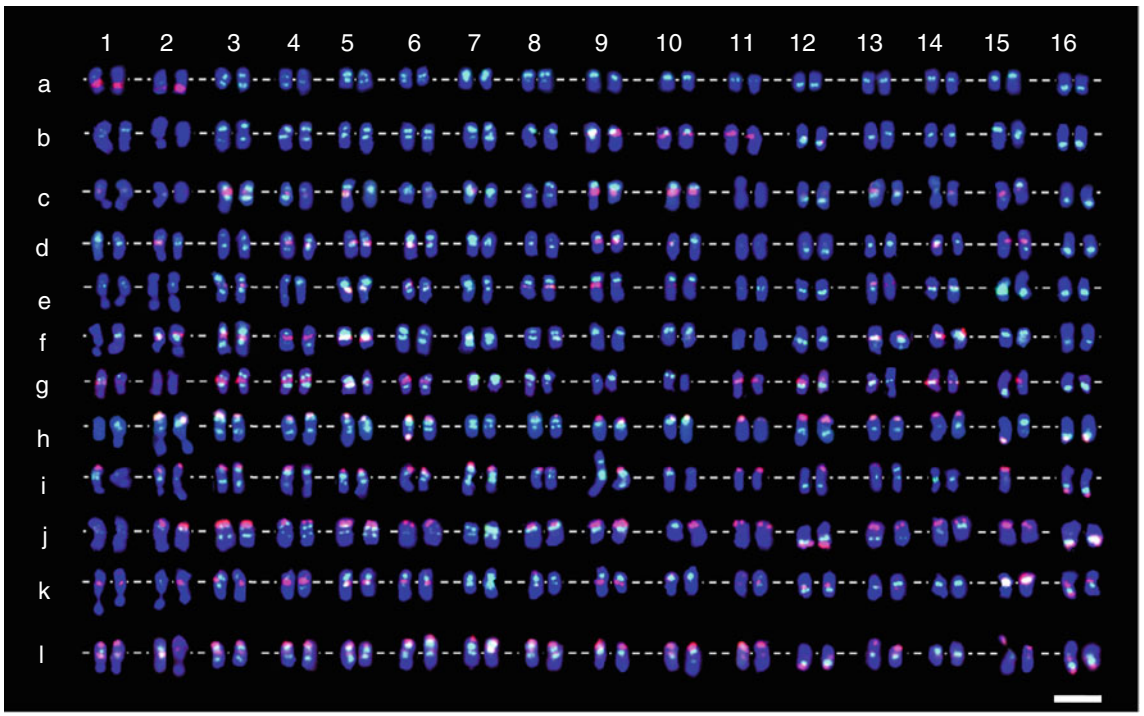

Fig. 5 Repeat localization on alfalfa 'Zhongmu No. 1' somatic chromosomes with the probes E180 (*green*) and: **a** 18S–26S rDNA; **b** 5S rDNA; **c** *Ms*CR-1; **d** *Ms*CR-2; **e** *Ms*CR-3; **f** *Ms*CR-4; **g** *Ms*CR-5; **h** *Ms*TR-1; **i** clone 65; **j** clone 74; **k** clone 68; **l** clone 87. *Scale bar* = 10 μm

