Theor Appl Genet (2002) 104:577–585

L. Wen · H.V. Tang · W. Chen · R. Chang · D.R. Pring
P.E. Klein · K.L. Childs · R.R. Klein

Development and mapping of AFLP markers linked to the sorghum fertility restorer gene *rf4*

Received: 20 June 2001 / Accepted: 3 August 2001

Abstract The restoration of male fertility in the sorghum IS1112 C (A3) male-sterile cytoplasm is through a two-gene gametophytic system involving complementary action of the restoring alleles *Rf3* and *Rf4*. To develop markers suitable for mapping *rf4*, AFLP technology was applied to bulks of sterile and fertile individuals from a segregating BC_3F_1 population. Three AFLP markers linked to *rf4* were identified and subsequently converted to STS/CAPS markers, two of which are co-dominant. Based on a population of 378 BC_1F_1 individuals, two STS/CAPS markers, LW7 and LW8, mapped to within 5.31 and 3.18 cM, respectively, of *rf4*, while an STS marker, LW9, was positioned 0.79 cM on the flanking side of *rf4*. Markers LW8 and LW9 were used to screen sorghum BAC libraries to identify the genomic region encoding *rf4*. A series of BAC clones shown to represent a genomic region of linkage group E were identified by the *rf4*-linked markers. A contig of BAC clones flanking the LW9 marker represent seed clones on linkage group E, from which fine mapping of the *rf4* locus and chromosome walking can be initiated.

Communicated by G. Wenzel

L. Wen · W. Chen · R. Chang
Department of Plant Pathology, University of Florida, Gainesville, FL 32611, USA

H.V. Tang · D.R. Pring (✉)
USDA-ARS, Department of Plant Pathology,
University of Florida, Gainesville, FL 32611, USA
e-mail: drpg@mail.ifas.ufl.edu
Tel.: +352-392-3638, Fax: –352-392-6532

P.E. Klein · K.L. Childs
Institute for Plant Genomics and Biotechnology,
and Department of Biochemistry and Biophysics,
Texas A&M University, College Station, TX 77843, USA

R.R. Klein
USDA-ARS, Southern Plains Agricultural Research Center,
College Station, TX 77845, USA

Keywords Sorghum · Cytoplasmic male sterility · Fertility restoration genes · Bulk segregant analysis-AFLP · Genetic mapping · BAC

Introduction

Cytoplasmic-nuclear male sterility (CMS) in plants is an important trait in hybrid seed production. Molecular studies have revealed that the expression of novel chimeric mitochondrial open reading frames is associated with CMS in many species. These chimeric genes often include segments of known mitochondrial genes (reviewed by Schnable and Wise 1998; Kempken and Pring 1999). Nuclear genes that restore fertility to plants expressing CMS are designated restorer of fertility (RF) genes. Little is known about the molecular action of RF genes, although it is assumed that they suppress or compensate for mitochondrial dysfunction during male gamete development. To-date, the only RF gene to be identified is maize *Rf2*, which encodes a mitochondrial aldehyde dehydrogenase (Cui et al. 1996; Liu et al 2001). In contrast, most other restorer genes are associated with changes in the transcription or translation of CMS-associated mitochondrial genes or open reading frames (reviewed in Schnable and Wise 1998; Kempken and Pring 1999). The isolation and characterization of RF genes will facilitate our understanding of how these nuclear genes regulate mitochondrial function and ensure successful male-gamete development.

Fertility restoration of the IS1112C (A3) source of CMS in sorghum [*Sorghum bicolor* (L.) Moench] is through a gametophytic mechanism, in which complementary action of restoring alleles of two genes, designated *Rf3* and *Rf4*, is required for pollen viability (Tang et al. 1998; Pring et al. 1999). The F1 resulting from the cross of A3Tx398 (*rf3rf3rf4rf4*) and the line IS1112C (*Rf3Rf3Rf4Rf4*) is heterozygous at both loci (*Rf3rf3Rf4rf4*), and 25% of the male gametes are predicted to be viable. Designations of the restoring alleles as dominant are tentative, since the nature of the restorers cannot yet be de-

termined in this gametophytic system (Tang et al. 1998). The chimeric mitochondrial open reading frame *orf107* (Tang et al. 1996b) is suspected as causal of CMS associated with the IS1112C cytoplasm. *Orf107* consists of 5′ and amino-terminal sequences duplicated from *atp9*, sequences of unknown origin, and carboxy terminal sequences highly similar to those of a mitochondrial open reading frame associated with CMS in rice. In male-sterile lines *orf107* is transcribed as three transcripts, 1100, 870 and 810-nt, and a low-abundance 380 nt transcript (Tang et al. 1996b, 1998). Plants restored to partial or full fertility exhibit a dramatically altered transcript pattern wherein the abundance of the large transcripts is reduced by about 80% concomitant with accumulation of the 380-nt transcript (Tang et al. 1996b). The 380-nt transcript results from enhanced transcript processing activity (TPA) and the 5′ terminus of the transcript is within *orf107*. Action of the restoring allele *Rf3* has been assigned to the enhanced TPA (Tang et al. 1998). The restoring *Rf3* allele is tightly linked to the dominant allele *Mmt1*, which confers enhanced transcript processing 5′ to the gene *urf209*, the sorghum counterpart of maize *orf25* (Tang et al. 1996a, 1998). No information is currently available on the possible mode of action of *Rf4*, the second gene required for fertility restoration.

Amplified fragment length polymorphisms (AFLPs, Vos et al. 1995) have been used effectively in genemarker development, particularly when used in combination with bulked segregant analyses (BSA, Michelmore et al. 1991). In sorghum, AFLP technology has been effectively used to map phenotypic traits including the fertility restoration locus *rf1* (Klein et al. 2001b) and quantitative trait loci (QTLs) governing grain mould resistance (Klein et al. 2001a). While the *mmt1* locus is tightly linked to the *rf3* locus in sorghum, no genetic markers or genes linked to *Rf4* have been previously identified, thereby hindering efforts to understand the mode of action of *Rf4*. The identification of markers tightly linked to trait loci such as *rf4* is a fundamental first step towards positional cloning of genes, to a better understanding of the mode of gene action and to elucidate the interaction of the mitochondrial and nuclear genomes that accompanies fertility restoration.

