Theor Appl Genet (1999) 99:464–472

M. Sun · H. Chen · F.C. Leung

Low-Cot DNA sequences for fingerprinting analysis of germplasm diversity and relationships in *Amaranthus*

Received: 19 October 1998 / Accepted: 8 January 1999

Abstract We examined genetic diversity and relationships among 24 cultivated and wild *Amaranthus* accessions using the total low-Cot DNA and five individual repetitive sequences as probes. These low-Cot DNA probes were obtained by the isolation of various classes of repetitive-DNA sequences, including satellites, minisatellites, microsatellites, rDNA, retrotransposon-like sequences, and other unidentified novel repetitive sequences. DNA fingerprints generated by different types of repetitive-DNA probes revealed different levels of polymorphism in the *Amaranthus* genomes. A repetitive sequence containing microsatellites was found to be a suitable probe for characterizing intraspecific accessions, whereas more conservative sequences (e.g. rDNA) were informative for resolving phylogenetic relationships among distantly related species. Genetic diversity, measured as restriction fragment length polymorphism (RFLP) and the similarity index at the low-Cot DNA level, was equally high among intraspecific accessions between the two species groups: grain amaranths (*A. caudatus*, *A. cruentus*, and *A. hypochondriacus*) and their putative wild progenitors (*A. hybridus*, *A. powellii*, and *A. quitensis*). At the interspecific level, however, the grain amaranth species are less divergent from each other than their wild progenitors. With the rare exceptions of certain *A. caudatus* accessions, grain amaranths were found to be closely related to *A. hybridus*. The results based on low-Cot DNA were comparable with previous RAPD and isozyme studies of the same set of species/accessions of *Amaranthus*, indicating that low-Cot DNA sequences are suitable probes for a fingerprinting analysis of plant germplasm diversity and for determining phylogenetic relationships.

Communicated by G. Wenzel

Dr. Mei Sun (✉) · H. Chen · F.C. Leung
Department of Zoology, The University of Hong Kong,
Pokfulam Road, Hong Kong
e-mail: meisun@hkucc.hku.hk
Fax: +852 2818 8136

Key words *Amaranthus* · Low-Cot DNA · Genetic diversity · DNA fingerprinting · Phylogenetic relationships

Introduction

Amaranthus is a genus of about 60 species, including both cultivated and wild forms. The genus has been the focus of research interest because it contains several species with high potential as future food, feed, or pigment crops. The three cultivated grain amaranths, *A. caudatus*, *A. cruentus* and *A. hypochondriacus*, are among the earliest New World grain crop domesticates and have been cultivated in Mexico, Central America, and the Andean highlands of South America for the past several thousand years (Sauer 1950, 1967, 1976). Sauer's studies suggested that the widely distributed weedy species *A. hybridus* was the common progenitor of the three grain species, with two other wild species *A. quitensis* and *A. powellii* also involved in the speciation of *A. caudatus* and *A. hypochondriacus*, respectively, through natural hybridization and introgression. An alternative hypothesis was that each of the three grain species was domesticated separately in different parts of the Americas, and that *A. cruentus*, *A. caudatus* and *A. hypochondriacus* evolved independently from *A. hybridus*, *A. quitensis* and *A. powellii*, respectively.

Previous molecular studies of the relationships among amaranth species produced varying results (e.g. Hauptli and Jain 1984; Transue et al. 1994; Lanoue et al. 1996; Chan and Sun 1997). To-date, the evolutionary origin of grain amaranths has not been unequivocally established. In addition to allozymes, RAPDs, and RFLPs of PCR-amplified chloroplast and nuclear DNA which have been employed in such studies, new genetic marker systems need to be developed to facilitate further investigations.

Various types of repetitive DNA exist in the genomes of eukaryotes, including several classes of tandemly repeated sequences (such as highly repetitive satellite DNA, and moderately repetitive minisatellite and micro-

satellite sequences), transposons and retrotransposons (moderately repetitive sequences), and rRNA genes (moderately repetitive sequences). In higher plants, the proportion of repetitive DNA can be as much as 80% of the genome (Nei 1987). High levels of polymorphism at some of the repetitive DNA loci, such as minisatellites and microsatellites, have frequently been utilized for fingerprinting individuals and cultivars, testing paternity, mapping genes, investigating population genetic structure, and addressing conservation issues (see examples and reviews in Jeffreys et al. 1985a, b; Nakamura et al. 1987; Burke et al. 1991; Antonius and Nybom 1994; Gray 1995; Akagi et al. 1997). However, DNA fingerprints have only occasionally been used for uncovering phylogenetic relationships at the interspecific level or above, largely because these loci are thought to evolve too rapidly to be phylogenetically informative over large time intervals (but see Aggarwal et al. 1994).

We have adopted a novel approach, named the low-Cot DNA probe method, to isolate various classes of repetitive DNA in *Amaranthus* species. Low-Cot DNA utilizes repetitive DNA reannealing kinetics, which has a faster rate of reannealing than non-repetitive DNA. The Cot value is determined by the starting concentration of single-stranded DNA (C_0) and the reassociation reaction time (t). Highly to moderately repetitive DNA sequences reanneal faster and thus have lower Cot values than mildly repetitive or single-copy DNA sequences. The low-Cot DNA includes microsatellites, minisatellites, ribosomal RNA genes (rDNA), interspersed retrotransposons or retrotransposon-like sequences, and other unidentified 'junk DNA' in *Amaranthus tricolor* (unpublished data). These repetitive DNA sequences could be developed as new classes of molecular markers suitable for simultaneously assessing both intra- and inter-specific genetic diversity and relationships. Such information is important for efficient utilization of the available plant germplasm resources, such as helping plant breeders to select suitable material for developing new cultivars.