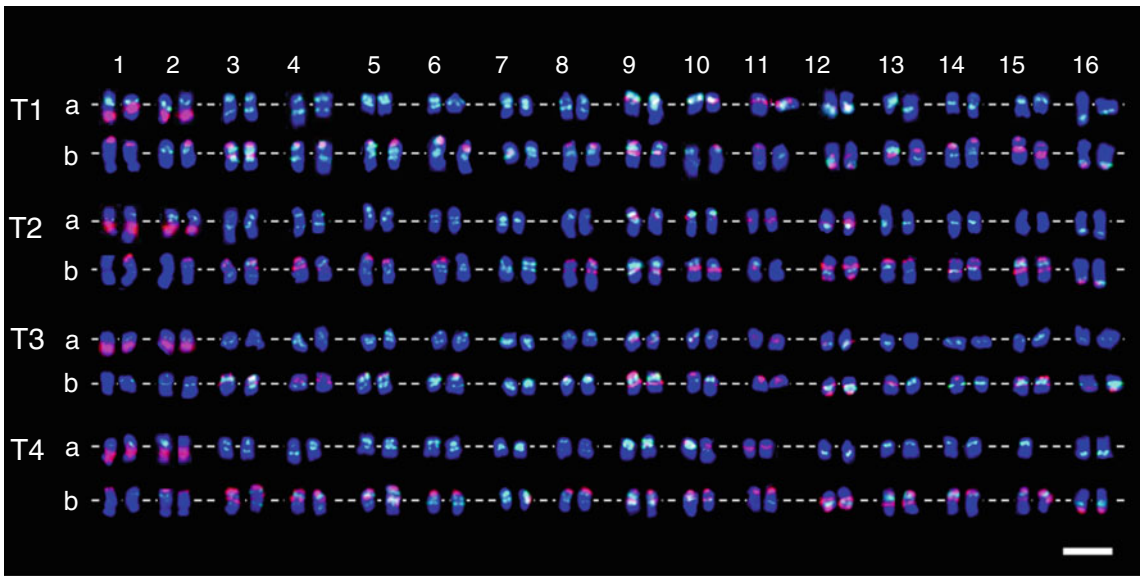

Fig. 6 Karyotyping of four alfalfa cultivars with different FISH combinations. T1, T2, T3, and T4 are cultivars 'Zhongmu No. 1', 'Aohan', 'Golden Empress', and 'Algonquin', respectively. **a** Probed by E180 (*green*) combined with 18S–26S rDNA (*red*) and 5S rDNA (*red*). **b** Probed by E180 (*green*) combined with *Ms*TR-1 (*red*) and *Ms*CR-3 (*red*). *Scale bar* = 10 μm

Chromosomes 9 and 10

5S rDNA signals are distributed between two E180 hybridizations on their short arms. The hybridization intensities of E180 discriminate chromosome 10 from chromosome 9.

Chromosomes 11 and 12

5S rDNA signals are present on the short arm and the long arm in chromosome 11 and 12, respectively. Chromosome 11 has weak or no E180 signals. The 5S rDNA signal on chromosome 12 is co-localized with E180 hybridization.

Chromosomes 13 and 14

Both are characterized by an absence of E180 hybridization around the centromere in most cases.

Chromosomes 15 and 16

E180 hybridizations are present on the terminal part of the short arm in chromosome 15, although a few variants were detected for this chromosome. Chromosome 16, showing E180 hybridization in the long arm, can be clearly discriminated from the others.

Significant chromosome variations were revealed in this study within and between varieties, possibly due to the outcrossing nature of alfalfa. Aneuploids, with a chromosome number of 31, were detected among the 'Zhongmu No. 1' (Figs. 4c and 5i) and 'Algonquin' (Figs. 4h, 6b T4) individuals. An exceptional variant featuring chromosome translocation of chromosome 9 was detected in 'Zhongmu No. 1' (Figs. 4c, 5i). Chromosomes 3, 12, and 16 were the most conserved, while the chromosome 15 was the most variable. The karyotype comparison of the four cultivars showed that 'Algonquin', in which chromosomes 13 and 14 showed E180 hybridization in the terminal parts of their short arms instead of around the centromere, was distinct from the others.

Discussion

Karyotyping of alfalfa is difficult because of the small size, high number, and autotetraploid nature of its chromosomes. A reference karyotype of tetraploid alfalfa was constructed using the combined techniques of C-banding and image analysis (Bauchan and Hossain 2001). It showed four similar series of chromosomes, eight sets of homologous chromosomes that were well characterized based on their chromosome morphology and C-banding. However, technique dependence hampered its wide application across different laboratories.