The genome of sorghum (750 Mbp) has been the target of a research effort to assemble an integrated genetic and physical map, thereby permitting the positional cloning of grass genes (Klein et al. 2000). In this report we have utilized AFLP technology to identify markers tightly linked to the *rf4* locus. Utilizing recently developed genomics technology (Klein et al. 2000; Childs et al. 2001), markers linked to the locus have been physically mapped on the sorghum genome. This information represents the initial efforts directed towards positional cloning of the *rf4* and *rf3* genes.

Materials and methods

Genetic stocks

A segregating BC_3F_1 population (Tang et al. 1998; Pring et al. 1999) was utilized to recover markers linked to the *rf4* locus. This population was prepared by pollinating the male-sterile line A3Tx398 (IS1112 C cytoplasm, Tx398 nucleus; *rf3rf3rf4rf4*) with IS1112C (*Rf3Rf3Rf4Rf4*), and backcrossing to A3Tx398 for three generations. The resulting BC3 plants (*Rf3rf3Rf4rf4*) were emasculated and pollinated with B3Tx398 (*rf3rf3rf4rf4*), generating the segregating BC_3F_1 population (A3Tx398*4/IS1112C//B3Tx398). B3Tx398 is the maintainer of the male-sterile A3Tx398 line and near-isogenic to A3Tx398. Three classes of BC_3F_1 plants were identified (Table 1). Class I, fertile plants that indicate the presence of both *Rf3* and *Rf4* (*Rf3rf3Rf4rf4*); Class II, sterile plants carrying *Rf3* identified by enhanced TPA, indicating absence of *Rf4* (*Rf3rf3rf4rf4*); and Class III, sterile plants that do not show enhanced transcript processing activity, indicating the absence of *Rf3*. This classification allows the concomitant identification of AFLP markers linked to either *rf3* or *rf4*, but only Classes I and II were utilized for *rf4* marker development.

A BC_1F_1 (A3Tx398//A3Tx398/IS1112 C) population of 378 plants was constructed for estimating the recombination fraction between *rf4* and identified markers. The BC_1F_1 plants are heterozygous (*Rf3rf3Rf4rf4*) and are fertile due to the gametophytic nature of fertility restoration, in which only pollen with both *Rf3* and *Rf4* restoring-alleles is viable. Although the population retained the *Rf3* and *Rf4* restoring-alleles, it would segregate for markers at unlinked loci.

Genomic DNA preparation for AFLP and CAPS/STS analyses

DNA for bulk segregant analysis-AFLP analyses was prepared by a CTAB procedure as described by Zhang et al. (1994). Leaf tissues, typically 5 g, were ground to powder in liquid nitrogen, extracted at 65°C, de-proteinized with chloroform/octanol, and precipitated with isopropanol prior to treatment with RNase. Analyses of STS/CAPS markers were conducted using genomic DNA isolated from 378 BC_1F_1 plants, based on the PEX DNA isolation method (Williams and Ronald 1994) with modification. Briefly, five leaf discs were excised with the cap of a 1.5-ml Eppendorf

Table 1 Strategy used to identify AFLP markers linked to the sorghum *Rf4* restoring allele. Two classes of BC_3F_1 plants were identified based on their male fertility and the *orf107*-enhanced transcript processing activity (TPA) conferred by the *Rf3* restoring allele. The parents and Class I and Class II bulks were used to identify potential *rf4*-linked markers

Item	IS1112C (bulk I)	A3Tx398 (bulk II)	Class I (bulk I)	Class II (bulk II)
Assigned genotype	*Rf3Rf3 Rf4Rf4*	*rf3rf3 rf4rf4*	*Rf3rf3 Rf4rf4*	*Rf3rf3 rf4rf4*
TPA	Yes	No	Yes	Yes
Fertility	Fertile	Sterile	Fertile	Sterile
rf4-linked pattern	+	–	+	–

Table 2 Description of *rf4*-linked AFLP markers LW7, LW8 and LW9, and the derived STS and CAPS markers

AFLP markers			AFLP-derived STS or CAPS markers				
Name	Primers	Size	Size	Primers	PCR conditions		
LW7	E-CTA/ M-CAC	214 bp	779 bp	5′-GATCTACATATGTGGAGCTC-3′ 5′-AGATCTCGGCGCTGCGGAGCAGAA-3′	94°C 30 s	–60°C –90 s	–72°C –120 s
LW8 (*Eco*RI)	E-CTC/ M-CAC	298 bp	310 bp/ 275 bp	5′-AGATGGAAAGCTCTGTGCTG-3′ 5′-AACACTCCACTAGGTGGGTCA-3′	94°C 30 s	–58°C –60 s	–72°C –60 s
LW9	E-AAG/ M-CAA	487 bp	366 bp/ 358 bp	5′-ACCTTCGACGGAATGTTTCAGG-3′ 5′-TTCGTCCGTTTGTGGTGAGTTG-3′	94°C 30 s	–58°C –60 s	–72°C –60 s

tube, and DNA was extracted in 800 µl of PEX buffer. The preparations were vortexed for 5–10 s every 15 min during incubation in PEX buffer at 65°C. DNA was precipitated, washed and dried, and dissolved in 100 µl of 0.1× TE buffer (pH 8.0). PCR was conducted with 1 µ l of DNA in a reaction volume of 10–20 µl.

Bulk segregant analysis-AFLP analyses

Genomic DNA was isolated from the BC_3F_1 individuals. Using the AFLP Small Genome Primer Kit (Life Technologies, Inc.), genomic DNA was digested with *Eco*RI and *Mse*I, ligated to *Eco*RI- and *Mse*I- adapters and pre-amplified by using *Eco*RI adapter (E-) and *Mse*I adapter+C (M-C) primers provided with the AFLP Small Genome Primer Kit. Pre-amplified AFLP templates from eight members from each class were bulked for selective amplification. Ninety six selective primer combinations of E+3/M-CA+1 and M-CT+1 were used for subsequent BSA-AFLP analyses. The E+3 primers were labeled with ^{33}P-γATP using T4 polynucleotide kinase (Boehringer Mannheim, Inc.). Selective amplification was conducted according to the manufacturer's instructions. AFLP products were separated on 5% polyacrylamide gels, fixed in 10% acetic acid for 20 min, rinsed with H_2O and dried on glass plates. AFLP products were visualized by exposure to Kodak Biomax MR1 film.