The objectives of the present study are to: (1) assess genetic diversity and relatedness among *Amaranthus* accessions at both intra- and inter-specific levels using low-Cot DNA and individual repetitive DNA sequences as fingerprinting probes; (2) further investigate the evolutionary origin of the three most important grain species, *A. caudatus*, *A. cruentus* and *A. hypochondriacus*, with emphasis on their relationships to each other and to their putative wild progenitors *A. hybridus*, *A. powellii* and *A. quitensis*; and (3) evaluate the applicability and effectiveness of various classes of repetitive DNA as molecular markers, in comparison to allozymes and RAPDs, in population-genetic and phylogenetic investigations.

Materials and methods

Seeds of 24 accessions of *Amaranthus*, representing three grain crop species and their three putative progenitors, were obtained from the USDA/ARS Plant Introduction Station at Ames, Iowa, USA. The same four accessions of each species that had been previously studied using isozyme and RAPD markers (Chan and Sun 1997) were employed in the present study for both intra- and inter-specific comparisons. These species (accessions) are: *A. caudatus* (PI 51167, PI 568132, PI 166045, and PI 553073), *A. cruentus* (Ames 5142, PI 566897, PI 566896, and Ames 5369), *A. hybridus* (Ames 5331, Ames 5684, Ames 2026, and Ames 14358), *A. hypochondriacus* (PI 477915, PI 477916, Ames 2178, and PI 540446), *A. powellii* (PI 572260, PI 572261, PI 538793, and PI 572257), and *A. quitensis* (PI 511734, PI 568154, PI 511745 and PI 511733). One accession of *A. tricolor* was collected from Hong Kong to serve as an outgroup source of DNA probes.

Seeds were germinated to provide leaf and root material for DNA extraction. Fresh samples of leaves and roots were first frozen and stored in liquid nitrogen, and genomic DNA extraction was then performed following the CTAB method (Doyle 1991). Four grams of fresh leaves were used from each of the 24 accessions for DNA extraction, yielding about 150 µg of DNA per accession.

Preparation of low-Cot DNA

Approximately 700 µg of DNA was obtained from 12 g of fresh leaves of *A. tricolor* for the preparation of low-Cot DNA. The genomic DNA from *A. tricolor* was sheared to 0.2–1.0-kb-long fragments by an ultrasonic apparatus (Sonifier 250, VWR Company) on ice. About 0.6 mg of the sheared DNA was dissolved with 1 ml of ddH_2O and denatured in boiling water for 15 min followed by chilling on ice for 5 min. NaCl was added to the denatured DNA to a final concentration of 0.18 M. The denatured DNA was allowed to re-anneal at 60°C for 4 h, and adjusted with phosphate buffer to a final concentration of 0.12 M phosphate. The double-stranded DNA was separated from single-stranded DNA with a hydroxylapatite column (BioRad) by adding the DNA sample to a pre-warmed (60°C) hydroxylapatite suspension (20 mg/ml in 0.12 M phosphate buffer, pH 7.4). The solution was gently shaken for 5 min and left to settle for 5 min at 60°C. The hydroxylapatite column was washed twice with pre-warmed (60°C) 0.12 M phosphate buffer to remove single-stranded DNA. To elute double-stranded DNA from the hydroxylapatite column, a pre-warmed (60°C) 0.4 M phosphate buffer was used to wash the column in a shaker at 60°C for 5 min. The column was then left in the same buffer for a further 5 min. The eluted double-stranded DNA was collected as the total low-Cot DNA.

Genomic DNA library construction

Two grams of fresh roots of *A. tricolor* was used for DNA extraction, yielding about 33 µg of DNA for constructing a genomic library. The genomic DNA was digested with *Sau*3AI and fragments separated on 1% agarose gels in 1×TAE buffer. The fragments between 0.3–1.6 kb were excised and purified using a Geneclean II kit (Bio 101, Inc.), and ligated to the *Bam*HI site of plasmid pBluescript SK (+) (Stratagene). The ligation mix was transformed into *Escherichia coli* (XL1 blue cells).

Screening the genomic DNA library and sequencing positive clones

The genomic library was screened with the total low-Cot DNA as a multilocus probe**,** which was labelled using the digoxigenin (DIG) non-radioactive nucleic-acid labelling and detection system (Boehringer Mannheim) following the manufacturer's protocol. Positive clones were grown overnight in LB medium, and the

plasmid DNA was extracted by the alkaline-lysis procedure. Seventy of the positive clones were selected for DNA sequence analysis. The inserts were each PCR-amplified and sequenced following the protocol described in the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer) using an automatic sequencer (ABI PRISM 310 Genetic Analyzer). The sequence homology was analyzed with the DNASIS and BLAST programs.

Southern-blot analysis

Approximately 8 µg of genomic DNA from each of the 24 accessions of *Amaranthus* (see Table 1) was digested with 32 units each of the restriction endonucleases *Eco*RI, *Bam*HI, *Dra*I, *Hin*fI, *Rsa*I, *Hae*III, and *Alu*I. The DNA fragments were separated on 1.0% agarose gels at 50 V for 17–18 h, and then transferred onto nylon membranes (Amersham Life Science Co.) following standard procedures.

Preparation of low-Cot DNA probes

The DNA inserts of selected clones that contained various types of repetitive sequences were individually labelled with DIG-dUTP using the PCR DIG Probe Synthesis Kit (Boehringer Mannheim), following the protocol supplied by the manufacturer. In addition to labelling individual low-Cot DNA clones, the total low-Cot DNA was also labelled as a multilocus probe using the DIG DNA Labelling Kit (Boehringer Mannheim), following the protocol supplied by the same manufacturer.

Hybridization and band detection

Hybridization of the membrane with individual inserts or the total low-Cot DNA probe was carried out at 60°C overnight. Membranes were washed twice for 15 min each with 2 × SSC and 0.1% SDS at room temperature, followed by two more washes with 0.5 × SSC and 0.1% SDS at 60°C for 15 min each. A non-radioactive chemiluminescent detection of DIG labelled-probes was performed with CDP-*Star*™ (Boehringer Mannheim Co.) following the detection procedure as described by the manufacturer. The membranes were exposed to X-ray films (Fuji MI-NPC 30) for 5–15 min at room temperature.

