ORIGINAL PAPER

X. Qi · T. S. Pittaway · S. Lindup · H. Liu ·
E. Waterman · F. K. Padi · C. T. Hash · J. Zhu ·
M. D. Gale · K. M. Devos

An integrated genetic map and a new set of simple sequence repeat markers for pearl millet, *Pennisetum glaucum*

Abstract Over the past 10 years, resources have been established for the genetic analysis of pearl millet, *Pennisetum glaucum* (L.) R. Br., an important staple crop of the semi-arid regions of India and Africa. Among these resources are detailed genetic maps containing both homologous and heterologous restriction fragment length polymorphism (RFLP) markers, and simple sequence repeats (SSRs). Genetic maps produced in four different crosses have been integrated to develop a consensus map of 353 RFLP and 65 SSR markers. Some 85% of the markers are clustered and occupy less than a third of the total map length. This phenomenon is independent of the cross. Our data suggest that extreme localization of recombination toward the chromosome ends, resulting in gaps on the genetic map of 30 cM or more in the distal regions, is typical for pearl millet. The unequal distribution of recombination has consequences for the transfer of genes controlling important agronomic traits from donor to elite pearl millet germplasm. The paper also describes the generation of 44 SSR markers from a $(CA)_n$-enriched small-insert genomic library. Previously, pearl millet SSRs had been generated from BAC clones, and the relative merits of both methodologies are discussed.

X. Qi · T. S. Pittaway · S. Lindup · H. Liu · E. Waterman ·
F. K. Padi · J. Zhu · M. D. Gale · K. M. Devos
John Innes Centre, Norwich Research Park,
Colney,
Norwich, NR4 7UH, UK

C. T. Hash
ICRISAT,
Patancheru, 502 324, Andhra Pradesh, India

H. Liu
Scottish Crop Research Institute,
Invergowry,
Dundee, DD2 5DA, UK

F. K. Padi
Savanna Agricultural Research Institute,
P.O. Box 52, Nyankpala/Tamale,
Northern Region, Ghana

K. M. Devos (✉)
Department of Crop and Soil Sciences and Department of Plant
Biology, The University of Georgia,
3111 Miller Plant Sciences Building,
Athens, GA, 30602-7271, USA
e-mail: kdevos@uga.edu
Tel.: +1-706-5420925
Fax: +1-706-5420914

Introduction

Pearl millet, *Pennisetum glaucum* (L.) R. Br. (2 n=2 x=14), is the seventh most important global cereal crop and is grown on about 13 million hectares each in India and Africa and on more than 2 million hectares in the United States. Grain yields are generally low (around 500–600 kg/ha), mainly because this crop is often cultivated for subsistence under extremely harsh conditions, on marginal soils, in areas of low rainfall and with little or no inputs. Furthermore, yields in Africa have decreased over the past two decades as millet cultivation has been expanded into yet more marginal lands. Although local landraces are reasonably tolerant to a range of biotic and abiotic stresses, their grain yield potential is low. Improved varieties, on the other hand, are higher yielding under low to moderate inputs, but are also more susceptible to adverse growing conditions and to pests and diseases. Nevertheless, hybrids and open-pollinated varieties have been fairly widely adopted by farmers for cultivation in more favorable environments. Among the targets identified for further improvement of pearl millet crop yields are the development of integrated pest and crop management strategies to reduce losses due to the millet stem borer and striga, a parasitic weed, and breeding for enhanced drought tolerance, improved phosphorus uptake from poorly soluble sources and durable resistance to downy mildew and rust diseases. In all these areas, biotechnology has an important role to play.

The first genetic map of pearl millet was generated in 1994 by Liu et al. (1994). The map contained 181

restriction fragment length polymorphism (RFLP) markers covering the seven pearl millet chromosomes and spanning a genetic distance of 303 cM. A subset of these markers has subsequently been transferred to a series of different crosses that segregate for agronomically important traits. Quantitative trait loci (QTL) have been mapped for downy mildew resistance (Jones et al. 1995, 2002), drought tolerance (Yadav et al. 2002, 2003, 2004) and for characteristics involved in domestication (Poncet et al. 2000, 2002). The integration of markers previously mapped in other grass species has provided the anchor points to align the pearl millet linkage groups to other cereal genetic maps, including the model species rice. Although the pearl millet genome appears to be relatively highly rearranged relative to rice, regions of colinearity between the two species can be clearly identified (Devos et al. 2000). These regions form a framework for exploitation of the rice genomic sequence as a source of new markers and candidate genes underlying traits in pearl millet.

Simple sequence repeats (SSRs), also known as microsatellites, remain the markers of choice for practical breeding applications, in particular in developing countries. Pearl millet breeding is conducted at a number of public and private breeding organizations. A central breeding programme, and the first to apply marker-assisted selection (MAS) is that of ICRISAT, Patancheru, India. Although ICRISAT has used RFLP markers in the transfer of downy mildew resistance QTL to elite seed parents, these markers proved too labor-intensive for large-scale genotyping. SSRs are abundant in eukaryotic genomes and provide a codominant, usually highly polymorphic marker system (Tautz 1989; Akkaya et al. 1992; Morgante and Olivieri 1993; Wu and Tanksley 1993; Roder et al. 1995; Bryan et al. 1997). The development of 50 SSRs from pearl millet BAC clones has been described in Qi et al. (2001) and Allouis et al. (2001). In this paper, we report the development of a further 44 SSRs from a $(CA)_n$-enriched small-insert library. The relative merits of the production of SSR markers from BAC clones and random small-insert libraries are discussed. We also present a consensus genetic map with mapping data from four pearl millet crosses and comprising 65 SSR markers, 220 homologous and 133 heterologous RFLP markers. The maps and markers provide the base for future genomic and comparative analyses of pearl millet and for application of MAS in breeding programs.

Materials and methods

Plant materials

Four F_2 populations, derived from crosses between inbred lines, were mapped. The LGD 1-B-10 × ICMP 85410 cross comprised 133 F_2 progeny, and was the original mapping population used by Liu et al. (1994). The parents of this cross differed by a translocation (Liu et al. 1994). A second mapping population was generated from a cross between the inbred lines 81B and ICMP 451 and consisted

of 157 F_2 progeny. Two further crosses, ICMB 841 × 863B (Yadav et al. 2004) consisting of 149 F_2 progeny, and PT 732B × P1449-2 consisting of 131 F_2 progeny, showed segregation for drought tolerance and downy mildew resistance, respectively.