Conversion of AFLP markers to STS and CAPS markers

AFLP reaction products from 96 E+3/M+3 unique primer combinations were examined for markers linked to fertility restoration in bulks from Classes I and II. AFLP markers co-segregating with fertility restoration in bulk segregant analysis were further examined in 16–21 individuals from Classes I and II. AFLP fragments, linked to fertility restoration and present in DNA from IS1112C, were cloned into pGEMT (Promega, Inc.). At least ten clones were analyzed on agarose gels, and several of the correct size were sequenced using T7 or SP6 primers. Nested primers were then selected and used to amplify the corresponding region from B3Tx398 genomic DNA. Inverse PCR (iPCR) was required to sequence flanking regions and to identify bases for the observed polymorphisms. A set of 12 tetra- or hexa-nucleotide-requiring restriction enzymes was used to digest genomic DNA from IS1112C and B3Tx398. Digested DNA fragments were self-ligated and used as a template. Outward primers were designed based on the AFLP fragment sequence and used to amplify the flanking DNA region, which was isolated and sequenced. Three AFLP markers, LW7, LW8 and LW9, were thus converted to sequence tagged site (STS) or cleaved amplified polymorphic sequence (CAPS) markers. PCR primer sequences and conditions for the three markers are listed in Table 2.

BAC screening with *rf4*-linked markers

A six-dimensional pooling array of BAC clones developed by Klein et al. (2000) was used to identify genomic clones harboring *rf4*-linked markers. Two sorghum BAC libraries were used in the construction of the pooling array with both libraries being derived from the elite sorghum genotype BTx623. The first library of 13440 clones was constructed from total sorghum DNA that had been partially digested with *Hin*dIII. An average insert size of 157 Kbp in this library was observed, although inserts up to 340 Kbp were noted. A second BAC library containing 12576 clones was constructed for this study from nuclear DNA partially restricted with *Bam*HI, with an average insert size of 150 Kbp. The construction of BAC pools and the use of pools to identify BACs harboring PCR-based markers have been previously detailed (Klein et al. 2000).

Pools of BAC clones were screened with *rf4*-linked STS markers and the address of BAC clones harboring these markers determined. *Rf4*-linked BAC clones, tentatively identified by screening BAC pools, were screened individually for the presence of *rf4*-linked markers. To further validate that PCR products from individual BACs encode a marker-derived sequence, PCR products were purified, cloned and sequenced. Sequences of BAC-derived PCR products were compared to the known sequence of the *rf4*-linked markers for validation.

Direct selection (exon capture) of cDNA with *rf4*-related BACs

Genetic markers from BAC clones flanking the *rf4* locus were identified to permit mapping of the *rf4* locus on the high-density genetic map of sorghum (Menz et al. 2001). This process was necessary as the *rf4*-linked markers were not polymorphic in the RIL population used for high-density genetic map construction (cross of BTx623 by IS3620 C). The sequence of genic regions from BTx623 and IS3620 C DNA was obtained using a cDNA direct-selection protocol (exon capture) as detailed by Childs et al. (2001). Briefly, cDNA synthesized from BTx623 and IS3620 C mRNA was captured (hybrid selected) by LW8-linked BAC clones. Following direct selection, hybrid-selected cDNAs were sequenced on an ABI 3700 as previously described (Childs et al. 2001). Sequence polymorphisms between BTx623 and IS3620C coding regions were identified and subsequently converted to genetic markers suitable for mapping in the RIL population of sorghum (Peng et al. 1999; Menz et al. 2001). The genetic map location of *rf4*-linked markers was determined by appending the raw data onto a data set used to construct the high-density genetic map of sorghum (Menz et al. 2001).

Restriction endonuclease fingerprint analysis of *rf4*-related BAC clones

BAC DNA was isolated and quantified by fluorometry as described (Klein et al. 1998). BAC DNA was digested with *Hin*dIII

and separated on 1% SeaKem agarose gels as detailed previously (Marra et al. 1997), except that a 30-well comb (1-mm thick) was used and 9 µl of each sample was loaded per each well. Gels were stained with SYBR Gold (Molecular Probes, Inc.) and images digitized with a Canon 3030 digital camera. Digitized images were imported into FPC (Soderlund 1999). The tolerance and cut-off values for automated contig assembly were as previously detailed (Marra et al. 1997; Klein et al. 2000).

Results

Identification of AFLP markers linked to *rf4*

The strategy used in isolating AFLP markers linked to *rf4*, as described above and illustrated in Table 1, was utilized with the BC_3F_1 population. Twenty one fertile Class I (*Rf3rf3Rf4rf4*) and 16 sterile Class II (*Rf3rf3rf4rf4*) members of the BC_3F_1 population were selected for AFLP analysis. Bulks of eight Class I and eight Class II individuals were utilized to identify *rf4*-linked markers. Each of the 96 primer combinations examined gave a range of 50–70 bands, with 10–15 polymorphic bands detected between the two parents. A representative gel from AFLP primer combination E-AAG/M-CAA is shown in Fig. 1A. Polymorphic bands that appeared in both IS1112 C and bulk I, but not in B3Tx398 and bulk II, were further examined among individual members of Classes I and II. Within this AFLP profile, three candidate AFLP markers, potentially linked to *rf4*, LW9, LW10 and LW11, were identified. To further examine potential linkage, AFLP templates from each of the sterile and fertile lines used to construct the Class-I and -II bulks were screened individually (Fig. 1B). Markers LW9, LW10, LW11 were present in 8, 7 and 3 of the eight Class-I fertile lines, respectively, while none of the markers were detected in any of the eight Class-II sterile lines. In total, seven AFLP markers linked to *rf4* were detected from the 96 primer combinations examined (data not shown). By screening 37 Class-I and -II BC_3F_1 individuals, it was determined that the AFLP markers LW7, LW8, LW9 were the most tightly linked of the seven AFLP markers identified. Based on their close proximity to the *rf4* locus, these three markers were converted to STS or CAPS markers for genetic mapping within a population of 378 BC_1F_1 individuals.