Data analysis

Restriction fragments of the DNA fingerprints ranged from 0.3 to 5.0 kb in size. In the data matrix, the presence of a band was coded as 1, whereas the absence of the same band was coded as 0. The similarities of DNA profiles between pairs of intra- and interspecific accessions were analyzed using the band-sharing coefficient (or similarity index) by the SIMQUAL program of NTSYS-pc (Version 1.8; Rohlf 1994). The Dice (1945) coefficient was estimated as $D=2a/(2a + b + c)$, where a is the number of positive matches (i.e. the presence of a band in both accessions), and $b + c$ is the total number of "unmatches" (i.e. the presence or absence of a band in one accession but not in the other). This estimate is equivalent to Nei and Li's (1979) band-sharing coefficient estimated by $F=2N_{XY} / (N_X + N_Y)$, where N_X and N_Y are the numbers of restriction fragments observed in accessions X and Y, respectively, whereas N_{XY} is the number of fragments shared by both accessions. The DNA fingerprints revealed by each of the single low-Cot probes were separately recorded and analyzed, and also combined to form a single data matrix for comparison with that generated by the total low-Cot DNA probe. Dendrograms were produced from the resultant similarity indices using the UPGMA (unweighted pair group method using arithmetic averages) clustering method.

Results

The total genomic DNA library consisted of about 4000 clones. Approximately 800 colonies were screened by hybridization with the total low-Cot DNA as a multilocus probe. Seventy positive clones containing repetitive sequences were obtained and the DNA inserts sequenced. Of the 22 clones tested for their suitability as fingerprinting probes, five containing respectively satellite DNA (*chy16*), micro-minisatellite (*c16*), a GC-rich repeat (*chy10*), and ribosomal DNA sequences (*chy14*

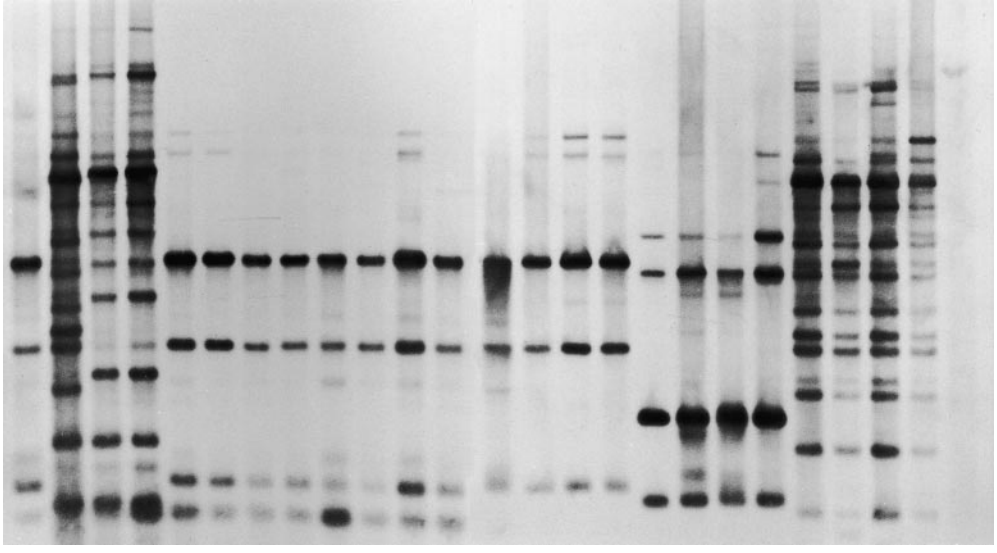

Fig. 1 Fingerprints of 24 accessions of *Amaranthus* species with the probe *chy16* and *Hinf* I restriction-enzyme digestion. *Lanes 1–4* four accessions of *A. caudatus* (PI 51167, PI 568132, PI 166045 and PI 553073); *lanes 5–8* four accessions of *A. cruentus* (Ames 5142, PI 566897, PI566896 and Ames 5369); *lanes 9–12* four accessions of *A. hybridus* (Ames 5331, Ames 5684, Ames 2026 and Ames 14358); *lanes 13–16* four accessions of *A. hypochondriacus* (PI 477915, PI 477916, Ames 2178 and PI 540446); *lanes 17–20* four accessions of *A. powellii* (PI 572260, PI 572261, PI 538793 and PI 572257); and *lanes 21–24* four accessions of *A. quitensis* (PI 511734, PI 568154, PI 511745 and PI 511733)

