Theor Appl Genet (1996) 93:722–732

K. Z. Lanoue · P. G. Wolf · S. Browning · E. E. Hood

Phylogenetic analysis of restriction-site variation in wild and cultivated *Amaranthus* species (Amaranthaceae)

Received: 19 December 1995 · Accepted: 19 April 1996

Abstract *Amaranthus* includes approximately 60 species, of which three are cultivated as a grain source. Many wild *Amaranthus* species possess agriculturally desirable traits such as drought and salt tolerance, and pathogen resistance. We examined relationships among wild and cultivated *Amaranthus* species based upon restriction-site variation in two chloroplast DNA regions and in a nuclear DNA region. The chloroplast regions consisted of (1) an intergenic spacer in transfer RNA genes and (2) the ribulose-1,5-bisphosphate carboxylase gene with a flanking open reading frame. The nuclear region was the internal transcribed spacers ITS-1 and ITS-2 flanking the 5.8S gene in the ribosomal DNA. These regions were amplified by the polymerase chain reaction and digested with a total of 38 restriction endonucleases. We detected 11 potentially informative restriction-site mutations and seven length-polymorphisms among the 28 *Amaranthus* species. Parsimony analysis was used to find the shortest tree for each separate data set (chloroplast, nuclear, and length) and for two combined matrices (chloroplast/nuclear and all data sets). Overall, there was a low level of variation which generated poorly resolved trees among the 28 species. Congruence analyses revealed that the chloroplast and nuclear data sets were congruent with each other but not to the length data set. The congruence of the chloroplast and nuclear data sets suggested that cytoplasmic gene flow may not be a confounding factor in our analyses. The phylogeny also suggested that drought tolerance evolved independently several times. The molecular phylogeny provides a basis for selection of species pairs for crop development.

Communicated by G. Wenzel

K. Z. Lanoue · P. G. Wolf · S. Browning · E. E. Hood (✉)[1]
Utah State University, Department of Biology, Logan, UT 84322-5305, USA

Present address:
[1]Pioneer Hi-Bred International, Incorporated, 7300NW 62nd Avenue, PO Box 1004, Johnston, IA 50131-1004, USA

Key words *Amaranthus* · Alternative crops · Molecular phylogenetics · PCR

Introduction

Three species of *Amaranthus* are commonly cultivated for grain production: *A. hypochondriacus*, *A. cruentus*, and *A. caudatus*. The three cultivated grain species have a relatively high nutritional value compared to monocot grains (Becker et al. 1981; Williams and Brenner 1995). For example, the crude protein content of the grain amaranths contains approximately 5% lysine and 4.4% sulfurous amino acids, which are limiting amino acids in other grains (Teutonico and Knorr 1985). In addition, the seeds of the three cultivated *Amaranthus* species have a higher mineral content than wheat, and contain vitamins A and C at nutritionally significant levels (Becker et al. 1981). Leaves from some *Amaranthus* species are grown as a nutritious vegetable similar to spinach. Wild *Amaranthus* species possess other agriculturally desirable traits such as drought and salt tolerance. Recently, amaranth has gained attention as an alternative crop (Reganold et al. 1990; Williams and Brenner 1995), and a phylogeny of the genus would aid in the development of a breeding program that utilizes wild *Amaranthus* species for crop improvement.

Several studies have examined evolutionary relationships among the cultivated *Amaranthus* species and a few wild species using hybrid analysis, morphology, allozymes, seed protein variation, and RAPD analysis (Sauer 1967; Pal and Khoshoo 1972, 1973; Jain et al. 1980; Hauptli and Jain 1984; Gudu and Gupta 1987; Sammour et al. 1993; Transue et al. 1994), but the results among these studies are variable. Recently, the chloroplast and the nuclear genomes have been utilized for phylogenetic analysis in other plant genera, because many regions exist that appear to be evolving at rates appropriate for interspecific comparison (Palmer et al. 1988; Hamby and Zimmer 1992; Liston 1992; Petersen and Doebley 1993; Kim and Jansen 1994). As a result,

examining specific regions in the chloroplast and nuclear DNA may provde additional information on which to base phylogenetic relationships among *Amaranthus* species.

The goal of the present study was to analyze restriction-site variation within specific regions of the chloroplast and nuclear DNA in *Amaranthus* in order to: (1) examine congruence between the variation within the chloroplast regions to that in the nuclear region, (2) estimate the intraspecific variation among populations of the three grain species, (3) explore patterns of evolution for drought tolerance and dioeciousness, and (4) identify possible wild species for future breeding.

Materials and methods

Accessions

Approximately 60 species are recognized in *Amaranthus* (Brenner 1990). Seed from 26 species was supplied by the USDA/ARS Plant Introduction Station in Ames, Iowa, and seed for nine populations of three grain amaranth species was obtained from the Rodale Research Center in Kutztown, Pa. (Table 1). Seed from two additional species was obtained from herbarium specimens (Table 1). The seed was germinated in a greenhouse, and 1–5 g of young leaves was used for DNA isolation. In addition, a mature plant of each species grown from seed was pressed and deposited in the Intermountain Herbarium at Utah State University.

Table 1 List of the 28 *Amaranthus* species used in this study. Seed for 26 species was obtained from the USDA/ARS Plant Introduction Station in Ames, Iowa. Seed for *A. pringlei* and *A. wrightii* was obtained from herbarium specimens on loan from the New York and Missouri Botanical Gardens, respectively. Seed for species denoted with a (*) was obtained from the Rodale Research Center in Kutztown, Pennsylvania

Amaranthus species	Accession	Seed origin	Habitat type
acutilobus Uline and Bray	Ames 13787	Germany	Unknown
albus L	Ames 13788	Canada	Dryroadsides and wasteplaces[1]
australis (A. Gray) J. Sauer	PI 553076	US/Fla.	Coastal marshes[2]
blitoides S. Watson	Not applicable	Utah cornfield	Dry roadsides and waste places[1]
cannabinus (L.) J Sauer	Ames 14359	US/Va.	Coastal margins[2]
caudatus L.	PI 553073	US	Dry areas[3]
caudatus L.*	713	Peru	Dry areas[3]
caudatus L.*	988	S. America	Dry areas[3]
caudatus L.*	1036	India	Dry areas[3]
crassipes Schldl.	Ames 10339	Czechoslovakia	Tropics[4]
cruentus L.	PI 477913	Mexico	Dry areas[3]
cruentus L.*	434	Mexico	Dry areas[3]
cruentus L.*	622	Guatemala	Dry areas[3]
cruentus L.*	1034	Africa	Dry areas[3]
deflexus L.	Ames 13779	Portugal	Coastal ports[5]
dubius C. Martius ex Thell.	Ames 5659	India	Tropics[6]
fimbriatus (Torrey) Benth. ex S Watson	Ames 15304	Mexico/Sonora	Desert[4,7,8]
floridanus (S Watson) Sauer	PI 553078	US/Fla.	Coastal dunes[2]
graecizans L.	PI 553079	US/Iowa	Dry roadsides and waste places[7]
hybrid (unknown hybrid)	Ames 16110	US/Calif.	Unknown
hybridus L	Ames 5684	US/Del.	Riverbanks[6]
hypochondriacus L.	PI 477917	Mexico	Dry areas[3]
hypochondriacus L.*	412	Mexico	Dry areas[3]
hypochondriacus L.*	646	Texmelucan	Dry areas[3]
hypochondriacus L.*	1221	Nepal	Dry areas[3]
lividus L.	Ames 5146	India	Wet areas
palmeri S. Watson	Ames 5370	US/Ariz.	Deserts[4]
powellii S. Wats.	Ames 13784	Germany	Deserts[6]
pringlei S. Wats.	N.Y. 16272	San Bernadino Co., Calif.	Moist areas[7]
pumilus Raf.	PI 553983	US/N.C.	Coastal dunes[4]
quitensis Kunth.	PI 511745	Ecuador	Riverbanks[6]
retroflexus L.	Ames 10826	US/Iowa	Roadsides, riverbanks, and waste places[5,7]
rudis J. Sauer	PI 553086	US/Iowa	Wet areas[9]
standleyanus L. Parodi ex Covas	Ames 15312	Argentina/La Pampa	Unknown
tricolor L.	Ames 2069	India/Tamil Nadu	Moist areas
viridis L	PI 540445	Indonesia/Java	Tropics[4,7]
wrightii S. Wats.	Mo. 400983	Hudspeth Co., Tex.	Dry grass and pinyon belts[5]