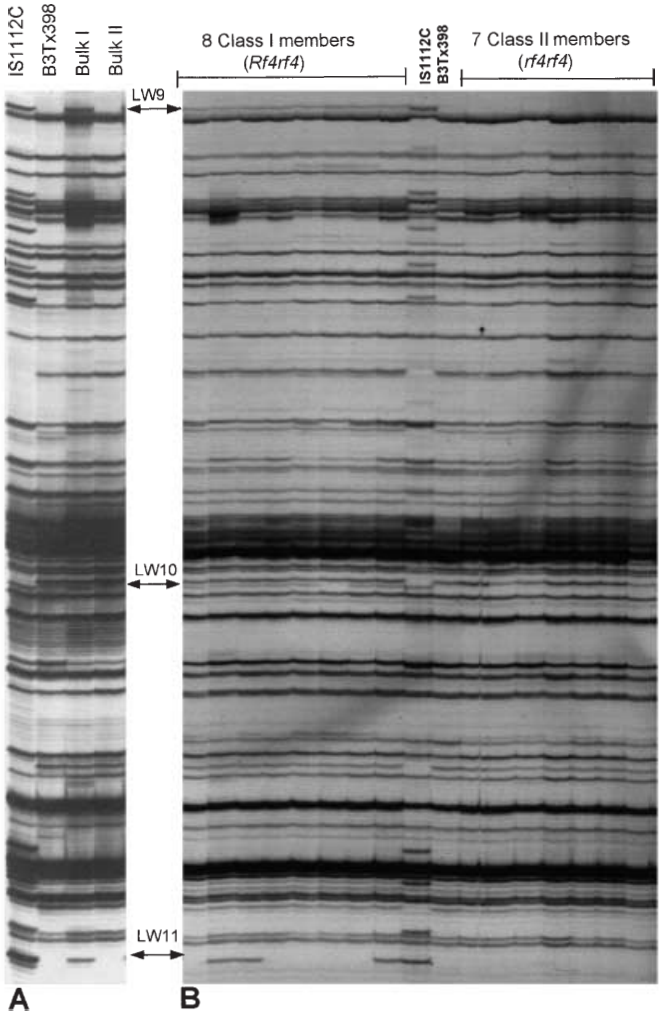

Fig. 1 A, B BSA-AFLP and analysis of Class I and II individuals using +3/+3 primer combinations *Eco*RI-AAG and *Mse*I-CAA. **A** AFLP products of parental lines IS1112C and B3Tx398, and bulks I and II, each consisting of eight members from Classes I and II, respectively. Three candidate fragments, LW9, LW10 and LW11, present in bulk I and absent in bulk II are marked with *arrows*. **B** AFLP analysis of individual members of bulks I and II

Conversion of AFLP markers to STS or CAPS markers for genetic mapping

To permit conversion of AFLP markers to CAPS or STS markers, the AFLP markers LW7, LW8, and LW9 were gel-purified, amplified, cloned and sequenced. Based on their sequence, primers were designed for amplifying B3Tx398 allelic fragments, which were then sequenced. A summary of each of these STS/CAPS markers is shown in Table 2.

The LW7 AFLP fragment (E-CTA/M-CA C) is small (214 bp; Table 2) and AT-rich, and it was not possible to design proper primers to recover corresponding B3Tx398 sequences. Inverse PCR was therefore conducted to identify 587 bp of flanking sequences from IS1112 C. Two primers from the flanking DNA regions were designed (Table 2), which generated a 779 bp fragment from IS1112 C, but no product from B3Tx398 (Fig. 2A). Confirmation of this product as linked to the *Rf4* allele was accomplished by examining 37 members of Classes I and II, which established the same segregation pattern as the AFLP fragment (Fig. 2A).

The AFLP marker designated LW8 (E-CTC/M-CAC) is 298 bp (Table 2). Internal primers were designed to amplify and sequence a 270-bp sequence from B3Tx398. A 3-bp deletion was detected in B3Tx398 as compared to IS1112 C. Inverse PCR was conducted to recover flanking DNA sequences from both parents. Sequence analyses showed that the *Eco*RI site flanking LW8 is

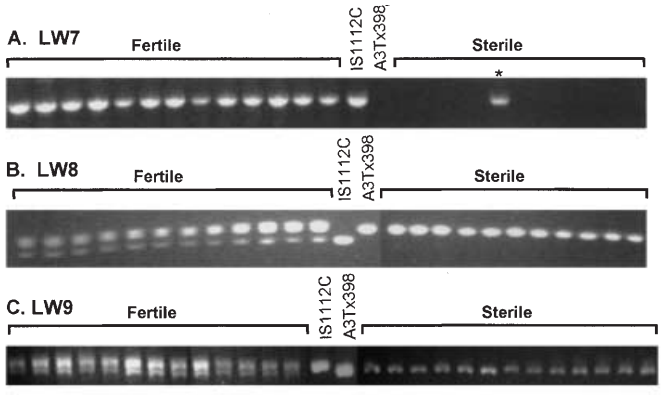

Fig. 2A–C Co-segregation of *rf4* AFLP markers LW7, LW8 and LW9, converted to STS or CAPS markers and scored on members of Class I (fertile) and II (sterile). **A** STS marker LW7, 779 bp, scored on 13 fertile and ten sterile individuals. **B** CAPS co-dominant marker LW8, 310- and 275-bp, scored on 12 fertile and 11 sterile individuals, analyzed in 2% agarose gels. **C** STS co-dominant marker LW9, 366- and 358-bp, scored on 13 fertile and 13 sterile individuals, analyzed in 4% NuSieve 3:1 agarose gels. Parental IS1112C and B3Tx398 are shown for each marker. The asterisk indicates a putative recombinant individual

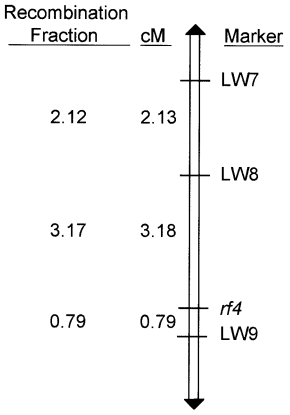

Fig. 3 Location of markers flanking the *rf4* locus on linkage group E. Recombination fraction and genetic distance in centiMorgans between adjacent loci were estimated based on the analysis of 378 BC_1F_1 plants

present in IS1112 C but absent in B3Tx398. By utilizing two suitable PCR primers flanking the *Eco*RI site for PCR and *Eco*RI digestion of the PCR products (Table 2), a polymorphism between the two mapping parents was created for the marker. *Eco*RI digestion of the PCR product amplified from the IS1112 C parent generated fragments of 275 bp and 35 bp, while the 307-bp PCR product amplified from the B3Tx398 parent cannot be digested by *Eco*RI. This marker was therefore converted to a co-dominant CAPS marker. The conversion of the AFLP marker LW8 was confirmed by examining 37 Class I and II members (Fig. 2B).