A panel of 20 pearl millet lines was used to conduct SSR polymorphism screens. In addition to the parents of the four mapping populations, the panel included the lines Tift 23D2B, Tift 238D1, Tift 383, IP 10401, IP 10402, IP 8214, P7-3, 843B, H 77/833-2, PRLT 2/89-33, ICMP 501 and 700481-21-8. Line 81B was the source of the DNA used in the generation of an SSR-enriched library, and a mapping parent. Its seed parent Tift 23D2B was the variety used in the construction of a pearl millet BAC library. The remaining lines were chosen because they were of interest to various pearl millet research and breeding projects.

Marker analyses

Restriction fragment length polymorphism analysis

The majority of the probe sources have been described in Liu et al. (1994) and Devos et al. (2000). The loci in the range *Xpsm1003 – Xpsm1009* were detected using cloned amplified fragment length polymorphism (AFLP; Vos et al. 1995) fragments as hybridization probes. Disease resistance gene analogues (RGAs) were isolated by Sivaramakrishnan (prefix "RMM" and "UOM"; unpublished data) and dehydrins were obtained from Close, University of California, Riverside ("TJC"; Close et al. 1989). Locus names followed the rules for wheat gene nomenclature (McIntosh et al. 1998). Methods for DNA isolation, restriction enzyme digestion, gel electrophoresis, Southern transfer, probe labeling and hybridization were as described in Devos et al. (1992). Four restriction enzymes were used; *Eco*RI, *Eco*RV, *Dra*I and *Hind*III.

Simple sequence repeat marker generation

The development of SSR markers from pearl millet BAC clones has been described in Qi et al. (2001) and Allouis et al. (2001). SSRs were also developed from a $(CA)_n$-enriched small-insert library using a protocol modified from Kijas et al. (1994). Briefly, genomic DNA from the inbred line 81B was digested with *Rsa*I and ligated to the adapter, top/bottom sequence 5'-CTCTTGCTT<u>A-GATCT</u>GGACTA-3'/5'-TAGTCC<u>AGATCT</u>AAGCAA-GAGCACA-3'. The *Bgl*II site in the adapter is underlined. After pre-amplification, DNA fragments were hybridized to biotinylated $(GT)_{15}$ oligonucleotides, and fragments carrying SSRs were selected using streptavidin-coated paramagnetic beads. The SSR-enriched DNA fragments were digested with *Bgl*II and cloned into the *Bam*HI site of either the plasmid pUC18 or pZErO. Recombinants were screened by colony hybridization using $(CA)_{15}$ as a probe. Plasmids were isolated from positive colonies using a

QIAprep Spin Miniprep Kit (Qiagen) and inserts were sequenced using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer) according to the manufacturer's recommendations. Primers were designed against the SSR flanking sequences using the program "Prime" (GCG Package Version 9, Wisconsin, Mass., USA), and PCR amplification was carried out in a 20 µl reaction volume containing 1× PCR buffer, 50 ng genomic DNA template, 30 ng each of forward and reverse primer, 4 mM dNTPs and 0.5 U *Taq* DNA polymerase. PCR conditions were as follows: denaturation at 94° for 3 min, followed by 35 cycles of denaturation at 94° for 30 s, annealing at 58–62° (depending on the melting temperature of the primers) for 30 s, and extension at 72° for 1 min. PCR products were separated on 5% denaturing polyacrylamide gels and visualized by silver staining. The loci detected by the SSR have been given the prefix "*Xpsmp*".

Map construction

Genetic maps were constructed using the software packages Mapmaker v.3.0 (Whitehead Institute for Biomedical Research, Cambridge, USA) and JoinMap 2.0 (Stam 1993). Consensus maps were first constructed by running JoinMap on the merged dataset. However, because marker orders on the consensus maps did not always reflect marker orders on the individual maps, the final consensus maps were generated by visual integration. This was achieved by using the genetic map generated in the 81B × ICMP 451 mapping population as a starting point, and interpolating the markers mapped in the other crosses into the base map according to their genetic map position. For markers that could not be placed unambiguously, a position range, indicated by a vertical bar, was assigned to the map.

Results

Isolation and characterization of SSR markers

Colony hybridization showed that on average 25% of the recombinant clones in the SSR enriched library contained a $(CA)_n$ repeat. 143 positive clones were sequenced. Clones that contained less than 50 bp of flanking sequence on one or both sides were discarded. A further 21 clones were present more than once. Primers were designed against 80 SSR-containing sequences. Of these, 44 produced simple PCR products and were screened for polymorphism against the panel of 20 pearl millet varieties (Fig. 1). The remaining primer pairs did not produce amplification products, lacked reproducibility or amplified multiple fragments. It is possible that some of these SSRs could be turned into useful markers by optimizing the PCR conditions. However, our aim was to generate markers that worked under standard sets of conditions and were easily transferable across laboratories. Optimization was there-

fore not attempted. The polymorphism information content (PIC) values varied from 0 (monomorphic) to 0.92, with an average of 0.71. The PIC values were estimated according to Anderson et al. (1993): $PIC_i = 1 - \sum_{j=1}^{n} p_{ij}2$ where p_{ij} is the frequency of the *j*th pattern of marker *i*, and the sum is made over *n* patterns, assuming homozygosity of the pearl millet inbred lines. The primer sequences, SSR repeat length, expected product size in the inbred line 81B and PIC values are given in Table 1.

Genetic maps

A total of 597 loci were mapped over the four crosses. The crosses LGD 1-B-10 × ICMP 85410 and 81B × ICMP 451 were mapped most extensively with RFLP markers. Markers evenly spread over the pearl millet genome were subsequently used to produce skeleton RFLP maps of the ICMB 841 × 863B and PT 732B × P1449-2 crosses. SSR markers were mainly mapped in the latter two crosses and in the 81B × ICMP 451 population. In total, 13 loci were mapped in all four crosses, 37 in three of the crosses, 66 in two crosses and 302 loci were mapped in only one cross. Seven linkage groups were obtained for all four mapping populations. Some distal markers showed linkage
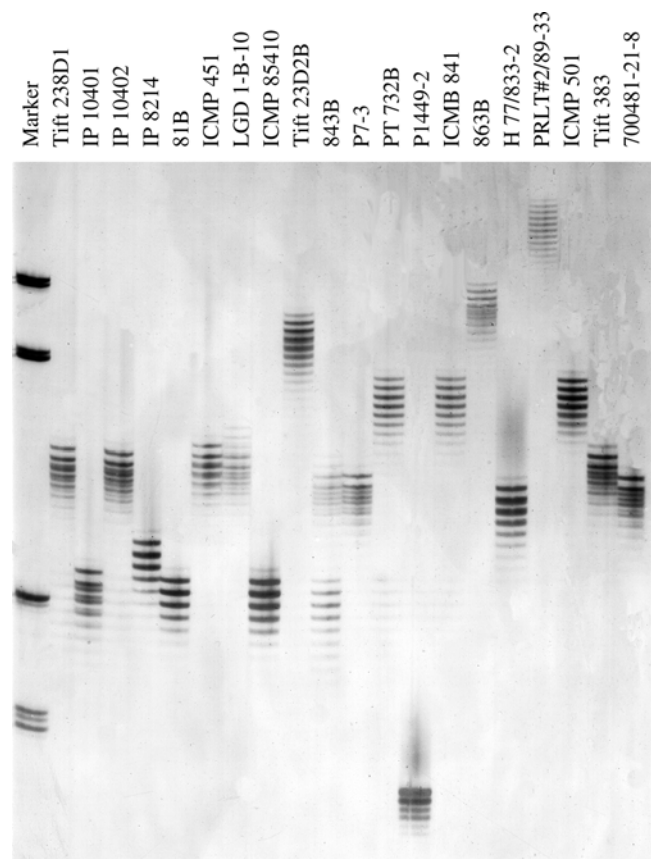