[1] Welsh et al. (1987)
[2] Sauer (1955)
[3] Sauer (1950)
[4] Wiggens (1980)
[5] Tidestrom and Kittell (1941)
[6] Sauer (1967)
[7] Kearney and Peebles (1960)
[8] McDougall (1973)
[9] Sauer (1972)

DNA isolation

The DNA extraction was based on the hexadecyltrimethylammonium bromide (CTAB) method of Doyle and Doyle (1990). However, two to three chloroform-isoamyl extractions were used, and the DNA pellets were then washed for 30 min at room temperature with 70% and 95% ethanol, respectively. Spinach (*Spinacia oleracea*) was used as the outgroup, and its DNA was isolated from leaves obtained from a local merchant. A pilot study determined that young leaves of a few amaranth species yielded DNA that was unsuitable for the polymerase chain reaction (PCR). Therefore, DNA from these species was obtained instead from seedlings, or else the young leaf DNA was purified on a cesium-chloride gradient (Maniatis et al. 1982).

PCR-amplification

The DNA from 28 *Amaranthus* species (including four populations each of *A. caudatus*, *A. cruentus*, and *A. hypochondriacus*) and from spinach was PCR-amplified with two sets of chloroplast primers and one set of nuclear primers (Fig. 1 A, B, and C, respectively). The two chloroplast regions are referred to here as BA-1 and ORF, respectively, and the nuclear region is referred to as ITS.

A Perkin Elmer DNA Thermal Cycler (TC-1) was used for all reactions. The PCR conditions and parameters for each region are described in Table 2. The *Taq* polymerase, $MgCl_2$, and reaction buffer were obtained from Promega (Madison, Wis, USA) and the dNTPs were obtained from Boehringer Mannheim (Indianapolis, Ind, USA).

The success of the PCR reactions was verified by electrophoresing 5 µl of each reaction through a 1% agarose gel in a 1 × Tris-borate (TBE) buffer, and staining the gel with ethidium bromide. The ORF region amplified the desired band as well as numerous bands of smaller size. Therefore, the entire reaction mixtures were size-separated on a 1% low-melting-temperature agarose gel in 1 × Tris-acetate (TAE) buffer. The desired band was excised and the DNA extracted with gelase (Epicentre Technologies, Madison, Wis. USA) which removes the agarose from the DNA. The gelase procedure was also done on *A. deflexus*, *A. floridanus* and *A. viridis* for the ITS region in order to isolate the desired amplified region from a second, smaller region. About 60% of the target DNA was recovered by the gelase method. The sizes of the amplified fragments were determined by comparison with known fragment sizes of lambda phage DNA digested separately with *Pst*I and *Hin*dIII.

The PCR-amplified DNA fragments were digested with the 38 restriction endonucleases (New England Biolabs, Beverly, Mass., USA) listed in Table 3. The restriction digestes were visualized on 1–4% agarose and 12% acrylamide gels stained with ethidium bromide.

Phylogenetic analysis

Data matrices were constructed based on the presence or absence of a restriction site in the chloroplast and nuclear regions (Table 4). Seven length-polymorphisms were also analyzed in a separate data set, and then were added to the combined chloroplast/nuclear data set (Table 4). Table 4 revealed four major groups of species. Each group was treated as a single entity in the parsimony analyses because all species contained within that group shared exactly the same character states.

The data matrices were analyzed using PAUP (phylogenetic analysis using parsimony) version 3.1.1 (Swofford 1993). Unweighted and character-state weighted (Wendel and Albert 1992) parsimony of 1.2:1 and 1.3:1 (gains to losses) were used. The trees for weighted and unweighted parsimony for all three data sets were identical (data not shown). Therefore, unweighted parsimony trees were used for comparisons of all data sets. A heuristic search was employed, and the shortest trees recovered for each data set were saved. The following

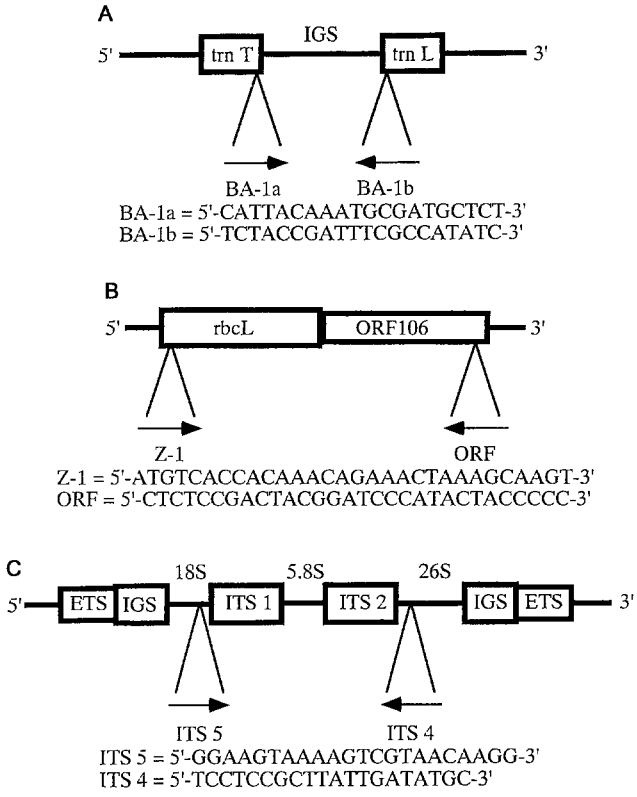


Fig. 1A–C Description of primer sets used for PCR. **A** Chloroplast BA-1 region (Taberlet et al. 1991) Transfer RNA-Threonine gene (*trn* T), transfer RNA-Leucine gene (*trnL*), and intergenic spacer (*IGS*). **B** Chloroplast ORF region (Rieseberg et al. 1992) Ribulose-1, 5-bisphosphate carboxylase gene (*rbcL*) and open reading frame 106 (*ORF 106*) from *Oryza sativa*. **C** Nuclear ITS region (White et al. 1990). Internal transcribed spacer (*ITS*), external transcribed spacer (*ETS*), and intergenic spacer (*IGS*)