The LW9 AFLP marker (E-AAG/M-CAA) is 487 bp (Table 2). Two internal primers were designed and used to generate a fragment of 366 bp from IS1112 C. Sequencing analysis of this fragment showed an 8-bp deletion in B3Tx398 relative to IS1112 C, generating a 358-bp fragment that can be distinguished from the 366-bp fragment. LW9 was therefore converted from an AFLP marker to a co-dominant STS marker. The conversion was confirmed by examining 37 Class I and II members (Fig. 2C).

Having successfully converted each of the three *Rf4*-linked markers to STS or CAPS markers, genetic mapping with the BC_1F_1 population was conducted. LW9 displayed three recombinant individuals out of the 378 BC_1F_1 plants examined, while LW8 and LW7 showed 12 and 20 recombinant BC_1F_1 individuals, respectively. The markers LW7, LW8 and LW9 displayed a recombination fraction with the *rf4* locus of 5.29±1.15%, 3.17±0.9%, and 0.79±0.46%, respectively. Analysis of the recombinant events amongst the three STS/CAPS markers linked to *rf4* provided the likely order of these three markers and the fertility restoration locus (Fig. 3). Each of the 12 recombinants for LW8 was also recombinant for LW7, indicating that both markers are on the same side of the *rf4* locus. One additional plant recombinant for LW8 retained the LW7 marker, indicating a double-crossover. The three recombinants for LW9 are unique and as such are not a subset of the recombinants observed for either LW8 or LW7. Based on these results, LW9 was mapped across the *rf4* locus from LW8 and LW7 (Fig. 3). Genetic distances for the three markers were estimated from their recombination fraction using the Kosambi function (Kosambi 1944), and are 5.31 cM, 3.18 cM and 0.79 cM for LW7, LW8 and LW9, respectively.

Identification of the BAC addresses of *rf4*-linked genetic markers

To identify the region of the sorghum genome that encodes *rf4,* two BAC libraries consisting of 24576 clones were screened with the *rf4*-linked genetic markers LW7, LW8 and LW9. Screening was facilitated by a BAC pooling strategy developed by the Texas A&M University–ARS Sorghum Genomics Research Group (Klein et al. 2000). Screening of the pools with the STS or CAPS markers revealed unique addresses of BACs encoding the markers LW8 and LW9, whereas a LW7-derived signal was detected in all BAC pools. The presence of a repetitive element within the sequence of LW7 most likely accounts for the detection of a strong signal for LW7 in all BAC pools examined. Thus, efforts to map *rf4*-linked markers to the sorghum genetic-physical map were restricted to the markers LW8 and LW9.

To permit integration of LW8-related BACs into the genetic-physical map of sorghum, sequence polymorphisms between BTx623 and IS3620 C (parental lines of the mapping population) were identified in the genomic region flanking the LW8 marker. To identify sequence polymorphisms, cDNAs from BTx623 and IS3620C were hybrid-selected (exon capture) with LW8-related BAC clones (Childs et al. 2001). From approximately 100 hybrid selected cDNAs (cDNA fragments of approximately 150 bp), several polymorphisms were detected between BTx623 and IS3620C. Of the sequence polymorphisms detected, one was successfully converted to