Fig. 1 Polymorphism analysis of marker PSMP2013 against a panel of 20 pearl millet lines

Table 1 SSR markers and their characteristics

Marker	Primer pair (F/R)	PCR product	Repeat type and length	PIC value
PSMP2001	CATGAAGCCAATTAGGTCTCACCATCTGACTTGTTCTTATCC	304	$(CT)_8(CA)_{48}$	0.82
PSMP2006	GACTTATAGTCACTGGGAAAGCTCGCTTTAATAACTTTGTGCGTATT	253	$(GT)_{51}$	0.87
PSMP2008	GATCATGTTGTCATGAATCACCACACTACACCTACATACGCTCC	238	$(GT)_{37}$	0.76
PSMP2013	GTAACCCACTAACCCTTACCGTCGCACAGAAAAAGAATAG	153	$(CT)_{19}(GT)_{16}$	0.88
PSMP2018	CGCAAGACATTTTAGTATCACCACAGTCATCCTCAGTCGTCC	203	$(GT)_{30}$	0.40
PSMP2019	TGTGCCACAGCTTGTTCCTCCAAGCAGCCAGTTCCTCATC	248	$(CA)_{38}$	0.85
PSMP2024	ACACCTTAGCAAAGGCTGCTACATGCTTAGTTTGCTATAACTTGGG	413	$(CA)_6(TC)_9(TG)_{30}$	0.50
PSMP2027	AGCAATCCGATAACAAGGACAGCTTTGGAAAAGGTGATCC	273	$(GT)_{31}$	0.86
PSMP2030	ACCAGAGCTTGGAAATCAGCACCATAATGCTTCAAATCTGCCACAC	107	$(CA)_{11}(GA)_{10}$	0.49
PSMP2031	CACATCCGCAAGAGACACCAAATTTTGGGGGTGTAGGTTTTGTTG	100	$(CCA)_3(TCC)_3$	0.00
PSMP2033	CTATACCATTGAATTGAAAGGTCCAATCTTTAGCTTTTTCAAGAGAC	200	$(CA)_9$	0.57
PSMP2040	CATTACACGTTTCTTCAAACGCTCTTCGGCCTAATAGCTCTAAC	163	$(CA)_{nd}$[a]	0.67
PSMP2043	TCATATTCTCCTGTCTAAAACGTCACAAATCGTACAAGTTCCACTC	192	$(CA)_{13}(GA)_6$	0.62
PSMP2045	TCATCTTCCCCTATCCGAAACACTTGCCAATGCTATCTTCAC	203	$(CA)_{11}(GA)_5$	0.74
PSMP2048	TGAATTGGGAATAAAGGAGACCACGTGTGCCTGCTTTTAGTAAC	252	$(AC)_{33}$	0.92
PSMP2050	ATCAAACGGCATCAGACAACGGATCTCTTAGTGTGGTGGAGAGC	102	$(AC)_{13}$	0.79
PSMP2052	ACTCAAGCTATGCATCAAGCAGAGGAGGGGGTAAGAATAGG	178	$(CA)_{13}$	ND
PSMP2056	ACCTGTAGCTTCAAAATTCAAAAAAATTCAGTGTGATTTCGATGTTGC	213	$(CA)_{28}(TA)_5$	0.83
PSMP2059	GGGGAGATGAGAAAACACAATCACTCGAGAGAGGAACCTGATCCTAA	119	$(AC)_{11}$	0.59
PSMP2060	AGTTATAATGTATGTGCGACACGTACCACAATTTCAATATACATGGC	220	$(CA)_{55}$	0.87
PSMP2063	GAGCACATGAAATAGGAAGCAGAAGGTAGTTATAGTTAGCTTGATC	166	$(AC)_{22}(AT)_5$	0.71
PSMP2064	ACCGAATTAAAGTCATGGATCGTTGATTCTTCTGACACAAATGAG	190	$(CA)_{56}$	0.90
PSMP2065	GCTTCACTAATCGTTTTATCCATTAGAGGATGAAGGAATTGATGAAG	207	$(AC)_{16}$	0.00
PSMP2066	ATATTAGAGCATTGCATCGCGCATAGCAGCATACAGCAGCAACTAA	267	$(AC)_{55}$	0.89
PSMP2068	CAATAACCAAACAAGCAGGCAGCTTCACTCCCACCCTTTCTAATTC	105	$(AC)_{14}$	0.90
PSMP2069	CCCATCTGAAATCTGGCTGAGAACCGTGTTCGTACAAGGTTTTGC	225	$(CA)_{26}$	0.80
PSMP2070	ACAGAAAAAGAGAGGCACAGGAGAGCCACTCGATGGAAATGTGAAA	226	$(CA)_{25}(TA)_6$	0.90
PSMP2071	TTGCAGTCCCACGAATTATTTGCTTTGAATTTATAATCCTCATACT	181	$(CA)_{14}$	0.49
PSMP2072	GAAATCTACACAAGGGTCTCCAGTACGGCAGAATGACATCTGAA	165	$(CA)_{24}$	0.90
PSMP2074	AGGACTGTAGGAGTGTGGACAACACAACCAGACCTACCAGTGAATGAGA	227	$(AC)_{11}$	0.75
PSMP2076	GGAATAGTATATTGGCAAAATGTGATACTACACACTGTAAGCATTGTC	161	$(AC)_{15}$	0.59
PSMP2077	GCCAATATTATTCCCAAGTGAACACTCTTGGTTGCATATCTTTCTTTT	180	$(CA)_{15}(TA)_8$	0.59
PSMP2078	CATGCCCATGACAGTATCTTAATACTGTTCGGTTCCAAAATACTT	172	$(CA)_{42}$	0.85
PSMP2079	AGCCGAAGGCTAATCAACAAGTGGTCAGCAGCAGATGTAA	165	$(AC)_{27}$	0.91
PSMP2080	CAGAATCCCCACATCTGCATTGCAACTGAGCGAAGATCAA	181	$(AC)_{14}$	0.77
PSMP2081	CTGTGCTGTCATTGTTACCATCAGATCACCTATTACTTTCCCT	167	$(AC)_{15}$	0.83
PSMP2082	GCAGGTCATATCTTTCGTGTCAAAAGCATCCTCAAATACCCAT	214	$(AC)_{12}$	0.26
PSMP2084	AATCTAGTGATCTAGTGTGCTTCCGGTTAGTTTGTTTGAGGCAAATGC	245	$(AC)_{42}$	0.80
PSMP2085	GCACATCATCTCTATAGTATGCAGGCATCCGTCATCAGGAAATAA	176	$(AC)_{11}$	0.73
PSMP2086	CGCTTGTTTTCCTTTCTTGCTGTTCCTTCTCAGATCCTGTGCTTTCTT	122	$(AC)_{14}$	0.83
PSMP2087	GGAACAGACTCCATACCTGAAATACCTGCCTGTGCTGTTAGT	126	$(AC)_{10}$	0.68
PSMP2088	AAGAAGCCACCAGCACAAAATGCATGAAAGTAGAGGATGGTAAA	149	$(CA)_{24}$	0.65
PSMP2089	TTCGCCGCTGCTACATACTTTGTGCATGTTGCTGGTCATT	127	$(AC)_{15}$	0.80
PSMP2090[b]	AGCAGCCCAGTAATACCTCAGCTCAGCCCTAGCGCACAACACAAACTC	178	$(CT)_{12}$	0.81