Table 2 PCR parameters used for the 28 *Amaranthus* species and the three gene regions. All PCR reactions were done in 100 µl and were topped with 75 µl of mineral oil

PCR set-up	BA-1 region	ORF region	ITS region
DNA	30–50 ng	30–50 ng	30–50 ng
$MgCl_2$	4.0 mM	2.0–4.0 mM	4.0 mM
Reaction buffer	10x	10x	10x
dNTPS	0.2 µM each	0.2 µM each	0.2 µM each
Primers	0.25 µM	0.25 µM	0.25 µM
Taq polymerase	1–2 Units	2 Units	1 Unit
PCR steps	1:94 °C, 1:30 min	1:94 °C; 5:00 min	1:94 °C: 5:00 min
	2:38 °C: 2:00 min	2:94 °C; 1:30 min	2:94 °C; 1:30 min
	3:72 °C: 5:00 min	3:38 °C; 2:00 min	3:50 °C; 1:00 min
	4:steps 1–3, 35 times	4:72 °C; 5:00 min	4:72 °C; 2:00 min
	5:72 °C; 10:00 min	5:steps 2–4, 35 times	5:steps 2–4, 35 times
		6:72 °C; 10:00 min	6:72 °C; 10:00 min

Table 3 Restriction endonucleases used to analyze the three gene regions PCR-amplified from 28 *Amaranthus* species. The three gene regions are denoted as BA-1, ORF, and ITS. Phylogenetically informative restriction sites are those shared by at least two species but not by all species. The Xs denote enzymes not used for the specified region

Restriction endonuclease	# Informative sites			Restriction endonuclease	# Informative sites		
	BA-1	ORF	ITS		BA-1	ORF	ITS
*Aci*I	0	0	2	*Hinf*I	1	0	0
*Alu*I	0	0	0	*Hpa*I	0	0	0
*Apa*I	0	0	0	*Hph*I	X	X	0
*a-Taq*I	1	0	0	*Kpn*I	0	0	0
*Bam*HI	0	0	0	*Mbo*I	0	1	0
*Ban*I	X	0	X	*Mse*I	0	0	2
*Bfa*I	0	0	1	*Msp*I	0	0	0
*Bgl*I	0	0	0	*Nci*I	0	0	0
*Bgl*II	X	0	X	*Nla*III	0	0	0
*Bst*EII	X	0	X	*Pst*I	0	0	0
*Bst*UI	0	0	0	*Pvu*II	X	0	X
*Bst*XI	0	0	0	*Rsa*I	0	0	1
*Cla*I	0	0	0	*Sac*I	0	0	0
*Dpn*I	0	X	X	*Sal*I	0	0	0
*Eco*RI	1	0	0	*Sau*3AI	0	0	0
*Eco*RV	0	1	0	*Sph*I	X	0	X
*Hae*III	0	0	0	*Sma*I	X	0	X
*Hha*I	0	0	0	*Tsp*509I	0	0	0
*Hind*III	0	0	0	*Xba*I	0	0	0

Table 4 Data matrix of the 18 characters used in the phylogenetic analyses. The data are coded so that the gain of a restriction site and the presence of a length polymorphism each equal (1), and the loss of a restriction site and the absence of a length polymorphism each equal (0). Missing data are represented by (?). Populations excluded in the analyses of between-data-set congruence are highlighted in bold

Amaranthus species	Nuclear data set						Chloroplast data set					Length data set						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
acutilobus	1	0	0	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0
Group 1: *albus, australis, blitoides, cannabinus, graecizans*	0	0	0	0	1	1	0	1	1	1	0	0	1	1	1	1	1	0
Group 2: *caudatus* (PI 553073), *cruentus* (PI 477913), hybrid, *hybridus*	1	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0
Group 3: *crassipes, palmeri, standleyanus*	1	0	0	0	1	1	0	1	1	1	0	0	1	1	1	1	1	0
Group 4: *caudatus* (713), *cruentus* (434,622,1034), *hypochondriacus* (412, 646, 1221)	1	0	0	0	1	0	0	0	?	1	0	?	?	?	?	?	?	?
caudatus (988)	1	0	0	0	1	1	0	0	?	1	0	?	?	?	?	?	?	?
caudatus (1036)	1	0	0	0	1	1	0	0	?	1	0	?	?	?	?	?	?	?
deflexus	1	0	0	1	0	0	0	1	1	1	1	0	1	1	1	1	1	0
dubius	1	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	1
fimbriatus	0	1	0	0	1	1	0	1	1	1	0	0	1	1	1	1	1	0
floridanus	0	0	1	1	0	0	0	1	1	1	0	0	1	1	1	1	1	0
hypochondriacus (PI 477917)	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
lividus	1	0	0	0	1	1	1	1	1	0	0	0	1	1	1	1	1	0
powellii	1	0	0	0	1	1	0	0	0	1	0	0	1	1	1	1	1	1
pringlei	1	0	0	0	1	1	0	1	1	1	0	0	0	0	0	0	0	0
pumilus	0	1	0	0	1	1	0	1	1	1	0	0	1	1	1	1	1	0
quitensis	1	0	0	0	1	0	0	0	1	1	0	1	1	1	1	1	1	0
retroflexus	1	0	0	0	1	1	1	0	0	1	0	0	0	0	0	0	0	0
rudis	0	0	1	0	1	1	0	1	1	?	?	0	1	1	1	1	1	?
tricolor	1	0	0	0	1	1	1	1	1	0	0	0	1	1	1	1	1	0
viridis	1	0	0	1	0	0	0	1	1	1	0	1	1	1	1	1	1	0
wrightii	0	0	1	0	1	1	0	1	1	1	0	0	1	1	1	1	1	0
spinach	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0

options in the heuristic search were used; TBR, MULPARS, COLLAPSE, and the random addition sequence using 100 replicates.

The skewness of the tree length distribution (g1 value) based upon 10000 random trees was calculated by PAUP (Swofford 1993) to determine whether the data represented random noise or phylogenetic signal (Hillis and Hulsenbeck 1992). In addition, the consistency, homoplasy, and retention indices were measured to evaluate the congruence among characters in a data set.