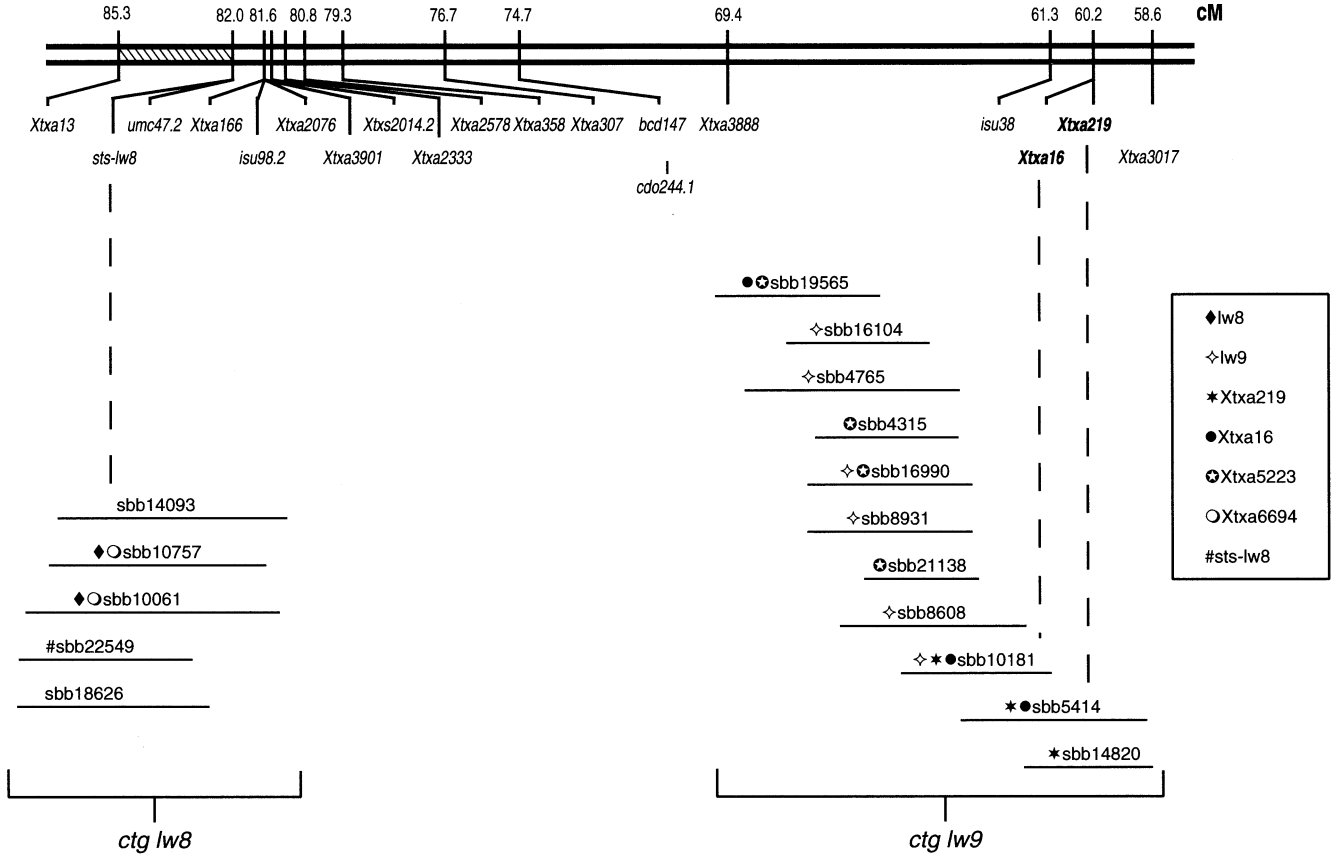

Fig 4 Integration of rf4-linked markers into the genetic/physical map of sorghum in a region of LG E. BACs linked to the rf4-genetic markers *lw9* and *STS-lw8* are shown below the framework (LOD≥3.00) genetic map with *dashed lines* extending down from each marker linked to that BAC contig. Markers *Xtxa5223*, *Xtxa6694* and *lw8* were not polymorphic in the RIL mapping population and as such were used to identify overlapping BAC clones but did not provide links to the genetic map. *Symbols* next to each BAC clone show the genetic markers assigned to that BAC clone. *STS-lw8* refers to the genetic marker isolated from lw8-related BAC sbb22549 that permitted the integration of these BACs to the physical/genetic map. The genomic region of low recombination frequency of LG E (*hatched*) is delineated. Genetic distances are expressed as cumulative map distances from position 0.0 (first genetic marker of LG) in cM (Kosambi estimates). Locus names appear adjacent to map positions

an insertion/deletion-type marker that was mapped (LOD threshold ≥ 3.00) to the region of LG E shown in Fig. 4.

BAC contigs harboring the *rf4*-linked marker LW9 were also identified. Examination of the genetic marker data for these BAC clones revealed several genetic markers that had been previously integrated into the genetic-physical map of sorghum, and the map location of these markers, and hence LW9, is depicted in Fig. 4. In agreement with the results obtained with the LW8 marker, LW9-related BACs reside on a genomic region of LG E. Two LG E genetic markers, *Xtxa16* and *Xtxa219*, were located on sorghum BAC clones (marked with asterices in Fig. 4) that also encoded the *rf4*-linked marker LW9. The two markers LW8 and LW9 define a broad region of the sorghum genome that should encompass the *rf4* locus. At present a minimum tiling path of BAC clones spanning the region defined by LW8 and LW9 has not been assembled. Based on the genetic mapping data, the *rf4* locus should reside nearer the genomic region defined by LW9 and, as such, map closure and fine mapping have commenced from the contig flanking this marker.

Discussion

The identification and development of male-sterile and fertility restorer lines is required for the exploitation of CMS systems for hybrid seed production. There are at least 22 sources of CMS in sorghum (Schertz et al. 1989), which have been assigned to several groups (Pring et al. 1995; Xu et al. 1995). Despite this wealth of genetic diversity, the A1 (milo) cytoplasm remains the primary CMS system used for hybrid seed production (Schertz et al. 1989; Pring et al. 1995). The IS1112C male-sterile cytoplasm is currently not used in hybrid sorghum seed production. F_1s resulting from the pollination of A3Tx398 with IS1112C exhibit about 50% seed set (Worstell et al. 1984), precluding utilization. The basis of this low seed set is the determination that fertility restoration is by a gametophytic system, wherein two genes, designated *Rf3* and *Rf4*, act in a complementary manner. Thus a heterozygous F_1 (*Rf3rf3Rf4rf4*) resulting from the cross A3Tx398/IS1112 C sheds only 25% viable pollen. A similar gametophytic system, the S male-

sterile cytoplasm of maize (Gabay-Laughnan et al. 1995), is dependent on a single gene for restoration, *Rf3*, which has been assigned a dominant function based on the development of tetraploids (Kamps et al. 1996). Conversion of the A3 CMS system to a single restorer gene system is predicted to result in 50% pollen viability in an F_1, which would mimic the maize S male-sterile cytoplasm system. We have begun efforts to develop molecular markers linked to *rf3* and *rf4* for assignment to linkage groups and positional cloning, and for marker-assisted selection for conversion to a single-gene restoration system. Furthermore, cloning of the restoring *Rf4* allele will permit detailed examination of the interaction of the mitochondrial and nuclear genomes that results in the restoration of pollen fertility.

We exploited the assigned action of the *Rf3* restoring allele, i.e., enhanced transcript processing of mitochondrial *orf107* (Tang et al. 1996a, 1998), to identify BC_3F_1 segregants that were male-sterile and thus heterozygous *Rf3rf3* and homozygous *rf4rf4*. These individuals were utilized in conjunction with fertile segregants, which are *Rf3rf3Rf4rf4*, for *rf4* marker development. This strategy was utilized to map the *rf4* locus in conjunction with bulk segregant analysis and AFLP high-volume marker technology. This approach was effective in identifying seven AFLP markers linked to *rf4* from a total of 96 unique AFLP primer combinations. The conversion of these AFLP fragments to STS or CAPS markers permitted rapid screening of a BC_1F_1 genetic-mapping population. The closest of these *rf4*-linked markers, LW9, mapped to within 0.79 cM of the target locus. Two other CAPS/STS markers, LW7 and LW8, were found to map to within 5 and 3 cM, respectively, of the locus. These *rf4*-linked markers apparently span the *rf4* locus based on marker analysis of the BC_1F_1 population. A parallel approach is being conducted to obtain genetic markers tightly linked to *rf3* locus in sorghum. Having genetic markers linked to both to the *rf3* and *rf4* loci will permit further genetic evaluation of this gametophytic restoration system and may permit germplasm screening for a single gene-restorer system for the A3 cytoplasm.