[a] Repeat length is unknown because the sequencing reaction was unable to proceed through the repeats

[b] This microsatellite was identified by sequence analysis of a BAC subclone

in only some of the crosses. In mapping populations where markers were segregating independently, they were added at a genetic distance of 50 cM to the linkage group with which linkage had been shown in one or more of the other crosses. The number of markers mapped and the genetic length of each linkage group in each of the four crosses are given in Table 2. The consensus maps are presented in Fig. 2. The individual maps generated in each of the four crosses can be viewed on MilletGenes (http://jic-bioinfo.bbsrc.ac.uk/cereals/millet.html).

Table 2 Number of markers mapped and genetic length for each linkage group in four crosses

LG	LDG-1-B-10 × ICMP 85410		81B × ICMP 451		ICMB 841 × 863B		PT 732B × P1449-2		Consensus map
	No. of markers	Length of LG (cM)	No. of markers	Length of LG (cM)	No. of markers	Length of LG (cM)	No. of markers	Length of LG (cM)	No. of markers
1	58	90.1	36	71.3	12 + 1[a]	34.0 (84.0)[c]	16	114.5	84
2	43	94.0	37 + 3[a]	51.5 + 16.2[b] (117.7)[c]	18	22.4	5 + 1[a]	42.3 (92.3)[c]	76
3	40	35.2	26	26.6	11	15.7	10	16.9	61
4	32	68.6	36	89.9	11	47.0	16	97.9	65
5	25	48.4	24	33.5	10	22.0	6	21.1	46
6	20	64.9	20	57.5	11	68.2	8	43.4	41
7	22 + 2[a]	69.0 + 5.3[b] (124.3)[c]	16	85.0	15	79.5	6	27.0	45
Total	242	472.6 (522.6)[c]	198	441.3 (491.3)[c]	89	295.8 (345.8)[c]	68	394.6 (444.6)[c]	418

[a] Markers or linkage groups that are unlinked to the main linkage group, but belong to the same chromosome
[b] The genetic length of each of the linkage groups
[c] The adjusted genetic map length, including 50 cM to account for every linkage gap

Segregation distortion

Regions of significant segregation distortion ($P \leq 0.01$) are present on LG 4 in the cross 81B x ICMP 451, on LGs 1, 2, 4, 5 and 7 in LGD 1-B-10 x ICMP 85410, on LGs 2, 4 and 5 in PT 732B x P1449-2, and on LGs 3 and 6 in ICMB 841 x 863B (see individual maps in MilletGenes). In most cases, the distortion is due to an excess of one of the parental alleles. The exceptions are LGs 1 and 2 in the cross LGD 1-B-10 x ICMP 85410, and LG 5 in PT 732B x P1449-2, where the number of heterozygotes is significantly higher than 50%. No attempt was made to investigate the cause of the distortions as most distortions appeared to be cross-specific. A possible exception is the distortion on LG 4, which is present in three of the four crosses analyzed. In all cases, there is an excess of one of the parental alleles, and the A:B ratio deviated most from 1:1 in the region spanned by *Xpsm265–Xpsm364*. This suggests that there may be a gene present in this region that affects gametophytic or zygotic viability.

Discussion

Distribution of recombination

An interesting feature of the genetic maps of pearl millet is the presence of large gaps in the distal chromosome regions. A concentration of mapped markers in centromeric regions, reflecting an unequal distribution of recombination, was first observed in the early molecular maps of wheat (Chao et al. 1989) and has since been seen in several species (Devos et al. 1992; Tanksley et al. 1992; Qi et al. 1996), but this effect appears to be extreme in pearl millet. The pearl millet map published by Liu et al. (1994) in the cross LGD 1-B-10 × ICMP 85410 contained 181 loci, but spanned only 303 cM. Currently, this map contains 242 loci and spans 473 cM. The difference in length is mainly due to the addition of 12 distally located markers. A further 50 cM can be added to the total length of the LGD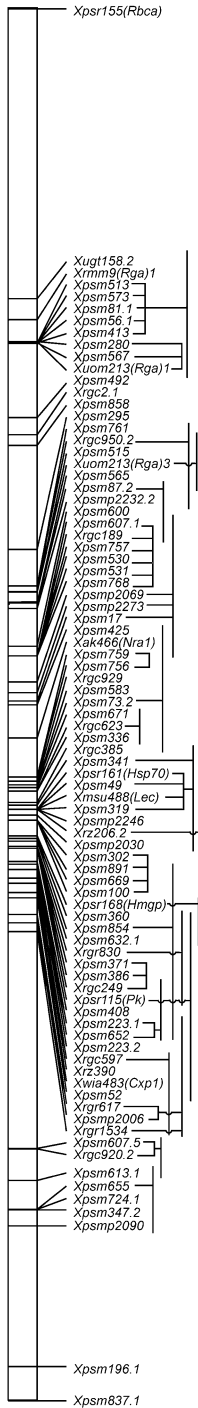 1-B-10 × ICMP 85410 map to include the markers *Xpsm160* and *Xpsm190*, shown to be located distally on LG 7 in other crosses, but forming a separate linkage group in the LGD 1-B-10 × ICMP 85410 cross (Table 2). When comparing the maps generated in the four crosses, marker orders and distribution are highly similar. The exception is the distal region of LG 1, which is compressed in the LGD 1-B-10 × ICMP 85410 cross relative to the other crosses (see the individual maps on http://jic-bioinfo.bbsrc.ac.uk/cereals/millet.html). The parents of this cross have been shown to differ by a translocation involving LGs 1 and 2 (Liu et al. 1994). In a translocation heterozygote, these linkage groups will form a quadrivalent during meiosis, with the result that markers across both linkage groups will appear genetically linked. The markers have been split into two linkage groups based on information from other crosses. This allows unambiguous ordering of the markers in that region, and the resulting map reflects the altered recombination rate caused by the presence of the translocation.