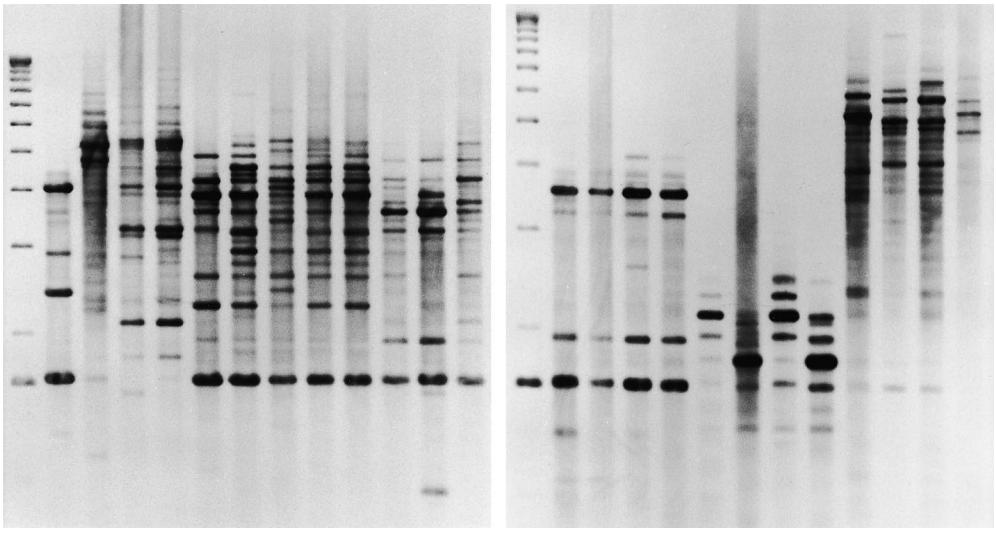

Fig. 2 Fingerprints of 24 accessions of *Amaranthus* species with the probe *c16* and *Alu*I restriction-enzyme digestion. *Lanes 1 and 14* DNA marker (396, 506, 1018, 1636, 2036, 3054, 4072 and 5090 bp); *lanes 2–5* four accessions of *A. caudatus* (PI 51167, PI 568132, PI 166045 and PI 553073); *lanes 6–9* four accessions of *A. cruentus* (Ames 5142, PI 566897, PI566896 and Ames 5369); *lanes 10–13* four accessions of *A. hybridus* (Ames 5331, Ames 5684, Ames 2026 and Ames 14358); *lanes 15–18* four accessions of *A. hypochondriacus* (PI 477915, PI 477916, Ames 2178 and PI 540446); *lanes 19–22* four accessions of *A. powellii* (PI 572260, PI 572261, PI 538793 and PI 572257); and *lanes 23–26* four accessions of *A. quitensis* (PI 511734, PI 568154, PI 511745 and PI 511733)

and *chy19*) were used as individual probes for fingerprinting analysis of the *Amaranthus* species/accessions, in addition to the use of total low-Cot DNA as a multisequence / multilocus probe.

Single repetitive-sequence probes

Each of the five clones was labelled with digoxigenin and used as individual probes to fingerprint the genomic DNA of 24 inter- and intra-specific accessions of *Amaranthus*. All the individual probes detected multiple bands with each of the restriction enzymes in the accessions surveyed, suggesting that these repetitive sequences are all multilocus probes, having multiple copies of the same or similar sequences distributed in the genome (Figs. 1 and 2). The DNA fingerprints of accessions exhibited probe and restriction enzyme-specific patterns, with a band size commonly ranging from 0.3 to 10 kb.

Total low-Cot DNA probe

The total low-Cot DNA contains different classes of repetitive sequences. When the low-Cot DNA was labelled and used as probe, it produced multilocus fingerprints. Multiple bands were observed with each restriction-enzyme digestion of the total genomic DNA (Fig. 3). The multiple repetitive sequences in the total low-Cot DNA include both highly repetitive satellite sequences, and moderately repetitive sequences, such as microsatellites, minisatellites, and chloroplast and nuclear rDNA, with various degrees of polymorphism. A higher proportion of monomorphic loci in each species was detected than those detected with single low-Cot DNA probes. However, the efficiency of the total low-Cot DNA is satisfactory given that, with a single restriction-enzyme digestion, nearly all individual accessions can be distinguished based on their unique multilocus DNA fingerprints.

Genetic diversity within and between species

The proportion of polymorphic restriction fragments detected with each probe within and between crop and wild species is given in Table 1. The total number of fragments detected in these species varies with the probe and the number of restriction enzymes employed. The highest polymorphism at the intraspecific level (averaged 75% over the six species studied), was detected using probe *c16*, a repetitive sequence containing microsatellite and minisatellite sequences. In comparison with this probe, the two nuclear rDNA probes, *chy14* and *chy19*, detected much less polymorphism within species (0–32%). Probe *chy16*, a satellite-containing sequence, revealed an intermediate level of polymorphism between the micro-minisatellite and rDNA probes at the intraspecific level. However, very high levels of polymorphism were revealed with all probes, ranging from 71 to 100%. Polymorphism at the intraspecific level detected by the five single low-Cot DNA probes combined was two-times higher than that detected by the total low-Cot DNA probe, indicating that the total low-Cot DNA includes many conservative sequences. At the interspecific level, however, the difference in the level of polymorphism detected was much less between the five probes combined and the total low-Cot DNA probe, suggesting that most of the conservative sequences contained in the total low-Cot DNA are informative for interspecific comparisons.