The Mickevich and Farris index (Swofford 1991) and the Miyamoto index (Swofford 1991) were calculated using information from PAUP and MacClade version 3 (Maddison and Maddison 1992) to determine the degree of congruence among data sets so that the accuracy of the combined matrices could be evaluated. To determine the Mickevich and Farris index (Swofford 1991), we first calculated the within data-set incongruence (i_w) which is the sum of the number of extra steps (e) for each separate data set ($e = s - c$,

Table 5 Restriction-site mutations and length polymorphisms detected in three gene regions. a 600-bp cpDNA fragment (BA-1), a 3200-bp cpDNA fragment (ORF), and an 800-bp nDNA fragment (ITS) among 28 species of *Amaranthus*. Restriction-site mutations (1–11) are described as changes in fragment size (bp) where the primitive character state precedes the derived character state, except for *Mbo*I where it is reversed. Fragments in parentheses denote presumed sizes unresolved due to small size. An increase in size (12–18) was scored as the derived character state and the length polymorphism is indicated in bold

Numbers shown on phylogenetic trees	Amplified region	Restriction endonuclease	Restriction-site polymorphisms or length polymorphisms
1	ITS	*Mse*I	800 = 430 + 300 + (70)
2	ITS	*Mse*I	300 = 230 + (70)
3	ITS	*Rsa*I	800 = 650 + (150)
4	ITS	*Bfa*I	800 = 400 + 400
5	ITS	*Aci*I	800 = 460 + 340
6	ITS	*Aci*I	460 = 300 + 160 and 340 = 240 + 100
7	BA-1	*a-Taq*I	330 = 270 + (60)
8	BA-1	*Eco*RI	600 = 350 + 250
9	BA-1	*Hin*fI	200 = 150 + 50
10	ORF	*Eco*RV	3200 = 1700 + 1200 + (?)
11	ORF	*Mbo*I	1350 = 850 + 500
12	BA-1	*Hin*fI	95 + **35** = 130
13	BA-1	*Alu*I	400 + **50** = 450
14	BA-1	None	600 + **50** = 650
15	BA-1	*Bst*UI	450 + **50** = 500
16	BA-1	*Sau*3AI	450 + **50** = 500
17	BA-1	*Mbo*I	450 + **50** = 500
18	ORF	*Bfa*I	700 + **50** = 750

where s = number of steps on the tree and c = number of characters). Then, we calculated the total incongruence (i_T) which is (e) for the combined data set. From this information the between-data-set incongruence ($i_B = i_T - i_W$) was determined so that the proportion of the total incongruence due to between-data-set incongruence (i_B / i_T)could be estimated.

The Miyamoto index also provides an estimate for between-data-set incongruence. First, the within-data-set incongruence was calculated in the same way as described above, but the total incongruence (i_T^*) equals the sum of the extra steps required for one data set to explain the minimal tree topology of another data set ($e^* = s - c$). Therefore the proportion of the total incongurene due to between-data-set incongruence equals $i_T^* - i_W / i_T^*$.

Results

PCR-amplified regions

The BA-1 primer set PCR-amplified a product that was approximately 600 bp with a 50-bp length polymorphism common to about half of the species. All species except *A. rudis* and *A.* hybrid (an unknown cross from the USDA/ARS Plant Introduction Station in Ames, Iowa) amplified with the ORF primer set which produced a PCR product that was approximately 3200 bp. *A. rudis* and *A.* hybrid were scored as missing (?) in the data matrix. In addition, all species amplified with the ITS primer set yielding an 800-bp product.

Eighteen restriction sites were detected in the BA-1 region, of which four were informative (shared by more than one species) (Table 5; Fig. 2). The ITS region produced eight informative restriction sites and the ORF region produced only two informative restriction sites out of 49 examined (Table 5).

Phylogenetic analysis

In order to avoid possible phylogenetic errors due to the use of a single data set (Chippindale and Wiens 1994; Kim and Jansen 1994; Olmstead and Sweere 1994), we

Fig. 2 An *Eco*RI restriction digest of the BA-1 region for 29 *Amaranthus* species including *A. spinosus*, *A. muricatus*, *A. caudatus* (713, 988, 1036), *A. cruentus* (434, 622, 1034), and *A. hypochondriacus* (412, 646, 1221, data not shown) (*lanes 1–29*), spinach (*lane 30*), uncut PCR product (*lane 31*), and lambda *Pst*I and *Hin*dIII size markers (*lanes 32 and 33*) The 29 *Amaranthus* species were run on a 1% agarose gel stained with ethidium bromide in alphabetical order as listed in Table 1. A species without an *Eco*RI recognition site (*lane 1*), with one recognition site (*lane 2*), and a partial digest of a species with one recognition site (*lane 11*) are shown

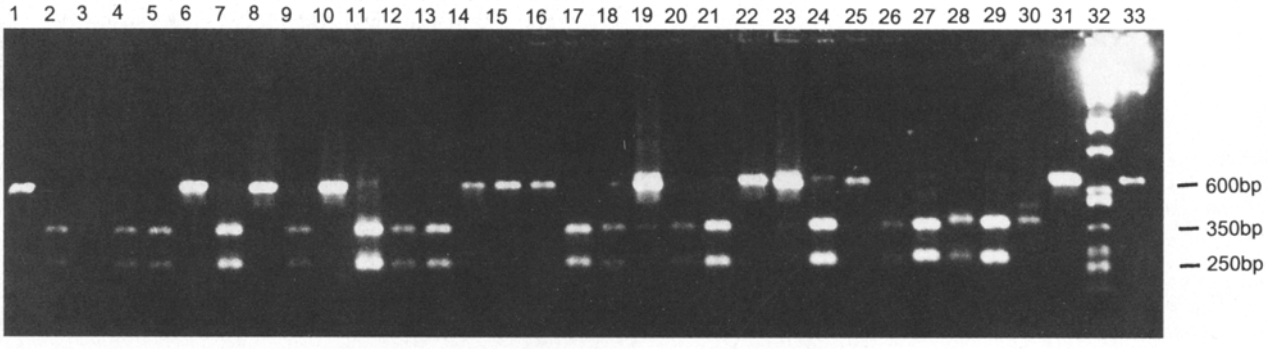


Table 6 Results of the parsimony analysis by PAUP version 3.1.1 (Swofford 1991) using Wagner (ordered) parsimony and the heuristic search. The statistics for each tree are based on the 50% majority rule consensus tree except for the chloroplast nuclear tree where only one tree was found