We have reason to believe that the large gaps in the distal regions of the genetic map indeed represent regions of high recombination, rather than being caused by a general lack of markers. First, none of the newly generated SSR markers mapped to the existing gaps. The few SSRs that mapped distally further extended the genetic maps. For example, addition of the marker *Xpsmp2086* on LG 4 extended the map by 22 and 26 cM in the 81B × ICMP 451 and PT 732B × P1449-2 crosses, respectively. Second, a concerted attempt to isolate markers in the region spanned by the loci *Xpsm464* and *Xpsm866.1* on LG 4 using AFLP markers in combination with a bulked segregant analysis (BSA) approach led to the identification of only one marker, *Xpsm1003*, that mapped in the gap between *Xpsm464* and *Xpsm866.1*. All other markers identified by BSA proved to be tightly linked to either *Xpsm464* (*Xpsm1004* and *Xpsm1005*) or to *Xpsm866.1* (*Xpsm1001*, *Xpsm1007* and *Xpsm1009*)(Fig. 2).
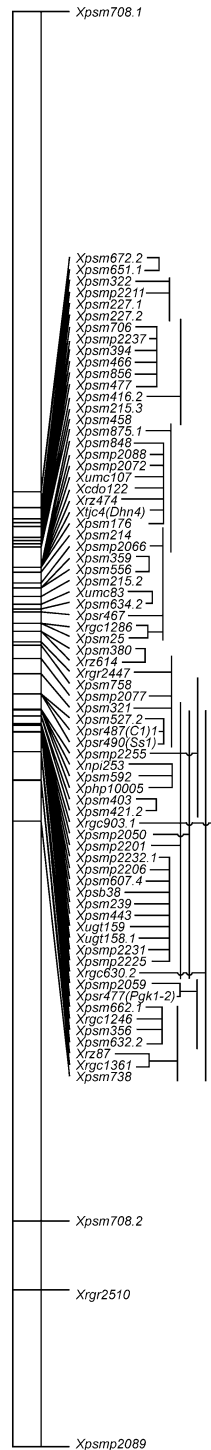
Two linkage groups that are still extremely short are LGs 3 and 5 with genetic lengths of around 30 and 40 cM, respectively. Using the comparative maps with rice as a guide, it would appear that these linkage groups are
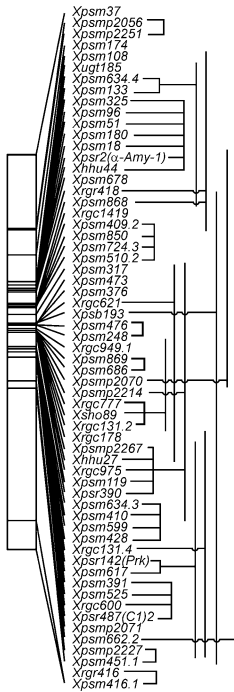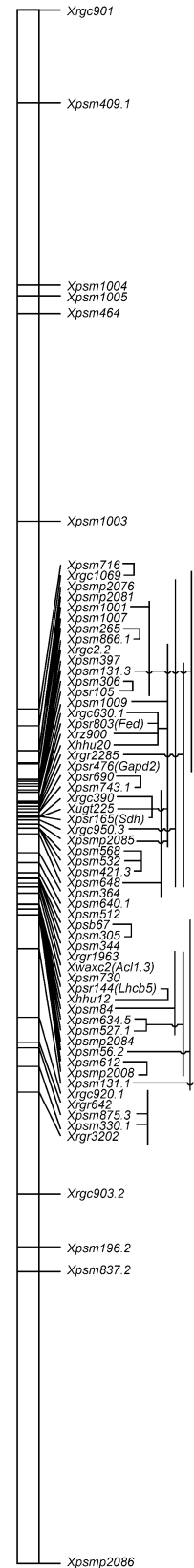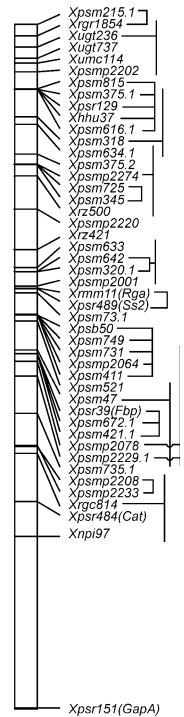
Pearl millet consensus maps

LG 1

Xpsr155(Rbca)
Xugt158.2
Xrmm9(Rga)1
Xpsm513
Xpsm573
Xpsm81.1
Xpsm56.1
Xpsm413
Xpsm280
Xpsm567
Xuom213(Rga)1
Xpsm492
Xrgc2.1
Xpsm858
Xpsm295
Xpsm761
Xrgc950.2
Xpsm515
Xuom213(Rga)3
Xpsm565
Xpsm87.2
Xpsmp2232.2
Xpsm600
Xpsm607.1
Xrgc189
Xpsm757
Xpsm530
Xpsm531
Xpsm768
Xpsmp2069
Xpsmp2273
Xpsm17
Xpsm425
Xak466(Nra1)
Xpsm759
Xpsm756
Xrgc929
Xpsm583
Xpsm73.2
Xpsm671
Xrgc623
Xpsm336
Xrgc385
Xpsm341
Xpsr161(Hsp70)
Xpsm49
Xmsu488(Lec)
Xpsm319
Xpsmp2246
Xrz206.2
Xpsmp2030
Xpsm302
Xpsm891
Xpsm669
Xpsm100
Xpsr168(Hmgp)
Xpsm360
Xpsm854
Xpsm632.1
Xrgr830
Xpsm371
Xpsm386
Xrgc249
Xpsr115(Pk)
Xpsm408
Xpsm223.1
Xpsm652
Xpsm223.2
Xrgc597
Xrz390
Xwia483(Cxp1)
Xpsm52
Xrgr617
Xpsmp2006
Xrgr1534
Xpsm607.5
Xrgc920.2
Xpsm613.1
Xpsm655
Xpsm724.1
Xpsm347.2
Xpsmp2090
Xpsm196.1
Xpsm837.1