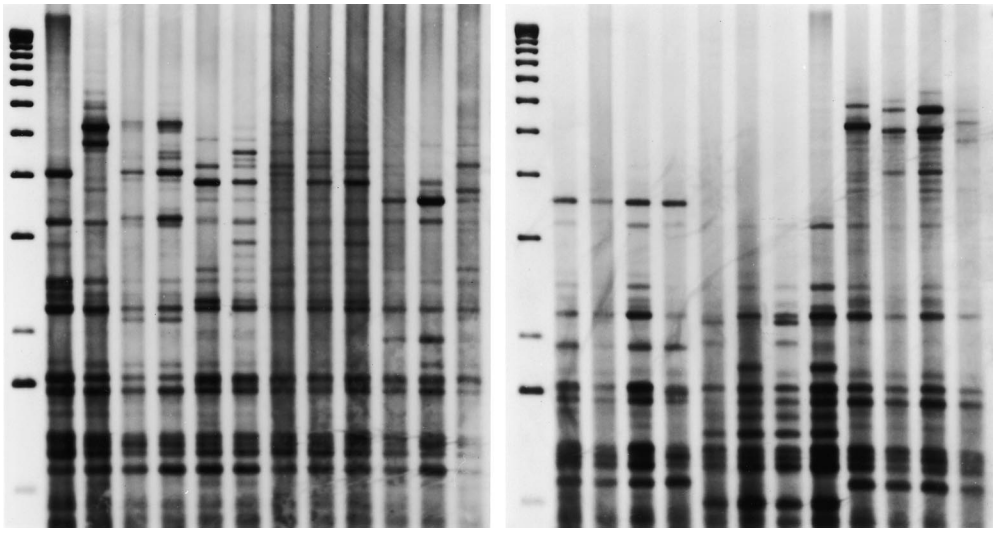

Fig. 3 Fingerprints of 24 accessions of *Amaranthus* species with the total low-Cot DNA probe and *Alu*I restriction enzyme digestion. *Lanes 1 and 14* DNA markers (396, 506, 1018, 1636, 2036, 3054, 4072 and 5090 bp); *lanes 2–5* four accessions of *A. caudatus* (PI 51167, PI 568132, PI 166045 and PI 553073); *lanes 6–9* four accessions of *A. cruentus* (Ames 5142, PI 566897, PI566896 and Ames 5369); *lanes 10–13* four accessions of *A. hybridus* (Ames 5331, Ames 5684, Ames 2026 and Ames 14358); *lanes 15–18* four accessions of *A. hypochondriacus* (PI 477915, PI 477916, Ames 2178 and PI 540446); *lanes 19–22* four accessions of *A. powellii* (PI 572260, PI 572261, PI 538793 and PI 572257); and *lanes 23–26* four accessions of *A. quitensis* (PI 511734, PI 568154, PI 511745, and PI 511733)

The average similarity index between accessions

The band-sharing coefficient or similarity index is an alternative measure of genetic diversity within and between species. A high average similarity index indicates low diversity. The extent of band-sharing varied widely between intra- and inter-specific accessions, ranging from 0 to 100% (Table 2). A complete lack of band-sharing existed between some of the inter-specific accessions, whereas 100% band-sharing was occasionally found between intraspecific accessions. The average similarity index between intraspecific accessions detected with probes *c16*, *chy10*, *chy16*, and *chy14* is given in Table 2. The most variable sequence, such as the micro-minisatellite probe *c16*, revealed the lowest average similarity or the highest diversity within species; whereas *chy14*, a probe containing 5 S rDNA sequence, detected the lowest diversity within species but sufficient divergence between species (matrices of pairwise similarity index between both intra- and inter-specific accessions are available upon request).

Phylogenetic relationships

Phylogenetic relationships within and between species were constructed based on the similarity index matrices

Table 1 RFLP polymorphism among intra- and inter-specific accessions of grain amaranths and their wild progenitors. Values under each species represent the proportion of polymorphic fragments (%) and the number of fragments detected with each probe (in parenthesis)

Probe	Type of Sequence	Grain amaranths				Wild progenitors			
		A. caudatus	*A. cruentus*	*A. hypochondriacus*	Interspecific accessions	*A. hybridus*	*A. powellii*	*A. quitensis*	Interspecific accessions
c16	Micro-minisatellite	96.1 (51)	61.7 (47)	53.6 (28)	98.8 (83)	77.6 (49)	77.4 (31)	81.0 (42)	100 (92)
chy10	GC-rich repeat	52.4 (42)	14.3 (35)	3.3 (30)	70.8 (48)	27.0 (37)	25.0 (28)	35.6 (45)	79.4 (63)
chy16	Satellite	86.0 (50)	40.7 (7)	5.6 (18)	87.1 (62)	50.0 (28)	38.9 (18)	40.4 (47)	97.3 (74)
chy14	rDNA	30.8 (13)	100 (10)	50.0 (14)	78.9 (19)	10.0 (10)	42.8 (7)	46.7 (15)	100 (22)
chy19	rDNA	0 (3)	0 (3)	0 (3)	0 (3)	0 (3)	0 (3)	0 (3)	100 (6)
All five probes combined		74.5 (161)	38.0 (121)	26.4 (91)	76.9 (221)	47.7 (132)	48.8 (84)	53.8 (160)	93.4 (256)
Total low-Cot DNA probe		41.6 (77)	20.3 (69)	8.1 (62)	58.8 (107)	23.5 (68)	24.6 (69)	23.9 (71)	58.9 (112)

Table 2 Mean similarity index between intraspecific accessions of grain amaranths and their wild progenitors; *n* is the total number of restriction fragments detected by each probe in the six species

Probe	*n*	Grain amaranths			Wild progenitors			Probe mean	Range[a]
		A. caudatus	*A. cruentus*	*A. hypochondriacus*	*A. hybridus*	*A. powellii*	*A. quitensis*		
c16 (micro-minisatellite)	106	0.441 ± 0.267	0.746 ± 0.082	0.774 ± 0.072	0.638 ± 0.165	0.584 ± 0.268	0.574 ± 0.155	0.626 ± 0.123	0.000–0.902
chy10 (GC-rich repeat)	68	0.795 ± 0.094	0.956 ± 0.023	0.992 ± 0.009	0.916 ± 0.046	0.907 ± 0.065	0.868 ± 0.053	0.906 ± 0.069	0.485–1.000
chy16 (satellite)	92	0.519 ± 0.229	0.862 ± 0.069	0.986 ± 0.016	0.832 ± 0.072	0.839 ± 0.124	0.837 ± 0.059	0.813 ± 0.155	0.121–1.000
chy14 (rDNA)	27	0.912 ± 0.038	0.973 ± 0.021	0.827 ± 0.134	0.980 ± 0.022	0.912 ± 0.096	0.852 ± 0.094	0.909 ± 0.062	0.000–1.000
All individual probes combined[b]	299	0.645 ± 0.154	0.869 ± 0.047	0.908 ± 0.016	0.818 ± 0.086	0.804 ± 0.127	0.793 ± 0.069	0.806 ± 0.090	0.223–0.928
Total low-Cot DNA probe	137	0.843 ± 0.059	0.939 ± 0.026	0.973 ± 0.018	0.918 ± 0.037	0.909 ± 0.065	0.913 ± 0.040	0.916 ± 0.043	0.667–1.000