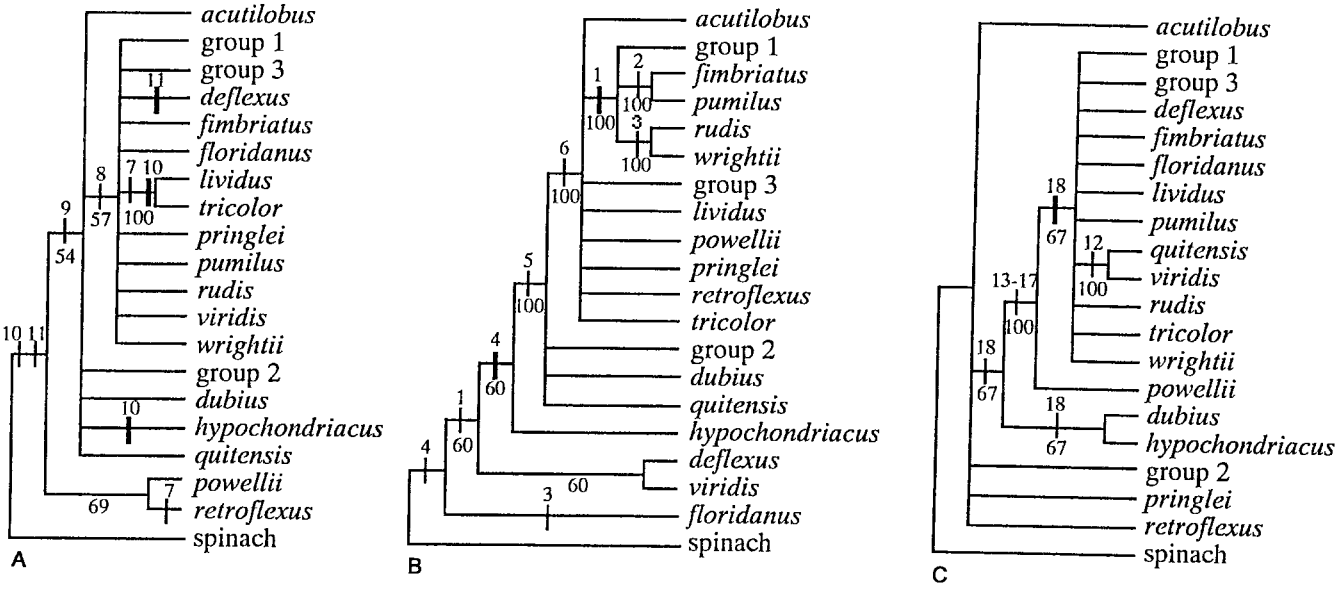
Item	Chloroplast	Nuclear	Chloroplast/nuclear	Length	All data sets
# Of taxa and/or groups	20	20	20	20	20
# Of characters	5	6	11	7	18
# Of MPT(s)	91	5	1	3	137
# Of steps on MPT(s)	9	9	19	8	33
Consistency index (CI)	0.56	0.67	0.58	0.88	0.51
Homoplasy index (HI)	0.44	0.33	0.42	0.13	0.49
Retention index(RI)	0.56	0.86	0.76	0.97	0.76
g1 value	−0.48	−0.56	−0.45	−0.72	−0.57

analyzed the data sets separately and then assembled them into the following two combined matrices: the chloroplast/nuclear matrix and a matrix which incorporated all three data sets.

The analyses of the three separate data sets differed in the number of characters, the consistency of the characters, and the number of minimal trees found (Table 6; Fig. 3 A, B, and C). To examine congruence between the chloroplast and nuclear data sets, they were assembled into one matrix from which one minimal tree was generated (Fig. 4 A). The combined chloroplast/nuclear tree was more resolved than the trees based upon the separate chloroplast and nuclear data sets, and had a consistency index of 0.58 which was higher than that for the chloroplast 50% majority rule consensus tree (0.56) and less than that for the nuclear 50% majority rule consensus tree (0.67). The resolution of the combined tree is indicative of a degree of congruence among data sets (Swofford 1991) which is also supported by a relatively stable consistency index. In addition, the shortest tree for the combined chloroplast/nuclear data set had a significant g1 value of −0.45 (Table 5), which also suggested that the combined data exhibited a phylogenetic signal comparable to the significant g1 values of the separate chloroplast and nuclear data sets which were −0.48 and −0.56, respectively. The major difference between the chloroplast and nuclear data sets was in the placement of *A. deflexus*, *A. floridanus* and *A. viridis*. The combined tree placed these three species similarly to their placement in the nuclear tree. The discrepancy found between the chloroplast and nuclear data sets may be due to the possibility that *A. deflexus*, *A. floridanus* and *A. viridis* could be heterozygous for the nuclear PCR product because all three species showed a doublet for this primer set (data not shown).

Addition of the length data set to the chloroplast/nuclear data set produced a matrix that incorporated all available characters and generated a resolved 50% majority rule consensus tree (Fig. 4 B) of 137 minimal trees. However, the consistency index of the combined tree

Fig. 3 A–C Fifty percent majority rule consensus trees of 28 *Amaranthus* species with spinach as the outgroup. Of the four populations sampled for *A. caudatus*, *A. cruentus* and *A. hypochondriacus*, only PI 553073, PI 477913, and PI 477917, respectively, were included in the analyses. Restriction-site mutations are labeled *1 to 18* as listed in Table 4, and *numbers* under the branches indicate the percentage of time that the proceeding node occurred among all trees. *Thin vertical lines* represent restriction-site gains, and *thickened vertical* lines represent reversals. **A** Fifty percent majority-rule consensus tree of the 91 MPTs using the chloroplast data set. **B** Fifty percent majority-rule consensus tree of the five MPTs using the nuclear data set. **C** Fifty percent majority-rule consensus tree of the three MPTs using the length data set



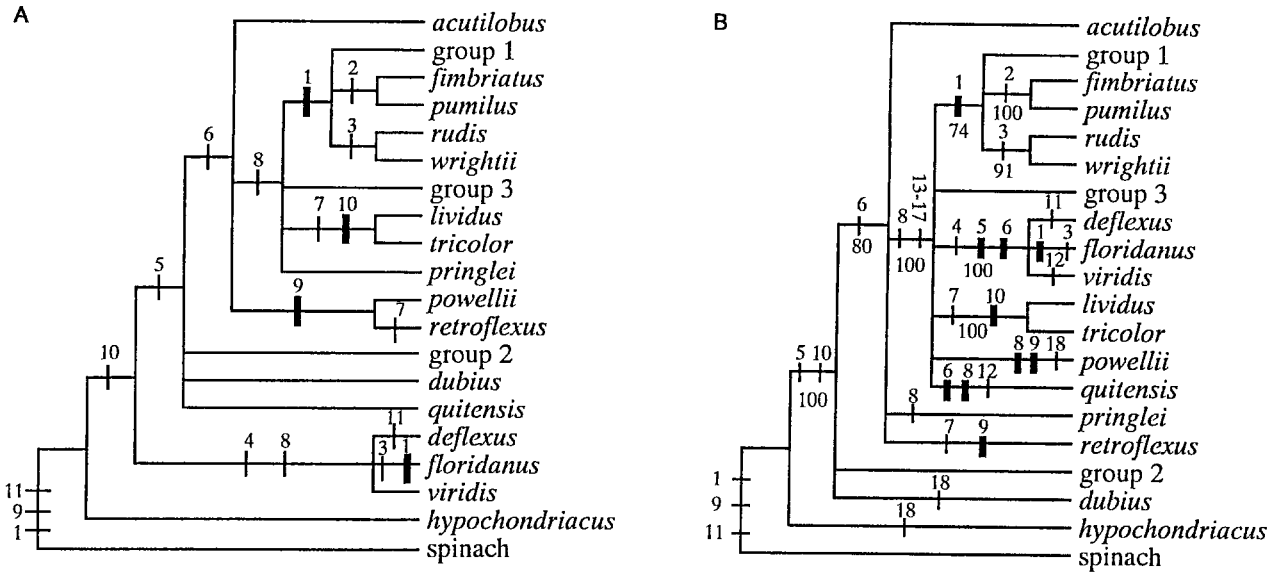


Fig. 4A, B Minimal trees of 28 *Amaranthus* species with spinach as the outgroup. Of the four populations sampled for *A. caudatus*, *A. cruentus* and *A. hypochondriacus*, only PI 553073, PI 477913, and PI 477917, respectively, were included in the analyses Restriction-site mutations are labeled *1 to 18* as listed in Table 4, and *numbers* under the branches indicate the percentage of time that the proceeding node occurred among all trees *Thin vertical lines* represent restriction-site gains, and *thickened vertical lines* represent reversals. **A** Minimal tree using the combined chloroplast/nuclear data set. **B** Fifty percent majority rule consensus tree of the 137 MPTs using all data sets combined

was 0.51, which was lower than the consistency indices for both the length 50% majority rule consensus tree (0.88) and for the chloroplast/nuclear tree (0.58). Even though the 50% majority rule consensus tree was resolved (Swofford 1991), these results suggested that the combination of all three data sets was not appropriate. The incompatibility of the three data sets is most likely due to the length data set not representing homologous characters, because it is not known for certain if the length mutation is identical between species by a size separation on an agarose gel (Swofford and Olsen 1990).