LG 2

Xpsm708.1
Xpsm672.2
Xpsm651.1
Xpsm322
Xpsmp2211
Xpsm227.1
Xpsm227.2
Xpsm706
Xpsmp2237
Xpsm394
Xpsm466
Xpsm856
Xpsm477
Xpsm416.2
Xpsm215.3
Xpsm458
Xpsm875.1
Xpsm848
Xpsmp2088
Xpsmp2072
Xumc107
Xcdo122
Xrz474
Xtjc4(Dhn4)
Xpsm176
Xpsm214
Xpsmp2066
Xpsm359
Xpsm556
Xpsm215.2
Xumc83
Xpsm634.2
Xpsr467
Xrgc1286
Xpsm25
Xpsm380
Xrz614
Xrgc2447
Xpsm758
Xpsmp2077
Xpsm321
Xpsm527.2
Xpsr487(C1)1
Xrgc490(Ss1)
Xpsmp2255
Xnpi253
Xpsm592
Xphp10005
Xpsm403
Xpsm421.2
Xrgc903.1
Xpsmp2050
Xpsmp2201
Xpsmp2232.1
Xpsmp2206
Xpsm607.4
Xpsb38
Xpsm239
Xpsm443
Xugt159
Xugt158.1
Xpsmp2231
Xpsmp2225
Xrgc630.2
Xpsmp2059
Xpsr477(Pgk1-2)
Xpsm662.1
Xrgc1246
Xpsm356
Xpsm632.2
Xrz87
Xrgc1361
Xpsm738
Xpsm708.2
Xrgr2510
Xpsmp2089

LG 3

Xpsm37
Xpsmp2056
Xpsmp2251
Xpsm174
Xpsm108
Xugt185
Xpsm634.4
Xpsm133
Xpsm325
Xpsm96
Xpsm51
Xpsm180
Xpsm18
Xpsr2(α-Amy-1)
Xhhu44
Xpsm678
Xrgr418
Xpsm868
Xrgc1419
Xpsm409.2
Xpsm850
Xpsm724.3
Xpsm510.2
Xpsm317
Xpsm473
Xpsm376
Xrgc621
Xpsb193
Xpsm476
Xpsm248
Xrgc949.1
Xpsm869
Xpsm686
Xpsmp2070
Xpsmp2214
Xrgc777
Xsho89
Xrgc131.2
Xrgc178
Xpsmp2267
Xhhu27
Xrgc975
Xpsm119
Xpsr390
Xpsm634.3
Xpsm410
Xpsm599
Xpsm428
Xrgc131.4
Xpsr142(Prk)
Xpsm617
Xpsm391
Xpsm525
Xrgc600
Xpsr487(C1)2
Xpsmp2071
Xpsm662.2
Xpsmp2227
Xpsm451.1
Xrgc416
Xpsm416.1

LG 4

Xrgc901
Xpsm409.1
Xpsm1004
Xpsm1005
Xpsm464
Xpsm1003
Xpsm716
Xrgc1069
Xpsmp2076
Xpsmp2081
Xpsm1001
Xpsm1007
Xpsm265
Xpsm866.1
Xrgc2.2
Xpsm397
Xpsm131.3
Xpsm306
Xpsm105
Xpsm1009
Xrgc630.1
Xpsr803(Fed)
Xrz900
Xhhu20
Xrgr2285
Xpsr476(Gapd2)
Xpsm690
Xrgc743.1
Xrgc390
Xugt225
Xpsr165(Sdh)
Xpsmp2085
Xpsm568
Xpsm532
Xpsm421.3
Xpsm648
Xpsm364
Xpsm640.1
Xpsm512
Xpsb67
Xpsm305
Xpsm344
Xrgr1963
Xwaxc2(Acl1.3)
Xrgc730
Xpsr144(Lhcb5)
Xhhu12
Xpsm84
Xpsm634.5
Xpsm527.1
Xpsmp2084
Xpsm56.2
Xpsm612
Xpsmp2008
Xpsm131.1
Xrgc920.1
Xrgr642
Xpsm875.3
Xpsm330.1
Xrgr3202
Xrgc903.2
Xpsm196.2
Xpsm837.2
Xpsmp2086

LG 5

Xpsm215.1
Xrgr1854
Xugt236
Xugt737
Xumc114
Xpsmp2202
Xpsm815
Xpsm375.1
Xpsr129
Xhhu37
Xpsm616.1
Xpsm318
Xpsm634.1
Xpsm375.2
Xpsm2274
Xpsm725
Xpsm345
Xpsm500
Xpsmp2220
Xrz421
Xpsm633
Xpsm642
Xpsm2001
Xrmm11(Rga)
Xpsr489(Ss2)
Xpsm73.1
Xpsb50
Xpsm749
Xpsm731
Xpsmp2064
Xpsm411
Xpsm521
Xpsm47
Xpsr39(Fbp)
Xpsm672.1
Xpsm421.1
Xpsmp2078
Xpsm2229.1
Xpsm735.1
Xpsmp2208
Xpsmp2233
Xrgc814
Xpsr484(Cat)
Xnpi97
Xpsr151(GapA)

◄ **Fig. 2** Consensus maps of the pearl millet genome based on data from four pearl millet crosses. The individual maps can be viewed on MilletGenes (http://jic-bioinfo.bbsrc.ac.uk/cereals/millet.html). The position range of markers that could not be placed unambiguously is indicated with a *vertical bar*

relatively complete (Devos et al. 2000). The two most distal markers on LG 3 for which comparative information is available, *Xpsr2(α-Amy-1)* and *Xrgr416*, map to the distal regions of rice chromosome arms 2L and 4L, respectively. Similarly, *Xpsr151(GapA)* on LG 5 maps distally on rice 3S. The top part of LG 5 corresponds to rice chromosome arm 4S, but only one rice 4S marker has been mapped on pearl millet LG 5. It is therefore difficult to estimate the coverage of that region of LG 5. The relative completeness of LG 3 and at least the bottom part of LG 5 based on the comparative data may, however, be deceptive. Because of the high recombination rates in the distal pearl millet regions, it is likely that markers that are separated by only a few centimorgans in rice will map tens of centimorgans apart in pearl millet. We therefore predict that markers will eventually be found that map 30 or more centimorgans beyond the current most distal marker on these short pearl millet linkage groups. Mapping telomeric sequences, as has been done in tomato (Ganal et al. 1992) and barley (Röder et al. 1993), will provide a definite answer on the completeness of these linkage groups. Alternatively, the rice genomic sequence can be used to develop new markers that are located distally on pearl millet, provided colinearity is maintained in the distal chromosome regions between rice and pearl millet.