A collection of complementary approaches is being utilized to construct an integrated genetic, physical and gene map of the sorghum genome (Klein et al. 2000; Childs et al. 2001). Integrated genetic and physical maps are extremely valuable for map-based gene isolation, comparative genome analysis, and as a source of sequence ready clones for genome-wide sequencing efforts. In the present study, we have utilized this technology to identify the genomic region encoding *rf4* and identify sorghum BAC clones that harbor *rf4*-linked genetic markers. By screening a six-dimensional pooling array of 24576 sorghum BAC clones, contigs harboring *rf4*-linked markers were identified. The pooling strategy developed by Klein et al. (2000) provides a rapid and efficient means of identifying overlapping BAC clones containing a common genetic marker. The *rf4*-linked marker LW9 identified a series of BAC clones that could be merged into a single contig by restriction endonuclease fingerprint analysis. These BAC clones had been shown to harbor the AFLP genetic markers, *Xtxa16* and *Xtxa219*, which have been mapped to a genomic region of LG E (Menz et al. 2001). The *rf4*-linked marker LW8 also identified a contig of five BAC clones but this contig had not been previously linked by markers to the genetic map. As such, the exon-capture protocol detailed by Childs et al. (2001) was utilized to obtain the sequence of genic regions from BAC clones flanking the LW8 marker. Exon capture effectively permits sequence determination of gene tags (100–200 bp) from any genotype by capturing cDNA from that genotype with immobilized BAC DNA (Childs et al. 2001). Utilizing cDNA from the inbred lines BTx623 and IS3620C and immobilized BAC DNA from LW8-related BAC clones, the sequence of genic regions near the LW8 locus was obtained from the parental lines of our RIL mapping population. Sequence polymorphisms were thus detected that could be mapped in the genetic mapping population, effectively permitting the integration of LW8-related BACs into the genetic-physical map. The results of this mapping effort showed that, in agreement with the coordinates of LW9, BAC clones encoding LW8 integrated into the genetic-physical map in a region of LG E approximately 15–20 cM proximal to that of LW9 (Fig. 4).

Examination of the genetic map of Menz et al. (2001) suggests that the marker LW8 resides in a region of LG E with low recombination frequency. In the genomic region of LW8-related BACs, 98 markers were assigned in a 3.3-cM region defined by the framework (LOD>3.0) markers *umc47.2* and *Xtxa13*. It is assumed that these clusters correspond to regions of the genome that are hyper-methylated and have relatively low rates of recombination, such as centromeres. The genomic region flanking LW9 does not show such a clustering of markers since suppressed genetic recombination between genetic markers was not observed. The genetic distance between LW9 and *rf4* was estimated at 0.79 cM, based on the BC_1F_1 population, and hence the locus should reside nearer the BAC contig assigned to this marker rather than LW8 (about 3 cM from *rf4*). As such, we have initiated fine mapping from the genomic region harboring LW9. Identifying DNA polymorphisms in the BACs flanking the LW9 marker will permit fine mapping of the *rf4* locus and thus allow a more-defined genomic region flanking the locus to be marked. During this process, additional BAC clones will be identified to permit expansion of BAC contigs and, hence, facilitate map closure. These efforts will permit the identification of a well-defined genomic region spanning the *rf4* locus from which candidate gene identification will commence.

The unified grass genome concept of Bennetzen and Freeling (1993) has received strong experimental support from recent comparative genome mapping studies of Gramineae species. To this end, we attempted to compare the map location of RF loci in maize and rice to *rf4* in sorghum. Peng et al. (1999) and Menz et al. (2001) utilized sorghum, maize, oat, barley and rice DNA clones to construct a consensus grass map that defines

syntenic regions of grass genomes. From these studies, it was determined that the region of LG E that harbors *rf4* corresponds to maize chromsome 4 (bin 4.05). Schnable and Wise (1994) generated detailed genetic maps of the *rf1*- and *rf2*-containing regions, while Kamps and Chase (1997) developed a RFLP map of maize gametophytic *rf3* locus. These studies clearly showed that *rf1* and *rf2* are encoded on chromosomes 3 and 9, respectively, while the *rf3* locus is on the long arm of maize chromosome 2. Hence, homologues of the *rf4* locus in sorghum and maize were not apparent based on colinearity of the genomes. Furthermore, colinear regions of the rice genome harboring fertility restoration loci (Akagi et al. 1996; Ichikawa et al. 1997; Yao et al. 1997; Tan et al. 1998) were not identified since the linkage information between rice and sorghum in this region of sorghum LG E (Peng et al. 1999; Ventelon et al. 2001) is too sparse to pinpoint homologues. Targeting rice probes to this region of sorghum LG E would permit better cross referencing of the sorghum and rice maps and would facilitate the identification of collinear regions of the rice and sorghum genomes.

Acknowledgments The authors thank Drs. C. D. Chase, E. Vallejos, and R. P. Wise for critical reviews of the manuscript, and acknowledge the technical assistance of Ye Li and Theresa Cao. Supported in part by the U.S. Department of Agriculture-National Research Initiative Competitive Grants Program, 99–35300–7725 to D. R. Pring and R. R. Klein. Sorghum physical map research supported in part by the NSF Plant Genome Research Grant DBI-0077713 (PEK). This material is based upon work supported by the National Science Foundation under Grant No. 9871390. Cooperative Investigations, U.S. Department of Agriculture-Agricultural Research Service and the Institute of Food and Agricultural Sciences, University of Florida. Journal Series number R-08197, Institute of Food and Agricultural Sciences, University of Florida.