The recently developed FISH technique is a powerful tool for chromosome identification. To date, only commonly used rDNAs have been used as chromosome markers for FISH in alfalfa. The 18S–26S rDNA sequences were mapped to two pairs of chromosomes, while 5S rDNA appeared in four pairs of sites (Falistocco 2000). Therefore, the identification of all chromosomes is not possible by FISH based on these limited markers. Simultaneous FISH with a combination of repetitive sequences and the 5S and 18-26S rDNA genes can be very helpful for physical map assembly (Galasso et al. 2001). In this study, the distribution of each identified sequence on the chromosomes was well described using 5S rDNA, 18S–26S rDNA, and E180 as reference markers. Because of the outcrossing nature of alfalfa, significant chromosome variations were detected within or between cultivars. However, comparable karyotypes could still be obtained across varieties using these markers. Alfalfa is regarded as an autotetraploid, and a few chromosomes showed similar FISH patterns. Our study suggests that a diploid-like karyotype is possibly available using the above chromosome markers. E180 produced different FISH patterns on most of the chromosomes, and 5S rDNA, 18S–26S rDNA, *Ms*CR-2, *Ms*CR-3, *Ms*CR-4, and *Ms*CR-5 produced hybridization on a portion of the chromosomes. Thus, they are more powerful markers for chromosome identification than the others we identified. Because of technical limitations, only two or three probes labeled by two-color fluorescence can be simultaneously applied in FISH detection. Using a multicolor FISH procedure with several repetitive sequences as the probe mixture, Kato et al. (2004) developed a detailed molecular karyotype for maize (*Zea mays* L.). More effective and detailed molecular karyotypes could be developed using a multicolor FISH procedure. BAC clone-based FISH has been developed specifically for constructing cytological maps of species with small genomes (Jiang and Gill 2006). It is a very efficient approach that integrates genetic linkage maps with chromosomal maps and has been successfully applied in small genome species such as common bean (Fonsêca et al. 2010), cucumber (Han et al. 2011), and sugar beet (Paesold et al. 2012). Further research to construct an integrated molecular cytogenetic map with BACs or fosmid clones would be very helpful for alfalfa genetic and genomic analysis. The chromosome number of alfalfa should be rationally assigned by an integrated map rather than just by the FISH patterns observed in the present study.

Tetraploid alfalfa has a haploid genome size of 1.7 pg of DNA (Blondon et al. 1994), which consists of 22 % highly repeated plus foldback sequences and 42 % moderately repeated sequences (Winicov et al. 1988). C-banding revealed that the constitutive heterochromatic DNAs that comprise thousands of copies of highly repeated sequences are distributed as centromeric, interstitial, and terminal bands on the short arms of most of the alfalfa chromosomes (Bauchan and Hossain 2001). In this study, 56.7 % of the Cot-1 library clones were characterized as centromeric (Type 1), 12.1 % were characterized as telomeric or subtelomeric (Type 2), and 20.0 % showed centromeric telomeric, subtelomeric, and interstitial bands simultaneously (Type 3). It can be inferred that the centromeric heterochromatin detected by C-banding comprises much more Type 1 sequences than Type 3 sequences. In addition, the fact that not a single repeat was uncovered in the centromeric, subtelomeric, and interstitial regions of all chromosomes reflects the heterogeneous nature of the constitutive heterochromatin detected by C-banding.

Plant centromere sequences are characterized by long arrays of highly repetitive satellite sequence, interspersed frequently with centromeric retrotransposons (Cheng et al. 2002; Jin et al. 2004). The five centromeric or pericentromeric repetitive sequences revealed in this study consist of three satellite DNAs and two minisatellite DNAs. *Ms*CR-3 showed 86.84 % homology to *Mt*R3, a satellite repeat located in the functional centromere of *M. truncatula* (Kulikova et al. 2004). This suggests that the sequence *Ms*CR-2 revealed in this study is possibly a functional portion of the centromeres of *M. sativa*, and that the sequence of *Mt*R-3 or *Ms*CR-2 is a well-conserved centromeric sequence across species in *Medicago*. *Ms*CR-3 hybridization produced strong signals on 14 chromosomes and weak signals on 10 chromosomes. The lack of signals on the other ten chromosomes could be related to its lower abundance, which could not be detected by our FISH method. The other two satellite and two minisatellite DNAs were characterized as centromeric or pericentromeric repeats based on their hybridization patterns. Whether they are functional centromeric components or just centromere-associated sequences needs to be confirmed by further study.