Duplications

Restrictions fragment length polymorphism markers tend to be preselected to be low copy, hence duplications present in the genome will remain largely undetected at the map level. However, there is strong evidence for at least one duplication, involving the distal regions of LGs 1 and 4. The duplication is identified by the markers PSM837, PSM196, RGC903, PSM724, RGC920 and PSM607 (Fig. 2 and unpublished data) and is orthologous to the duplication between the short arms of rice chromosomes 11 and 12 (Devos et al. 2000). The rice 11S/12S duplication has been dated based on amino acid substitutions between paralogous protein pairs to 25 million years ago (Goff et al. 2002), which is after the divergence of the Ehrhartoideae and Panicoideae lineages. Although we cannot exclude the possibility that the pearl millet and rice duplications are independent events, a more likely explanation is that the duplicated genes have evolved more slowly than the values of 9×10^{-10} nonsynonymous substitutions per site per lineage per year and 2.25 nonsynonymous substitutions per amino acid change used for the molecular clock by Goff and coworkers (Goff et al. 2002) or that a degree of gene conversion has taken place between the duplicated regions. A second putative duplication in pearl millet includes markers

PSM81, PSM866, RGC2, RGC950 and PSM515 on LGs 1 and 4 (Devos et al. 2000).

Simple sequence repeat markers

Sixty-three SSR markers, detecting a total of 65 loci, have been mapped (Fig. 2). This includes 35 SSRs from the enriched library described in this paper and 28 SSRs developed from BAC clones (Allouis et al. 2001; Qi et al. 2001). A comparison of the pearl millet SSRs isolated using the two methods shows that SSRs from the enriched library have a relatively higher PIC value (0.71) compared to those from the BAC clones (0.48). This is almost certainly due to the overall higher mean number of repeat units present in the former compared to the latter SSRs (26 units versus 11 units). From a technical point, it is easier to generate SSR markers from enriched libraries than from BAC clones. Nevertheless, although the use of 3' anchored SSR primers in combination with suppression PCR on pooled BAC clones is more cumbersome, this method provided a relatively higher number of useful SSRs. Nearly 50% of the fragments amplified by the suppression PCR could be converted to SSR markers, and 40% of these provided simple patterns and had a PIC value ≥0.40. BAC SSRs have the added value of anchoring individual BACs to the genetic maps, and thus provide a first step in the construction of BAC contigs. In contrast, only around 7% of the enriched library clones yielded SSR markers with a simple amplification pattern and a PIC value ≥0.40. Although this value appears low, the overall efficiency of our methodology was comparable to that obtained in other studies. In cotton, SSR enrichment using biotinylated oligonucleotides and capture with streptavidin-coated paramagnetic beads, followed by high stringency screening of the clones by hybridization with the relevant SSR oligonucleotides resulted in an enrichment of 20% (Connell et al. 1998) versus 25% in our study. Similar enrichment values were obtained by Edwards et al. (1996). The reduction from 25 to 7% of useful SSRs was due to the presence of duplicate clones (15%), small clones with insufficient flanking DNA (29%) and primer amplification problems. Approximately 55% of the designed primer sequences gave clear amplification products. This value is higher than the 32 and 36% success rate of obtaining analyzable products reported by Bryan et al. (1997) and Röder et al. (1995).

In tomato, GT and GA SSRs have been shown to be predominantly associated with the centromeres (Areshchenkova and Ganal 2002). In other species such as rice and wheat, dinucleotide SSRs have been shown to be randomly distributed (McCouch et al. 1997; Röder et al. 1998; Stephenson et al. 1998). In pearl millet, the distribution pattern of SSRs is difficult to assess because of the uneven distribution of recombination. However, the SSRs follow the distribution pattern of the RFLP probes with most loci being located in the centromeric cluster and few loci mapping to the more distal chromosome regions. As a consequence, the available set of SSRs should be
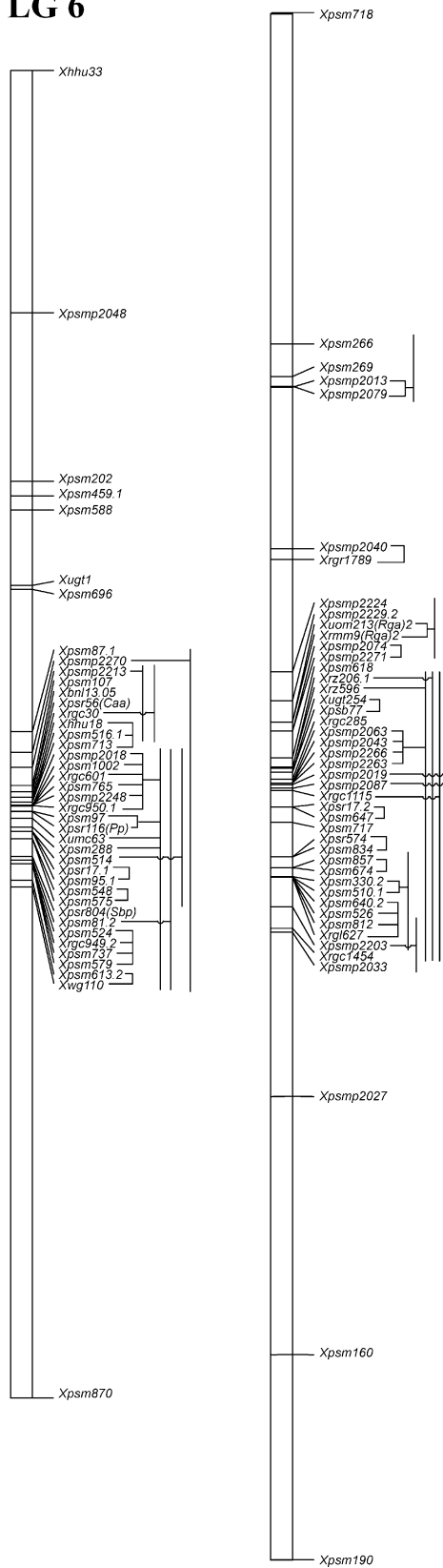
LG 6

LG 7

Fig. 2 (continued)

sufficient to conduct marker-assisted breeding of traits that are located within the regions of relatively low recombination. However, more markers are needed for the distal chromosome regions. The most efficient way to achieve this may be to screen BAC clones, identified with distally located RFLP markers, for the presence of SSR markers. This approach has been successfully used to isolate SSR marker PSMP2090 (Padi 2002). This SSR is present on BAC 356L1, which also carries the RFLP marker PSM196. Both markers detect loci in the distal region of LG 1.