Character incongruence due to the combination of data sets was also examined by calculating the Mickevich and Farris index (Swofford 1991) and the Miyamoto index (Swofford 1991). The results in Table 7 indicated a large amount of variation between the two indices. According to the Mickevich and Farris index (I_{MF}), only 12.5% of the character incongruence of the combined chloroplast/nuclear data set was due to the assembly of the chloroplast and nuclear data sets. However, the Miyamoto index (I_M) suggested that this incongruence was 36%. The same type of situation was also seen for the combination of the chloroplast/nuclear and length data sets (Table 7). However, there was a much higher degree of incongruence ($I_{MF} = 36\%$ and $I_M = 72\%$) which also implied a lack of congruence between the length and chloroplast/nuclear data sets, as discussed above.

Swofford (1991) suggested that the Mickevich and Farris index may underestimate the degree of character incongruence between data sets due to an unequal

Table 7 Character incongruence among the *Amaranthus* data sets. Number of characters (c), number of steps of the MPT (s), the number of extra steps ($e = s - c$), the number of extra steps when data set (x) is applied to tree (y) ($e^* = s - c$), within-data-set incongruence ($i_W = \Sigma e$), between-data-set incongruence ($i_B = i_T - i_W$), total incongruence for the Mickevich and Farris index ($i_T = s - c$ for combined data sets), total incongruence for the Miyamoto index ($i_T{}^* = \Sigma e^*$), Miyamoto index ($I_M = (i_T{}^* - i_W)/i_T{}^*$), and the Mickevich and Farris index ($I_{MF} = i_B/i_T$)

Tree (y) Data set (x)	Chloroplast	Nuclear	Length	Chloroplast/nuclear	All data sets
Chloroplast	$c = 5; s = 9; e = 4$	$c = 5; s = 11; e^* = 6$	$c = 5; s = 12; e^* = 7$		
Nuclear	$c = 6; s = 11; e^* = 5$	$c = 6; s = 9; e = 3$	$c = 6; s = 13; e^* = 7$		
Length	$c = 7; s = 19; e^* = 12$	$c = 7;, s = 25; e^* = 18$	$c = 7; s = 8; e = 1$ $c = 11; s = 25; e^* = 14$	$c = 7; s = 25; e^* = 14$ $c = 11; s = 19; i_W = 7$ $i_B = 1; i_T = 8; i_T{}^* = 11$ $I_M = 0.36; I_{MF} = 0.125$	
Chloroplast/nuclear					
All data sets					$c = 18; s = 32; i_W = 9$ $i_B = 5; i_T = 14; i_T{}^* = 32$ $I_M = 0.72, I_{MF} = 0.36$

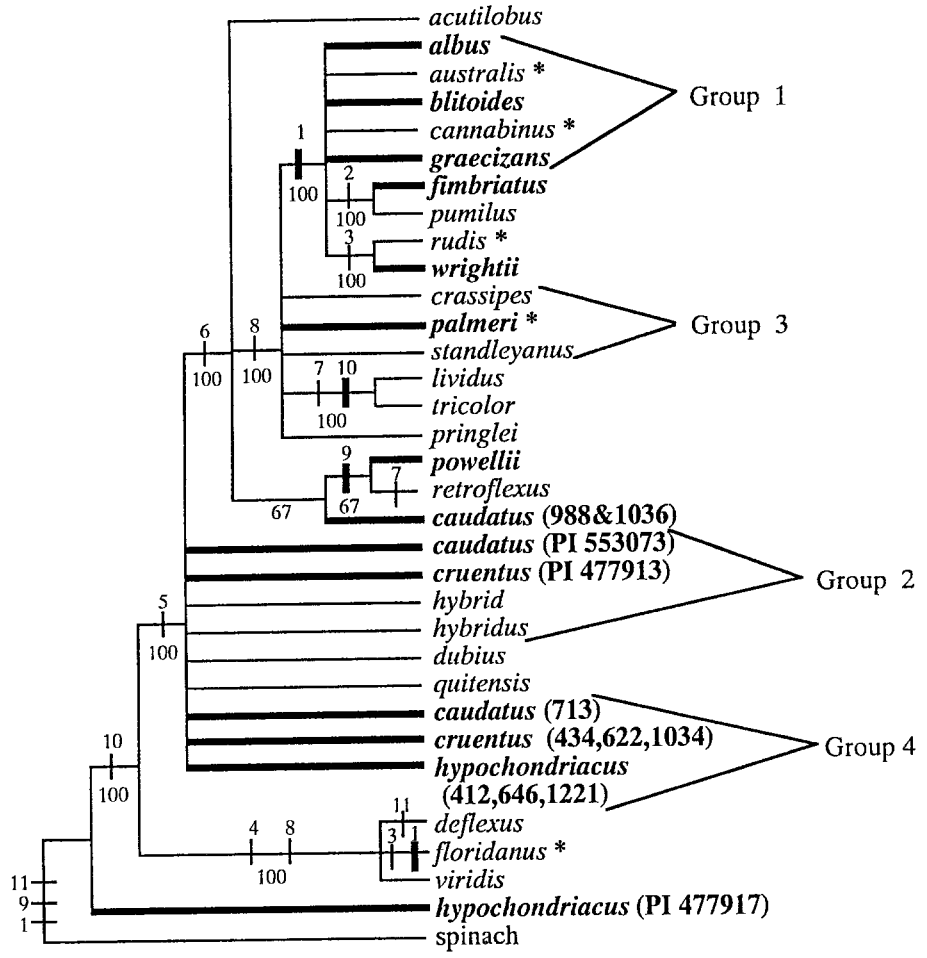


Fig. 5 The fifty percent majority rule consensus tree of three MPTs constructed from the chloroplast/nuclear data set of 28 *Amaranthus* species and three additional populations each of *A. caudatus*, *A. cruentus* and *A. hypochondriacus*, with spinach as the outgroup. Restriction-site mutations are labeled *1 to 11* as listed in Table 4, and *numbers* under the branches indicate the percentage of time that the proceeding node occurred among the three MPTs. *Thin vertical lines* represent restriction-site gains, and *thickened vertical lines* represent reversals. Drought-tolerant species are highlighted in *bold* and dioecious species are denoted with an (*)

number of characters in data sets that are combined. However, the *Amaranthus* chloroplast and nuclear data sets only varied by one character. Swofford (1991) also suggested that a complication may arise when calculating the Miyamoto index if more than one minimal tree exists for either data set. Indeed, more than one minimal tree was found by the parsimony analysis for each *Amaranthus* data set except for the chloroplast/nuclear data set. Therefore, the Miyamoto index was calculated by using the 50% majority rule consensus tree when necessary. On the other hand, the Mickevich and Farris index would not be sensitive to the use of a consensus tree because the number of steps remains the same among all minimal trees. Therefore, for the *Amaranthus* data sets, the Mickevich and Farris index may be the more accurate measure of between-data-set incongruence. Consequently, a Mickevich and Farris value of 12.5% indicated that the between-data-set incongruence for the combined chloroplast/nuclear tree is less than the within-data-set incongruence, and therefore it is appropriate to combine the chloroplast and nuclear restriction-site data sets (Kluge 1989).