References

Akagi H, Yokozeki Y, Inagaki A, Nakamura A, Fujimura T (1996) A codominant DNA marker closely linked to the rice nuclear restorer gene, *Rf-1*, identified with inter-SSR fingerprinting. Genome 39:1205–1209

Childs KL, Klein RR, Klein PE, Morishige DT, Mullet JE (2001) Mapping genes on an integrated sorghum genetic and physical map using cDNA selection technology. Plant J 27:243–255

Cui X, Wise RP, Schnable PS (1996) The rf2 nuclear restorer gene of male-sterile T-cytoplasm maize. Science 272:1334–1336

Gabay-Laughnan S, Zabala G, Laughnan JR (1995) S-type cytoplasmic male sterility in maize. In: Levings CS III, Vasil IK (eds) The molecular biology of plant mitochondria. Kluwer Acad pubs, Dordrecht, pp 395–452

Ichikawa N, Kishimoto N, Inagaki A, Nakamura A, Koshino Y, Yokozeki Y, Oka M, Samoto S, Akagi H, Higo K, Shinjyo C, Fujimura T, Shimada K (1977) A rapid PCR-aided selection of a rice line containing the *Rf-1* gene which is involved in restoration of the cytoplasmic male sterility. Mol Breed 3:195–202

Kamps TL, McCarty DR, Chase CD (1996) Gametophytic genetics in *Zea mays* L. dominance of a restoration-of-fertility allele (*Rf3*) in diploid pollen. Genetics 142:1001–1007

Kamps TL, Chase CD (1997) RFLP mapping of the maize gametophytic restorer-of-fertility locus (*rf3*) and aberrant pollen transmission of the nonrestoring *rf3* allele. Theor Appl Genet 95:525–531

Kempken F, Pring DR (1999) Male sterility in higher plants – fundamentals and applications. Prog Bot 60:140–166

Klein PE, Klein RR, Cartinhour SW, Ulanch PE, Dong J, Obert JA, Morishige DT, Schlueter SD, Childs KL, Ale M, Mullet JE (2000) A high-throughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. Genome Res 10:789–807

Klein RR, Morishige DT, Klein PE, Dong J, Mullet JE (1998) High throughput BAC DNA isolation for physical map construction of sorghum (*Sorghum bicolor*). Plant Mol Biol Rep 16:351–364

Klein RR, Rodriguez-Herrera R, Schlueter JA, Klein PE, Yu ZH, Rooney WL (2001a) Identification of genomic regions that affect grain mould incidence and other traits of agronomic importance in sorghum. Theor Appl Genet 102:307–319

Klein RR, Klein PE, Chhabra AK, Dong J, Pammi S, Childs KL, Mullet JE, Rooney WL, Schertz KF (2001b) Molecular mapping of the *rf1* gene for pollen fertility restoration in sorghum (*Sorghum bicolor* L.). Theor Appl Genet 102:1206–1212

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Liu F, Cui X, Horner HT, Weiner H, Schnable PS (2001) Mitochondrial aldehyde dehydrogenase activity is required for male fertility in maize. Plant Cell 13:1063–1078

Marra MA, Kucaba TA, Dietrich NL, Green ED, Brownstein B, Wilson RK, McDonald KM, Hillier LW, McPherson JD, Waterston RH (1997) High throughput fingerprint analysis of large-insert clones. Genome Research 7:1072–1084

Menz MA, Klein RR, Mullet JE, Obert JA, Unruh NC, Klein PE (2001) A high-density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP, RFLP and SSR markers. Plant Mol Biol (in press)

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

Peng Y, Schertz KF, Cartinhour S, Hart G (1999) Comparative genome mapping of *Sorghum bicolor* (L.) Moench using an RFLP map constructed in a population of recombinant inbred lines. Plant Breed 118:225–235

Pring DR, Tang HV, Schertz KF (1995) Cytoplasmic male sterility and organelle DNAs of sorghum. In: (Levings CS III, Vasil IK (eds) The molecular biology of plant mitochondria. Kluwer Academic publishers, Dordrecht, pp 461–495

Pring DR, Tang HV, Chen W, Howad W, Kempken F (1999) A unique two-gene gametophytic male sterility system in sorghum involving a possible role of RNA editing in fertility restoration. J Hered 90:386–393

Schertz KF, Sotomayor-Rios A, Torres-Cardona S (1989) Cytoplasmic-nuclear male sterility – opportunities in breeding and genetics. Proc Grain Sorg Res Util Conf 16:175–186

Schnable PS, Wise RP (1994) Recovery of heritable, transposon-induced, mutant alleles of the *rf2* nuclear restorer of t-cytoplasm maize. Genetics 136:1171–1185

Schnable PS, Wise RP (1998) The molecular basis of cytoplasmic male sterility and fertility restoration. Trends Plant Sci 3:175–180

Soderlund C (1999) FPC V4: user's manual fingerprinted contigs. Technical Report SC-01–99. The Sanger Centre, Cambridge, UK

Tan XL, Vanavichit A, Amornsilpa S, Trangoonrung S (1998) Genetic analysis of rice CMS-WA fertility restoration based on QTL mapping. Theor Appl Genet 96:994–999

Tang HV, Pring DR, Muza FR, Yan B (1996a) Sorghum mitochondrial orf25 and a related chimeric configuration of a male-sterile cytoplasm. Curr Genet 29:265–274

Tang HV, Pring DR, Shaw LC, Salazar RA, Muza FR, Yan B, Schertz KF (1996b) Transcript processing internal to a mitochondrial open reading frame is correlated with fertility restoration in male-sterile sorghum. Plant J 10:123–133

Tang HV, Chang R, Pring DR (1998) Cosegregation of single genes associated with fertility restoration and transcript processing of sorghum mitochondrial *orf107* and *urf209*. Genetics 150:383–391

Ventelon M, Deu M, Garsmeur O, Doligez A, Ghesquière A, Lorieux M, Rami JF, Glaszmann JC, Grivet L (2001) A direct comparison between the genetic maps of sorghum and rice. Theor Appl Genet 102:379–386

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new concept for DNA fingerprinting. Nucl Acids Res 23:4407–4414

Williams CE, Ronald PC (1994) PCR template-DNA isolated quickly from monocot and dicot leaves without tissue homogenization. Nucleic Acids Res 22:1917–1918

Worstell JV, Kidd HJ, Schertz KF (1984) Relationships among male-sterility inducing cytoplasms of sorghum. Crop Sci 24: 186–189

Xu GW, Cui YX, Schertz KF, Hart GE (1995) Isolation of mitochondrial DNA sequences that distinguish male-sterility inducing cytoplasms in *Sorghum bicolor* (L.) Moench. Theor Appl Genet 90:1180–1187

Yao FY, Xu CG, Yu SB, Li JX, Gao YJ, Li XH, Zhang Q (1997) Mapping and genetic analysis of two fertility restorer loci in the wild-abortive cytoplasmic male sterility system in rice. Euphytica 98:183–187

Zhang Q, Gao YJ, Yang SH, Ragab RA, Saghai Maroof MA, Li ZB (1994) A diallel analysis of heterosis in elite hybrid rice based on RFLPs and microsatellites. Theor Appl Genet 89:185–192