Conclusions

Genetic maps are an important tool for both the pearl millet geneticist and breeder. They form a framework for trait mapping and QTL analysis. The availability of SSR markers will facilitate marker-assisted transfer of the identified QTL to elite germplasm. Furthermore, the incorporation of heterologous RFLP probes into the pearl millet map has allowed alignment with the genetic map of rice. This makes it possible to exploit the rice genomic sequence for targeted marker development for specific regions of the pearl millet genome. Comparative analyses may also aid in the identification of genomic regions or genes involved in resistance to drought and disease, two important breeding targets for pearl millet. The markers also provide the entry points for physical mapping. It is expected that the overall organization of the pearl millet genome, with a size comparable to that of maize, will be similar to that of maize and other large-genome cereals. This needs to be confirmed by sequence analysis of selected regions. Of specific interest is the high recombination rate in the distal chromosome regions. Why are the rates so high, and where precisely does recombination take place? And how does this differ from recombination patterns in orthologous regions in other grass species? The tools are available in pearl millet to answer these questions. The maps thus provide a first step in understanding the genome structure of pearl millet.

References

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1131–1139

Allouis S, Qi X, Lindup S, Gale MD, Devos KM (2001) Construction of a BAC library of pearl millet, *Pennisetum glaucum*. Theor Appl Genet 102:1200–1205

Areshchenkova T, Ganal MW (2002) Comparative analysis of polymorphisms and chromosomal location of tomato microsatellite markers isolated from different sources. Theor Appl Genet 104:229–235

Bryan GJ, Collins AJ, Stephenson P, Orry A, Smith JB, Gale MD (1997) Isolation and characterisation of microsatellites from hexaploid bread wheat. Theor Appl Genet 94:557–563

Chao S, Sharp PJ, Kilian A, Gale MD (1989) RFLP based genetic maps of homologous group 7 chromosomes of wheat and barley. AFRC Institute of Plant Science and John Innes Institute annual report. IPSR, Norwich, pp 23–24

Close TJ, Kortt AA, Chandler PM (1989) A cDNA-based comparison of dehydration-induced proteins (dehyrdins) in barley and corn. Plant Mol Biol 13:95–108

Connell JP, Pammi S, Iqbal MJ, Huizinga T, Reddy AS (1998) A high through-put procedure for capturing microsatellites from complex plant genomes. Plant Mol Biol Rep 16:341–349

Devos KM, Atkinson MD, Chinoy CN, Liu C, Gale MD (1992) RFLP based genetic map of the homologous group 3 chromosomes of wheat and rye. Theor Appl Genet 83:931–939

Devos KM, Pittaway TS, Reynolds A, Gale MD (2000) Comparative mapping reveals a complex relationship between the pearl millet genome and those of foxtail millet and rice. Theor Appl Genet 100:190–198

Edwards KG, Varker GHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques 20:758–760

Ganal MW, Broun P, Tanksley SD (1992) Genetic mapping of tandemly repeated telomeric DNA sequences in tomato (*Lycopersicon esculentum*). Genomics 14:444–448

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun W-L, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296:92–100

Jones ES, Liu CJ, Gale MD, Hash CT, Witcombe JR (1995) Mapping quantitative trait loci for downy mildew resistance in pearl millet. Theor Appl Genet 91:448–456

Jones ES, Breese WA, Liu CJ, Singh SD, Shaw DS, Witcombe JR (2002) Mapping quantitative trait loci for resistance to downy mildew in pearl millet: field and glasshouse screens detect the same QTL. Crop Sci 42:1316–1323

Kijas JMH, Fowler JCS, Garbett CA, Thomas MR (1994) Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. BioTechniques 16:657–662

Liu CJ, Witcombe JR, Pittaway TS, Nash M, Busso CS, Hash CT, Gale MD (1994) An RFLP-based genetic map of pearl millet (*Pennisetum glaucum*). Theor Appl Genet 89:481–487

McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, Huang N, Ishii T, Blair M (1997) Microsatellite marker development, mapping and applications in rice genetics and breeding. Plant Mol Biol 35:89–99

McIntosh RA, Hart GE, Devos KM, Gale MD, Rogers WJ (1998) Catalogue of gene symbols for wheat. In: Slinkard AE (ed) Proceedings of the 9th international wheat genetics symposium. University Extension Press, University of Saskatchewan, Saskatoon, pp 1–325

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3:175–182

Padi FK (2002) Genetic analysis of adaptive traits in pearl millet (*Pennisetum glaucum* (L.). R. Br.). Dissertation, University of East Anglia

Poncet V, Lamy F, Devos KM, Gale MD, Sarr A, Robert T (2000) Genetic control of domestication traits in pearl millet (*Pennisetum glaucum* L., Poaceae). Theor Appl Genet 100:147–159

Poncet V, Martel E, Allouis S, Devos KM, Lamy F, Sarr A, Robert T (2002) Comparative analysis of QTLs affecting domestication traits between two domesticated × wild pearl millet (*Pennisetum glaucum* L., *Poaceae*) crosses. Theor Appl Genet 104:965–975

Qi X, Stam P, Lindhout P (1996) Comparison and integration of four barley genetic maps. Genome 39:379–394

Qi X, Lindup S, Pittaway TS, Allouis S, Gale MD, Devos KM (2001) Development of simple sequence repeat markers from bacterial artificial chromosomes without subcloning. BioTechniques 31:355–362

Röder MS, Lapitan NLV, Sorrells ME, Tanksley SD (1993) Genetic and physical mapping of barley telomeres. Mol Gen Genet 238:294–303

Roder MS, Plashk J, Konig SU, Borner A, Sorrells ME, Tanksley SD, Ganal MW (1995) Abundance, variability and chromosomal location of microsatellites in wheat. Mol Gen Genet 246:327–333

Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier M-H, Leroy P, Ganal MW (1998) A microsatellite map of wheat. Genetics 149:2007–2023

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J 3:739–744

Stephenson P, Bryan GJ, Kirby J, Collins AJ, Devos KM, Busso CS, Gale MD (1998) Fifty new microsatellite loci for the wheat genetic map. Theor Appl Genet 97:946–949

Tanksley SD, Ganal MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Govannoni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Röder MS, Wing RA, Wu W, Young ND (1992) High-density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160

Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res 17:6463–6471

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 technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wu J, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet 241:225–235

Yadav RS, Hash CT, Bidinger FR, Cavan GP, Howarth CJ (2002) Quantitative trait loci associated with traits determining grain and stover yield in pearl millet under terminal drought stress conditions. Theor Appl Genet 104:67–83

Yadav RS, Bidinger FR, Hash CT, Yadav YP, Yadav OP, Bhatnagar SK, Howarth CJ (2003) Mapping and characterization of QTL × E interactions for traits determining grain and stover yield in pearl millet. Theor Appl Genet 106:512–520

Yadav RS, Hash CT, Bidinger FR, Devos KM, Howarth CJ, Skot K (2004) Genomic regions associated with grain yield and aspects of post-flowering drought tolerance in pearl millet across stress environments and testers background. Euphytica 136:265–277