2007-07-13

Development and Characterization of Microsatellite Markers for the Grain Amaranths (Amaranthus spp. L.)

Melanie Ann Mallory
Brigham Young University - Provo

DEVELOPMENT AND CHARACTERIZATION OF MICROSATELLITE MARKERS FOR THE GRAIN AMARANTHS (*AMARANTHUS* SPP. L.)

by

Melanie Ann Mallory

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the degree requirements for

Master of Science

Department of Plant and Wildlife Sciences

Brigham Young University

August 2007

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Melanie A. Mallory

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	P. Jeffery Maughan, Chair
Date	Eric N. Jellen
Date	Donald B. Pratt

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Melanie A. Mallory in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and departmental style requirements; (2) its illustrative material including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

_____ _____

Date P. Jeffery Maughan
Chair, Graduate Committee

Accepted for the Department

_____

Loreen A. Woolstenhulme
Graduate Coordinator

Accepted for the College

_____

Rodney J. Brown
Dean, College of Life Sciences

ABSTRACT

DEVELOPMENT AND CHARACTERIZATION OF MICROSATELLITE MARKERS FOR THE GRAIN AMARANTHS (*AMARANTHUS* SPP. L.)

Melanie A. Mallory

Department of Plant and Wildlife Sciences

Master of Science

The grain amaranths (*Amaranthus hypochondriacus* L., *A. cruentus* L., and *A. caudatus* L.) are important pseudocereals native to the Americas that have received increased attention for their nutritional content, specifically their balance of amino acids. The objective of this project was to produce and characterize a set of highly informative, reproducible microsatellite markers for the grain amaranths. A total of 1457 clones were sequenced from three genomic libraries enriched for the microsatellite motifs AAC, AAT and AC. Of these, 353 (24%) contained unique microsatellites. An additional 29 microsatellite loci were identified among 728 BAC-end sequences of a newly developed amaranth BAC library. Flanking primers were designed for 319 of the microsatellite loci and all were screened on a panel of diverse amaranths, including grain and weedy *Amaranthus* species. A total of 179 (56%) microsatellites were polymorphic across

accessions from the three grain amaranths. Among these polymorphic microsatellite loci, a total of 731 alleles were identified with average of four alleles per locus. Heterozygosity values ranged from 0.14 to 0.83 with a mean value of 0.62. Thirty-seven (21%) of the markers were polymorphic between the parents of a segregating population and were shown to be inherited in a normal Mendelian fashion based on chi-squared analysis, demonstrating the utility of these markers for linkage mapping of the amaranth genome. Phylogenetic analysis using the marker data showed *A. hybridus* accessions in two of the three major grain amaranth clades, suggesting the polyphyletic evolution of the three cultivated species from different *A. hybridus* ancestors. The microsatellite markers reported here will be useful for further evaluating the relationships among the grain amaranths and their relatives and are an ideal resource for use in marker-assisted breeding programs, germplasm analysis and varietal identification. The transferability of these markers to *A. hybridus*, *A. powellii*, and *A. retroflexus* as reported here suggests that the markers may be useful to other species with the genus *Amaranthus*, including economically important weeds, vegetable amaranths, and ornamentals.

Key words: amaranth – *Amaranthus* – microsatellites – SSRs – heterozygosity – linkage

ACKNOWLEDGEMENTS

My sincere appreciation is owed to all those who have assisted me in this project. I express my gratitude for my supportive advisor, Dr. Jeff Maughan, and committee members, Drs. Eric Jellen and Donald Pratt, for their expertise, support, patience, and guidance. I gratefully acknowledge Dr. David McClellan for his assistance with phylogenetic analyses. My further appreciation goes out to all of the students who have given their assistance with alacrity in the more tedious aspects of this project: Jimena B. Alvarez, M. Njeri Karanu, Sixto Tuston, Yoselyn Paez, Juan Badal, and Jared Allen. I am also grateful to Jared Benson and Nat Cannon for their assistance with statistical analyses, as well as to Maria Sederberg and Andrea McNabb for their help with the writing process and, more importantly, for their friendship. My special thanks are owed to Rozaura Vivas Hall for being part of this project from the beginning and for sharing so much of her time and herself with me. Finally, I express appreciation for my family, especially for my mother, who has always believed in me more than I believe in myself. This research was funded by a Stephen F. Austin State University faculty research grant, as well as grants from the Ezra Taft Benson Agriculture and Food Institute and the Holmes Family Foundation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Chapter 2

CHAPTER 1: DEVELOPMENT AND CHARACTERIZATION OF MICROSATELLITE MARKERS FOR THE GRAIN AMARANTHS (*AMARANTHUS* SPP. L.)

INTRODUCTION

The grain amaranths (*A. hypochondriacus*, *A. cruentus*, and *A. caudatus*) belong to the genus *Amaranthus* L., which includes 60-70 species (Sauer 1976). The three grain amaranths are classified along with their putative progenitor species (*A. hybridus* L., *A. quitensis* H.B.K., *and A. powellii* S. Wats.) in what is termed the *A. hybridus* complex and are thought to be paleo-allotetraploids ($2n=4x=32$), although chromosome counts of both 32 and 34 have been reported for *A. cruentus* (Pal et al. 1982, Greizerstein and Poggio 1994, 1995). While the grain amaranths have been cultivated for centuries in the Americas, they have been underutilized since the Spanish Conquest when they were replaced by Old World crops and their cultivation suppressed due to their deeply rooted use in indigenous religious practices (Sauer 1976, 1993; Iturbide and Gispert 1994). In the last few decades, the grain amaranths have begun to reclaim some of their importance, largely due to the recognition of the nutritional value of their seed for human consumption (Bressani et al. 1992; Tucker 1986).