Alfalfa is regarded as an autotetraploid with four nearly identical genomes. The diploid subspecies *M. sativa* ssp. *caerulea* (Less. ex Ledeb.) Schmalh. ($2n = 2x = 16$) is considered to be the progenitor of cultivated tetraploid alfalfa (McCoy and Bingham 1988). *M. sativa* ssp. *falcata* (L.) arcangeli differs from ssp. *caerulea* by the morphologies of yellow flowers and sickle-shaped pods. It has both diploid ($2n = 2x = 16$) and tetraploid ($2n = 4x = 32$) forms. *M. sativa* ssp. *falcata* is a valuable germplasm resource for the improvement of alfalfa due to its extreme winter hardiness, broad crowns, creeping root habit, and some foliar disease resistance (Barnes et al. 1977). Cytological research demonstrated that *M. sativa* ssp. *falcata* and *M. sativa* ssp. *caerulea* have a close genetic relationship (Lesins 1957; Cleveland and Stanford 1959). Our study revealed that a few tetraploid alfalfa chromosomes could be separately stained with *Ms*CR-3 and *Ms*CR-4 by FISH. Double-target FISH performed on an ssp. *falcata* diploid accession showed that only 1 chromosome carries an *Ms*CR-3 signal and 12 chromosomes bear distinct *Ms*CR-4 signals. Similar results could be obtained in double-target FISH probed by *Ms*CR-3 combined with *Ms*CR-2 and *Ms*CR-5 (Fig. 3i, j). This indicates that *M. sativa* ssp. *falcata* might have a close relationship with some of the chromosomes in alfalfa.

We identified several new chromosome markers in this study. Investigation with these chromosome markers in related taxa will provide valuable information for understanding the phylogenetics of alfalfa. These FISH markers will also be useful for germplasm resource identification and breeding applications.

Acknowledgments We thank Professor Tao Wang (China Agriculture University) for providing the seeds of some alfalfa materials. This study was financially supported by the "Joint Scholars" program of the "Lights in the Western Region" talent cultivation plan of the Chinese Academy of Sciences and partly supported by the Main Direction Program for Knowledge Innovation of the Chinese Academy of Sciences (KSCX2-EW-Q-23).

References

Allen G, Flores-Vergara M, Krasynanski S, Kumar S, Thompson W (2006) A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. Nat Protoc 1:2320–2325

Barnes D, Bingham E, Murphy R, Hunt O, Beard D, Skrdla W, Teuber L (1977) Alfalfa germplasm in the United States: genetic vulnerability, use, improvement and maintenance. USDA Technical Bulletin 1571. ARS, Washington DC, p 21

Bauchan G, Hossain M (1997) Identification of Alfalfa chromosomes using Geimsa banding and image analysis techniques. In: Proceedings of the XVIII International Rangeland Congress, Canada, pp 61–62

Bauchan GR, Hossain MA (2001) Distribution and characterization of heterochromatic DNA in the tetraploid African population alfalfa genome. Crop Sci 41:1921–1926

Begum R, Alam SS, Menzel G, Schmidt T (2009) Comparative molecular cytogenetics of major repetitive sequence families of three *Dendrobium* species (Orchidaceae) from Bangladesh. Ann Bot 104:863–872

Blondon F, Marie D, Brown S, Kondorosi A (1994) Genome size and base composition in *Medicago sativa* and *Medicago truncatula* species. Genome 37:264–270

Calderini O, Pupilli F, Paolocci F, Arcioni S (1997) A repetitive and species-specific sequence as a tool for detecting the genome contribution in somatic hybrids of the genus *Medicago*. Theor Appl Genet 95:734–740

Cheng Z, Dong F, Langdon T, Ouyang S, Buell CR, Gu M, Blattner FR, Jiang J (2002) Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. Plant Cell 14:1691–1704

Cleveland R, Stanford E (1959) Chromosome pairing in hybrids between tetraploid *Medicago sativa* L. and diploid *Medicago falcata* L. Agron J 51:488–492

Dou QW, Chen ZG, Liu YA, Tsujimoto H (2009) High frequency of karyotype variation revealed by sequential FISH and GISH in plateau perennial grass forage *Elymus nutans*. Breed Sci 59:651–656

Falistocco E (2000) Brief communication. Physical mapping of rRNA genes in *Medicago sativa* and *M. glomerata* by fluorescent in situ hybridization. J Hered 91:256–260

Falistocco E, Falcinelli M, Veronesi F (1995) Karyotype and C-banding pattern of mitotic chromosomes in alfalfa, *Medicago sativa* L. Plant Breed 114:451–453