[a] Including the similarity index between interspecific accessions. The low range represents the bottom value between interspecific accessions, and the high range represents the top value between intraspecific accessions

[b] The similarity index between accessions was not estimated for fingerprints with the *chy19* probe itself, but the data was included in the analysis of fingerprints by all five individual probes

with all five individual probes combined (Fig. 4) and with the total low-Cot DNA probe (Fig. 5). A strict consensus tree of the two phenograms (data not shown) showed a good fit of the two data sets, indicating similar relationships among the species/accessions. For comparisons between different probes and between different molecular markers, phenograms (data not shown) were also computed based on the similarity index for each individual probe (this study) and for RAPD data (Chan and Sun 1997). It was found that the micro-minisatellite probe, *c16*, which revealed the highest level of polymorphism, was most suitable for characterizing intraspecific accessions, whereas the rDNA probe, *chy14*, which detected the lowest diversity within species as well as between closely related species, was satisfactory for resolving relationships among more distantly related taxa.

Discussion

Repetitive DNA sequences are ubiquitous in eukaryotic genomes (Dover and Flavell 1982; Cavalier-Smith 1985; Nei 1987; John and Miklos 1988). Due to their varied rate of mutation, they can be very valuable sources of genetic marker systems, with potential applicability for population studies, gene mapping, and for assessing phylogenetic relationships at various taxonomic levels.

A major advantage of the low-Cot DNA fingerprinting approach is that many different types of repetitive DNA could be obtained from total genomic DNA in a single isolation of the fast-reannealing fraction of denatured sequences. For example, of the 70 clones sequenced, 12 were identified as homologous to nuclear rDNA genes; three contained satellite sequences; ten contained minisatellites, microsatellites and GC- or AT-rich repeats; two contained retrotransposon-like sequences; two were homologous to genes coding for the cytochrome oxidase subunit and pckA; 14 contained various chloroplast DNA sequences, including genes coding for rRNA, RNA polymerase, ATPase, and rbcL; while the remaining 22 clones contained unidentified repetitive sequences (unpublished data). It is not surprising, however, to pick up a large proportion of chloroplast DNA sequences by the low-Cot DNA method, because of the presence of multiple copies of the chloroplast genome in total genomic DNA. Given the highly conservative nature of chloroplast DNA (Wolfe et al. 1987), this class of slowly evolving sequences in the total low-Cot DNA is valuable for constructing phylogenetic relationships at interspecific or higher taxonomic levels, whereas the fast-evolving nuclear repetitive DNA sequences, such as satellites, micro- or mini-satellites, GC- or AT-rich repeats and retrotransposon-like sequences, would be suitable for resolving relationships within a species or among closely related species.

As shown in the present study, different types of repetitive DNA clones revealed different levels of polymorphism in the *Amaranthus* genomes, and the level of polymorphism detected by the total low-Cot DNA probe

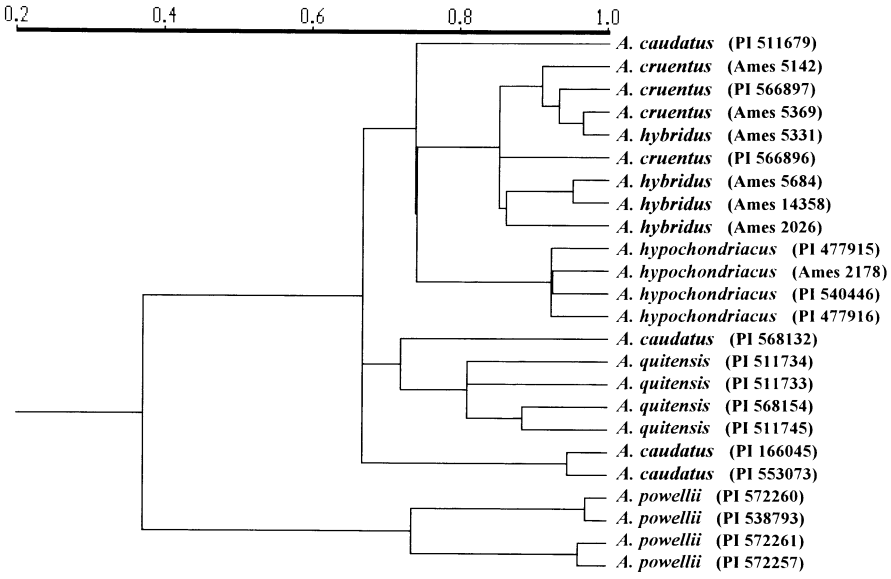

Fig. 4 UPGMA phenogram computed from the similarity matrix based on the combined DNA fingerprints using all five individual low-Cot DNA probes, *chy10*, *chy14*, *chy16*, *chy19* and *c16*

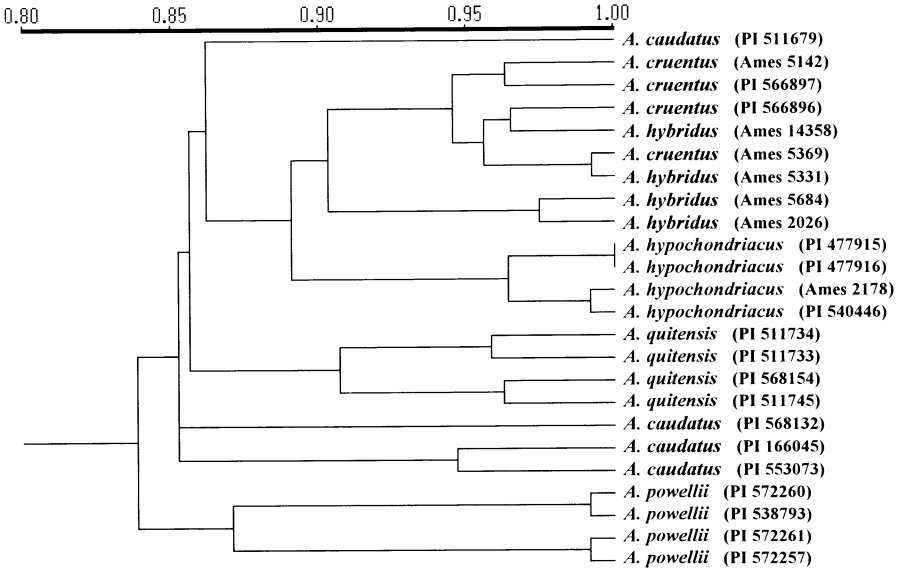

Fig. 5 UPGMA phenogram computed from the similarity matrix based on the total low-Cot DNA fingerprint data