Discussion

Grain amaranths

Two common grain amaranths, *A. caudatus* and *A. cruentus*, shared identical character states across all three data sets, and are included in group 2 along with *A. hybridus* and *A.* hybrid (Table 4). In addition, *A. dubius* and *A. quitensis* fall into group 2 when the length data set is removed from the phylogenetic analysis (Fig. 5). The putative progenitors of *A. caudatus* and *A. cruentus* are considered to be *A. quitensis* and *A. hybridus*, respectively (Williams and Brenner 1995). Our analysis finds these two grain species identical to their putative progenitor in all phylogenetically informative characters. This finding is in accordance with Doebley (1992) who postulates that a crop species and its progenitor should maintain a close phylogenetic relationship due to the recent time since domestication.

However, other authors have found that *A. caudatus* and *A. cruentus* are not as closely related to each other as

each is to *A. hypochondriacus* (Transue et al. 1994; Williams and Brenner 1995). In our study, *A. hypochondriacus* differs from *A. caudatus* and *A. cruentus* by two character states. This discrepancy is most likely a result of the accessions used in our DNA analysis. Accessions of the grain amaranths have been found to vary in morphological characters (Transue et al. 1994), hybrid viability (Pal and Khoshoo 1972, 1974; Transue et al. 1994), and chromosome number (Greizenstein and Poggio 1994). Therefore, we did additional analyses, using only the phylogenetically informative nuclear and chloroplast data sets (Table 4), for three additional accessions of each of the grain *Amaranthus* species. For the most part, the populations of each species grouped together (Fig. 5). The four populations of *A. cruentus* were the most similar (although complete identity was uncertain because of a missing data point). One character state in the nuclear DNA was found to divide the populations of *A. caudatus*, and a single *A. hypochondriacus* population was separated from the remaining three due to one character-state difference in the chloroplast DNA and one in the nuclear DNA. A total of three characters were found to differ among the 12 populations of the three species, of which two were only variable in *A. hypochondriacus* PI 477917 (Table 4; Fig. 5). Overall, three out of four accessions each of *A. cruentus* and *A. hypochondriacus* were found to be identical to each other (group 4) suggesting that they are more closely related than either is to *A. caudatus*.

Introgression

Phylogenetic analyses in many plant genera have noted the introgression of cytoplasmic genomes between species, causing incongruence among characters in the nuclear and chloroplast genomes (Rieseberg and Soltis 1991). As a result, chloroplast-based phylogenies were found to depict different relationships than those based upon non-chloroplast data sets. The results of both the parsimony and congruence analyses in our study indicate that the nuclear and chloroplast data sets are congruent, suggesting that the introgression of foreign organellar DNA was not of frequent occurrence among the *Amaranthus* species sampled. In addition, there was little variation among populations within a species (Fig. 5), especially within the chloroplast DNA, which also suggests infrequent introgression. However, the overall low level of variation found in our study may not be sufficient to reveal the true frequency of cytoplasmic gene flow. The extent of introgression in *Amaranthus* needs to be further evaluated by examining intraspecific variation within the wild species, because the cultivated species may not be representative of the entire genus. Introgression may be less common within the cultivated species due to heavy selection against weedy traits which may be a result of introgression from neighboring weedy relatives (Doebley 1992).

Drought tolerance and dioeciousness

Drought tolerance appears to have evolved independently of the conserved gene regions examined in this study. The phylogenetic tree constructed from the chloroplast/nuclear data set (Fig. 5) reveals that species adapted to dry environments do not form a monophyletic group. For example, *A. hydrochondriacus* (PI 477917), *A. fimbriatus*, *A. palmeri* and *A. wrightii* are often found in dry areas. The three latter species are separated from *A. hypochondriacus* (PI 477917) by six, four, and six character-state changes, respectively. Therefore, the distribution of drought tolerance may reflect parallel evolution due to independent occurrence of similar habitats.

Sauer (1955) describes the dioecious amaranths *A. australis*, *A. cannabinus*, *A. floridanus*, *A. palmeri* and *A. rudis*, and suggests a single origin independent of the rest of the genus. This hypothesis is inconsistent with the results of our phylogenetic analysis. Three dioecious species, *A. australis*, *A. cannabinus* and *A. rudis*, are very closely related, while two others, *A. palmeri* and *A. floridanus*, are less related to the former cluster than are many monoecious amaranths (Fig. 5). Our phylogeny (Fig. 5) suggests a minimum of three transitions between dioeciousness and monoeciousness.

Future prospects

The need for sustainability in agriculture has become accepted by farmers, government, and environmentalists, because the consequences of monoculturing have become apparent in the last 20 years: namely, decreasing soil fertility, high energy costs, low farm incomes, excessive ground and air pollution, destruction of wildlife habitats, and an overall decrease in biodiversity (Reganold et al. 1990; Schaller 1993). Due to an increasing awareness for sustainability over monoculturing, the development of alternative crops is finding favor (Reganold et al. 1990).

Amaranth could be a valuable alternative crop due to its high nutritional content, the many desirable traits expressed by the grain amaranths, and the unexplored potential of the wild species. Unlike the traditional monocot grains, the grain amaranths possess lysine in nutritionally significant levels. In addition, some varieties of grain amaranths tolerate drought and heat and, as a result, lessen the demand for costly irrigation. Peruvian varieties of *A. caudatus* have been found that are resistant to damping off and root rot, common diseases of crop plants (National Research Council 1989). Grain amaranths have been found to tolerate acid, saline, and alkaline soils and also soils high in aluminum (Williams and Brenner 1995).

There are also amaranth species that can be grown as a nutritious vegetable source. *Amaranthus tricolor*, *A. dubius*, *A. cruentus*, *A. hybridus* and *A. lividus* have all been used as vegetables in hot, humid areas of Africa,

Southeast Asia, China, and India (National Research Council 1985). Some vegetable amaranths have been known to withstand torrential rains, while others, like *A. palmeri*, thrive in North American deserts.

As described above, disease or drought resistance have been associated with a few selected varieties of cultivated amaranths. This implies that the related wild species may be valuable reservoirs for desirable traits common to all populations, unlike the cultivated species where the expression of desirable traits is only seen in a few populations. Traits that would most likely be found among the wild species include drought and salt tolerance, and resistance to a variety of viral and bacterial diseases. Based upon our phylogeny (Fig. 4 A), a number of wild species closely related to cultivated amaranths should be explored for breeding purposes: *A. floridanus* and *A. pumilus* for salt tolerance, *A. powellii* and *A. fimbriatus* for drought resistance, and *A. hybridus*, *A. quitensis* and *A. retroflexus* for pest, viral, and bacterial resistance. Not only would these wild species be valuable for breeding with the cultivated amaranths, but they would also be favorable species to utilize in the genetic engineering of desirable traits into other crops.