Fonsêca A, Ferreira J, dos Santos TR, Mosiolek M, Bellucci E, Kami J, Gepts P, Geffroy V, Schweizer D, dos Santos KGB, Pedrosa-Harand A (2010) Cytogenetic map of common bean (*Phaseolus vulgaris* L.). Chromosome Res 18:487–502

Fukui K, Kamisugi Y, Sakai F (1994) Physical mapping of 5S rDNA loci by direct-cloned biotinylated probes in barley chromosomes. Genome 37:105–111

Galasso I, Schmidt T, Pignone D (2001) Identification of *Lens culinaris* ssp. *culinaris* chromosomes by physical mapping of repetitive DNA sequences. Chromosome Res 9(3):199–209

Gillies CB (1970) Alfalfa Chromosomes 1. Pachytene karyotype of diploid *Medicago falcata* L. and its relationship to *Medicago sativa* L. Crop Sci 10:169–171

Han Y, Zhang Z, Huang S, Jin W (2011) An integrated molecular cytogenetic map of *Cucumis sativus* L. chromosome 2. BMC Genet 12:18

Ho KM, Kasha KJ (1972) Chromosome homology at pachytene in diploid *Medicago sativa, Medicago falcata* and their hybrids. Can J Genet Cytol 14:829–838

Jiang JM, Gill BS (2006) Current status and the future of fluorescence in situ hybridization (FISH) in plant genome research. Genome 49:1057–1068

Jin W, Melo JR, Nagaki K, Talbert PB, Henikoff S, Dawe RK, Jiang J (2004) Maize centromeres: organization and functional adaptation in the genetic background of oat. Plant Cell 16:571–581

Kato A, Lamb JC, Birchler JA (2004) Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. Proc Natl Acad Sci USA 101:13554

Kulikova O, Geurts R, Lamine M, Kim DJ, Cook DR, Leunissen J, De Jong H, Roe BA, Bisseling T (2004) Satellite repeats in the functional centromere and pericentromeric heterochromatin of *Medicago truncatula*. Chromosoma 113:276–283

Lesins K (1957) Cytogenetic study on a tetraploid plant at the diploid chromosome level. Can J Bot 35:181–196

Masoud SA, Gill BS, Johnson LB (1991) C-Banding of alfalfa chromosomes standard karyotype and analysis of a somaclonal variant. J Hered 82:335–338

McCoy TJ, Bingham ET (1988) Cytology and cytogenetics of alfalfa. In: alfalfa and alfalfa improvement. Agron. Monogr. 29. ASA, CSSA and SSSA. Madison, WI, pp 737–776

Paesold S, Borchardt D, Schmidt T, Dechyeva D (2012) A sugar beet (*Beta vulgaris* L.) reference FISH karyotype for chromosome and chromosome-arm identification, integration of genetic linkage groups and analysis of major repeat family distribution. Plant J 72:600–611

Richards EJ, Ausubel FM (1988) Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. Cell 53:127–136

Rosato M, Galian JA, Rossello JA (2012) Amplification, contraction and genomic spread of a satellite DNA family (E180) in *Medicago* (Fabaceae) and allied genera. Ann Bot 109:773–782

Schmidt T, Heslop-Harrison JS (1998) Genomes, genes and junk: a model of the large-scale organization of plant chromosomes. Trends Plant Sci 3:195–199

Winicov I, Maki DH, Waterborg JH, Riehm MR, Harrington RE (1988) Characterization of the alfalfa (*Medicago sativa*) genome by DNA reassociation. Plant Mol Biol 10:369–371

Xia X, Erickson L (1993) An AT-rich satellite DNA sequence, E180, in alfalfa (*Medicago sativa*). Genome 36:427–432

Zakrzewski F, Wenke T, Holtgrawe D, Weisshaar B, Schmidt T (2010) Analysis of a Cot-1 library enables the targeted identification of minisatellite and satellite families in *Beta vulgaris*. BMC Plant Bio 10:8

Zhang L, Xu C, Yu W (2012) Cloning and characterization of chromosomal markers from a Cot-1 library of peanut (*Arachis hypogaea* L.). Cytogenet Genome Res 137:31–41

Zwick MS, Hanson RE, Islam-Faridi MN, Stelly DM, Wing RA, Price HJ, McKnight TD (1997) A rapid procedure for the isolation of Cot-1 DNA from plants. Genome 40:138–142