was much lower than that detected by the single probes combined (Table 1). Ribosomal RNA genes represent a class of moderately repetitive DNA organized in tandem duplication with non-transcribed spacers. Amounts of these genes vary from several hundreds of copies up to 10% of the genome in some plants. Various rDNA sequences have been successfully used to estimate phylogeny at different levels (see reviews in Clegg and Durbin 1990; Soltis et al. 1997). The rDNA probes employed in the present study, *chy14* and *chy19*, revealed much lower levels of polymorphism than the other three probes, *c16*, *chy10* and *chy16*, suggesting the conservative nature of these rDNA sequences. The difference in the level of polymorphism detected by the individual probes and the total low-Cot DNA probe indicates that many of the additional repetitive sequences that are present in the total low-Cot DNA are more conservative than most of the single probes used; thus, many more monomorphic loci could be included in the sampling of the genome. This property of the total low-Cot DNA probe should be further explored for comparative genetic diversity assessment.

Also, because the total low-Cot DNA probe consists of many different types of repetitive sequences, the highly repetitive sequences would become the dominant probes to affect the fingerprint profile, whereas the low-copy number sequences may fail to be detected, resulting in differences between the total low-Cot DNA fingerprints and those generated by individual probes. Despite the potential sampling bias due to differences in copy number among the repetitive sequences, the relationships revealed by the total low-Cot DNA probe (Fig. 5) were very similar to those revealed by the five individual probes combined (Fig. 4). As the total low-Cot DNA

probe contains a range of repetitive DNA sequences, from rapidly to slowly evolving, it should be more suitable than the single-sequence probes for simultaneously investigating phylogenetic relationships at both intra- and inter-specific levels. In addition, the efficiency of using the total low-Cot DNA as a probe is evident, as no genomic cloning steps or sequencing information is required when the total low-Cot DNA is developed and used as a probe, in comparison to the use of individual repetitive sequences as probes.

Information on genetic diversity and relationships within and among crop species and their wild relatives is essential for the efficient utilization of plant genetic-resource collections. Among several molecular approaches that have been employed in assessing genetic diversity and relationships, isozyme and RAPD (random amplified polymorphic DNA, Williams et al. 1990) data can be generated faster and with less labor than other methods such as microsatellites and DNA fingerprints. However, the number of isozyme loci that can be detected is limited, and often insufficient allozyme polymorphism restricts its application as a genetic marker system. Although polymorphism at RAPD loci is usually high, the dominant nature of RAPD markers and the low homologies between co-migrating bands from congeneric species have limited their application in population and phylogenetic studies (Lynch and Milligan 1994; Thormann et al. 1994; Rieseberg 1996). RAPD markers are highly effective in clustering intraspecific accessions of *Amaranthus*, but cannot resolve relationships among the species (Chan and Sun 1997). More recently, microsatellites have become increasingly important in genome mapping, paternity and kinship analysis, and for studying genetic structure of natural populations. There are several advantages using this marker system in such studies. Microsatellites are highly polymorphic DNA markers with discrete loci and co-dominant alleles. For estimating phylogeny among closely related taxa or intraspecific accessions, rapidly evolving characters such as microsatellites would be the markers of choice. However, similar to RAPDS, the information generated by microsatellites may not be accurate for the reconstruction of phylogeny among distantly related species.

Thus, for germplasm characterization and phylogenetic studies of crops and their wild relatives, a complementary approach should be developed for simultaneously generating both highly polymorphic markers for fingerprinting intraspecific cultivars/accessions and the less-variable markers suitable for studying phylogenetic relationships at higher taxonomic levels. As demonstrated in the present study, a single isolation and application of the total low-Cot DNA as a multilocus probe may meet this contrasting need. The efficiency of the low-Cot DNA approach is apparent in providing a large number of bands/loci across the genome. The total number of bands/loci that can be generated is directly related to the number of restriction enzymes employed. A large sample size of loci (e.g. >200) can be easily achieved with about ten restriction enzymes. In addition, the information extracted from DNA-fingerprint analyses can be related to conventional population genetic parameters (Lynch 1991). The techniques, limitations and advantages of DNA fingerprinting have already been well reviewed (Burke et al. 1991; Bruford et al. 1998).

Nearly all dendrograms based on the low-Cot DNA data indicated a close relationship between *A. cruentus* and *A. hybridus*. With the exception of some accessions of *A. caudatus*, the three crop species usually clustered close to *A. hybridus*, with *A. quitensis* clustered next closest to the group. This general relationship is consistent with that obtained based on isozyme data (Chan and Sun 1997) as in these *Amaranthus* species/accessions, sufficient allozyme polymorphism existed allowing for both intra- and inter-specific applications. However, due to their divergent fingerprints (e.g. Figs. 1 and 2), the four accessions of *A. caudatus* were not well clustered based on the low-Cot DNA data. One of the *A. caudatus* accessions (*PI 568132*) is distinctive from the rest of accessions in most of the dendrograms, but occasionally grouped with the *A. quitensis* accessions (Fig. 4). Accession *PI 511679* of *A. caudatus* was frequently clustered with *A. hypochondriacus* or with other grain amaranths (Figs. 4, 5), whereas accessions *PI 166045* and *PI 553073* of *A. caudatus* are most closely related to each other and often distinctive from the rest of the intra- and inter-specific accessions. However, a general pattern of relationships exists between *A. caudatus* and the *A. hypochondriacus* - (*A. hybridus – A. cruentus*) group and with *A. quitensis*, which is consistent with the hypothesis for its evolutionary origin through the route from *A. hybridus* to *A. cruentus* domestication and subsequent gene flow from *A. quitensis* through introgression. *A. powellii*, however, is clearly the most distantly related to the grain amaranth group. Its putative role in the origin of *A. hypochondriacus* as a progenitor has not been supported by most data sets, although it is difficult to ascertain whether the species had historically contributed genes to *A. hypochondriacus* through natural hybridization and introgression. Similar conclusions could be drawn from previous studies where many other wild species of *Amaranthus* were also included in phylogenetic analyses (e.g. Lanoue et al. 1996; Chan and Sun 1997).