References

Becker R, Wheeler EL, Lorenz K, Stafford AE, Grosjean OK, Betschart AA, Saunders RM (1981) A compositional study of amaranth grain. J Food Sci 46:1175–1180

Brenner D (1990) The grain amaranth gene pools. In: Proc 4th Amaranth Symp, amaranth: perspectives on production, processing and marketing. Minneapolis, Minnesota, pp 193–194

Chippindale PT, Wiens JJ (1994) Weighting, partitioning, and combining characters in phylogenetic analysis. Syst Biol 43: 278–287

Doebley J (1992) Molecular systematics and crop evolution. In: Soltis PS, Soltis DE, Doyle JJ (eds) Molecular systematics of plants. Chapman and Hall, New York, pp 202–222

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue tissue. Focus 12:13–15

Greizenstein EJ, Poggio L (1994) Karyological studies in grain amaranths. Cytologia 59:25–30

Gudu S, Gupta VK (1987) Electrophoresis as an aid for identification of various species and cultivars of grain amaranths. E Afr Agric For J 52:244–250

Hamby RK, Zimmer EA (1992) Ribosomal RNA as a phylogenetic tool in plant systematics. In: Soltis PS, Soltis DE, Doyle JJ (eds) Molecular systematics of plants. Chapman and Hall, New York, pp 50–91

Hauptli H, Jain S (1984) Allozyme variation and evolutionary relationship of grain amaranths (*Amaranthus* spp.) Theor Appl Genet 69:153–165

Hillis DM, Huelsenbeck JP (1992) Signal, noise, and reliability in molecular phylogenetic analyses. J Hered 83:189–195

Jain SK, Wu L, Vaidya KR (1980) Levels of morphological and allozyme variation in Indian amaranths: a striking contrast. J Hered 71:283–285

Kearney TH, Peebles RH (1960) Arizona flora. University of California Press, Berkeley and Los Angeles, California, pp 265–267

Kim KJ, Jansen RK (1994) Comparisons of phylogenetic hypotheses among different data sets in dwarf dandelions (*Krigia*, Asteraceae): additional information from internal transcribed spacer sequences of nuclear ribosomal DNA. Pl Syst Evol 190:157–185

Kluge AG (1989) A concern for evidence and a phylogenetic hypothesis of relationships among Epicrates (Boideae Serpentes). Syst Zool 38:7–25

Liston A (1992) Variation in the chloroplast genes rpoCl and rpoC2 of the genus *Astragalus* (Fabaceae): evidence from restriction-site mapping of a PCR-amplified fragment. Am J Bot 79:953–961

Maddison WP, Maddison DR (1992) MacClade: analysis of phylogeny and character evolution, version 3. Sinauer Associates, Inc., Sunderland, Massachusetts

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

McDougall WB (1973) Seed plants of northern Arizona. The Museum of Northern Arizona, Flagstaff, Arizona, pp 156–157

National Research Council (1985) Vegetable amaranths. In: National Research Council (eds) Amaranth: modern prospects for an ancient crop Rodale Press Inc, pp 39–47

National Research Council (1989) Kiwicha. In: National Research Council (eds) Lost crops of the Incas: little known plants of the Andes with promise for worldwide cultivation. National Academy Press, Washington D.C., pp 139–147

Olmstead RG, Sweere JA (1994) Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. Syst Biol 43:467–481

Pal M, Khoshoo TN (1972) Evolution and improvement of cultivated amaranths. V. Inviability, weakness, and sterility in hybrids J Hered 63:78–82

Pal M, Khoshoo TN (1973) Evolution and improvement of cultivated amaranths VI. Cytogenetic relationships in grain types Theor Appl Genet 43:242–251

Pal M, Khoshoo TN (1974) Grain amaranths, evolutionary studies in world crops. In: Hutchinson J (ed) Diversity and change in the Indian sub-continent University Press, Cambridge, pp 129–137

Palmer JD, Jansen RK, Michaels HJ, Chase MW, Manhart JR (1988) Chloroplast DNA variation and plant phylogeny. Ann Missouri Bot Gard 75:1180–1206

Petersen G, Doebley JF(1993) Chloroplast DNA variation in the genus *Secale* (Poaceae). Pl Syst Evol 187:115–125

Reganold JP, Papendick RI, Parr JF (1990) Sustainable agriculture. Sci Amer June 1990.112–120

Rieseberg LH, Soltis DE (1991) Phylogenetic consequences of cytoplasmic gene flow in plants. Evol Trends Plants 5:65–84

Rieseberg LH, Hanson MA, Philbrick CT (1992) Androdioecy is derived from dioecy in Datiscaceae; evidence from restriction-site mapping of PCR-amplified chloroplast DNA fragments. Syst Bot 17:324–336

Sammour RH, Hammoud MA, Abd Alla SA (1993) Electrophoretic variations in *Amaranthus*. Bot Bull Acad Sin 34:37–42

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

Sauer JD (1955) Revision of the dioecious amaranths. Madrono 13: 5–46

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

Sauer JD (1972) The dioecious amaranths: a new species name and major range extensions Madrono 21:426–434

Schaller N (1993) The concept of agricultural sustainability. Agric Ecol Environ 46:89–97

Swofford DL (1991) When are estimates from molecular and morphological data incongruent. In: Miyamoto MM, Cracraft J (eds) Phylogenetic analysis of DNA sequence. Oxford University Press, New York, pp 295–333

Swofford DL (1993) PAUP Phylogenetic analysis using parsimony, version 3.1. Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois

Swofford DL, Olsen GJ (1990) Phylogeny reconstruction In: Hillis DM, Moritz C (eds) Molecular systematics Sinauer Associates, Inc. Sunderland, Massachusetts, pp 411–502

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105–1109

Teutonico RA, Knorr D (1985) Amaranth: composition, properties, and applications of a rediscovered food crop Food Technol April 1985:49–60

Tidestrom I, Kittell T (1941) A flora of Arizona and New Mexico. The Catholic University of America Press, Washington, D.C., pp 663–666

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

Welsh SL, Atwood ND, Higgins LC, Goodrich S (1987) A Utah flora. Brigham Young University, Provo, Utah, pp 44–45

Wendel JF, Albert VA (1992) Phylogenetics of the cotton genus (*Gossypium*). character-state weighted parsimony analysis of chloroplast-DNA restriction-site data and its systematic and biogeographic implications. Sys Bot 17:115–143

White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of the fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) PCR protocols: a guide to methods and applications. Academic Press, San Diego, Calif., pp 315–324

Wiggens IL (1980) Flora of Baja California. Stanford University Press, Stanford, California, pp 90–92

Williams JT, Brenner D (1995) Grain amaranth (*Amaranthus* species). In: Williams JT (eds) Cereals and pseudocereals Chapman and Hall, London, pp 129–186