Acknowledgments This research was supported by a grant from the Research Grants Council of Hong Kong to M. Sun. We thank Mr. D. Brenner for supply of germplasm from the USDA/ARS collection and for valuable discussions on *Amaranthus* genetics.

References

Aggarwal RK, Majumdar KC, Lang JW, Singh L (1994) Genetic affinities among crocodilians as revealed by DNA fingerprinting with a Bkm-derived probe. Proc Natl Acad Sci USA 91:10601–10605

Akagi H, Yokozeki Y, Inagaki A (1997) Highly polymorphic microsatellites of rice consist of AT repeats, and a classification of closely related cultivars with these microsatellite loci. Theor Appl Genet 94:61–67

Antonius K, Nybom H (1994) DNA fingerprinting reveals significant amounts of genetic variation in a wild raspberry *Rubus idaeus* population. Mol Ecol 3:177–180

Bruford MW, Hanotte O, Brookfield JFY, Burke T (1998) Multilocus and single-locus DNA fingerprinting. In: Hoelzel AR (ed) Molecular genetic analysis of populations, a practical approach. Oxford University Press, Oxford, pp 287–336

Burke T, Dolf G, Jeffreys A, Wolff R (1991) DNA fingerprinting: approaches and applications. Birkhauser Verlag, Basel

Chan KF, Sun M (1997) Genetic diversity and relationships detected by isozyme and RAPD analysis of crop and wild species of *Amaranthus*. Theor Appl Genet 95:865–873

Cavalier-Smith T (1985) The evolution of genome size. John Wiley, Chichester

Clegg MT, Durbin ML (1990) Molecular approaches to the study of plant biosystematics. Austr Systematic Bot 3:1–8

Dice LR (1945) Measures of the amount of ecological associations between species. Ecology 26:297–302

Dover GA, Flavell RB (1982) Genome evolution. Academic Press, London

Doyle JJ (1991) DNA protocols for plants – CTAB total DNA isolation. In: Hewitt GM (ed) Molecular techniques in taxonomy. Springer, Berlin, pp 283–293

Gray EM (1995) DNA fingerprinting reveals a lack of genetic variation in northern populations of the western pond turtle (*Clemmys marmorata*). Conserv Biol 9:1244–1254

Hauptli H, Jain S (1984) Allozyme variation and evolutionary relationships of grain amaranths (*Amaranthus* spp.). Theor Appl Genet 69:153–65

Jeffreys AJ, Wilson V, Thein SL (1985a) Hypervariable "minisatellite" regions in human DNA. Nature 314:67–73

Jeffreys AJ, Wilson V, Thein SL (1985b) Individual-specific fingerprints of human DNA. Nature 316:76–79

John B, Miklos GLG (1988) The eukaryote genome in development and evolution. Allen and Unwin, London

Lanoue KZ, Wolf PG, Browning S, Hood EE (1996) Phylogenetic analysis of restriction-site variation in wild and cultivated *Amaranthus* species (Amaranthaceae). Theor Appl Genet 93: 722–732

Lynch M (1991) Analysis of population genetic structure by DNA fingerprinting. In: Burke et al. (eds) DNA fingerprinting approaches and applications. Birkhäuser Verlag, Basel, pp 113–126

Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. Mol Ecol 3:91–99

Nakamura Y, Leppert M, O'Connell P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kumlin E, White R (1987) Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235:1616–1622

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

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Rieseberg LH (1996) The homology among RAPD fragments in interspecific comparisons. Mol Ecol 5:99–105

Rohlf FJ (1994) Numerical taxonomy and multivariate analysis system. Exeter Software, Setauket, New York

Sauer JD (1950) The grain amaranths: a survey of their history and classification. Ann Miss Bot Gard 37:561–619

Sauer JD (1967) The grain amaranths and their relatives: a revised taxonomic and geographic survey. Ann Miss Bot Gard 54: 103–137

Sauer JD (1976) Grain amaranths. In: Simmonds NW (ed) Evolution of crop plants. Longman Group Limited, London, pp 4–7

Soltis DE, Soltis PS, Nickrent DL, Johnson LA, Hahn WJ, Hoot SB, Sweere JA, Kuzoff RK, Kron KA, Chase MW, Swensen SM, Zimmer EA, Chaw SM, Gillespie LJ, Kress WJ, Sytsma KJ (1997) Angiosperm phylogeny inferred from 18 S ribosomal DNA sequences. Ann Miss Bot Gard 84:1–49

Thormann CE, Ferreira ME, Camargo LEA, Tivang JG, Osborn TC (1994) Comparison of RFLP and RAPD markers for estimating genetic relationships within and among cruciferous species. Theor Appl Genet 88:973–980

Transue DK, Fairbanks DJ, Robison LR, Andersen WR (1994) Species identification by RAPD analysis of grain amaranth genetic resources. Crop Sci 34:1385–1389

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18: 6531–6535

Wolfe KH, Li W-H, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054–9058