Amaranth grain is 50 to 60% starch, with higher fiber (8%) and more fat (7 to 8%) than the grain of most cereals (Pedersen et al. 1987, Breene 1991). They are noted for their relatively high protein content and balance of essential dietary amino acids. Crude protein content from pale-seeded grain types is substantially higher than most cereal grains and has been reported to range from 12.5 to 22.5% on a dry matter basis, with an average of about 15% (Bressani 1989, Breene 1991). Lysine is often the limiting amino acid in other cereal grains; however, amaranth seed protein is rich in this essential amino acid, ranging from 0.73 to 0.84% of the total seed protein content (Bressani et al. 1987). Amaranth oil is also of high nutritional quality, containing a relatively high

content of squalene (7-8%; Bressani et al.1987) and is thought to be effective in reducing cholesterol levels in humans (Berger et al. 2003, Martirosyan et al 2007). The grain amaranths have also been noted for their ability to thrive under extreme abiotic stress (Brenner 2000). They are impressive producers of biomass under warm and dry conditions, an attribute likely related to their C_4 photosynthetic apparatus (Kadereit et al. 2003). Thus, it has been suggested by several researchers that amaranth may be a useful alternative crop in developing nations, especially in overpopulated and undernourished areas (Pal and Khoshoo 1974, Sauer 1993).

The evolutionary origin of the grain amaranths is still under debate, although two hypotheses have been proposed by Sauer (1967, 1976). The first hypothesis is based on geography and suggests that all three grain amaranths evolved independently, while the second hypothesis is based on morphological features and proposes that all three grains are descended mainly from *A. hybridus*. Molecular studies, including analyses with isozymes (Chan and Sun 1997); random amplified polymorphic DNAs (Transue et al. 1994, Chan and Sun 1997); and amplified fragment length polymorphisms (Xu and Sun 2001) have attempted to clarify the relationships among the grain amaranths and their relatives. While these studies support Sauer's second hypothesis of a monophyletic evolution of each of the three grain amaranth species from *A. hybridus*, they have highlighted the need for new methods with greater resolving power to clarify taxonomic relationships within the *A. hybridus* complex.

Microsatellites are short repeated nucleotide motifs usually one to four base pairs in length which are flanked by conserved sequences and occur ubiquitously throughout eukaryotic genomes (Tautz and Renz 1984). They are a widely-considered the genetic

marker system of choice due to their characteristics of being highly reproducible, informative, locus-specific, multiallelic, and codominant (Morgante and Olivieri 1993, He et al. 2002). Microsatellites have been extremely useful in determining taxonomic relationships among closely related individuals and assessing diversity within a species (Ni et al. 2002, Fukunaga et al. 2005, Ellwood et al. 2006). Sun et al. (1999) noted that among probes designed from various types of repetitive sequences, a probe consisting of microsatellite and minisatellite sequence showed the highest polymorphism across the grain amaranths and their close relatives, suggesting that microsatellites may be extremely valuable for characterizing inter- and intraspecific relationships within the *A. hybridus* complex. While the initial cost of developing microsatellites markers is high, once developed these PCR-based markers are inexpensive to use and require less technical expertise relative to other types of molecular markers. Thus, the goals of this project were to i) develop a collection of reproducible microsatellite markers for the grain amaranths, ii) assess the informativeness of these microsatellite markers by screening them against a panel of grain amaranth accessions and iii) use the markers to characterize the relationships of the grain amaranths and their putative ancestors. Moreover, we show the Mendelian inheritance of 37 of these microsatellites in a segregating F_2 population.

MATERIALS AND METHODS

Plant material and DNA extraction. For microsatellite development and characterization, seeds from a total of 35 diverse amaranth individuals, representing ten *A. hypochondriacus* accessions, nine *A. caudatus* accessions, eight *A. cruentus* accessions, five *A. hybridus* accessions, two *A. powellii* accessions and one *A. retroflexus*

accession, were obtained from the USDA collection (USDA, Iowa State University, Ames, IA; Table 1). For linkage analysis, an F_2 population was developed from a cross of 'PI 482049' (*A. cruentus*) and 'PI 477914' (*A. cruentus*). The F_2 population consisted of 92 plants produced by self-fertilizing a single F_1 plant that was kindly provided by David Brenner (USDA, Iowa State University, Ames, IA). All plants were greenhouse grown in Provo, Utah, USA in 15cm (6in) pots using Sunshine Mix II (Sun Grow, Inc., Bellevue, WA) and supplemented with Osmacote fertilizer (Scotts, Marysville, OH). Plants were maintained at 25°C under broad-spectrum halogen lamps with a 12-h photoperiod.

Total genomic DNA was extracted from 30mg freeze-dried leaf tissue according to procedures described by Sambrook et al. (1989) with modifications described by Todd and Vodkin (1996). Extracted DNA was quantified using a Nanodrop (ND 1000 Spectrophotometer, NanoDrop Techonologies Inc., Montchanin, DE) and diluted to 30 ng μl^{-1} in water.

Microsatellite enriched library construction. Four libraries, enriched for microsatellites consisting of AC, AG, AAC or AAT motifs, were produced by Genomic Identification Services, Inc. (Chatsworth, CA) using genomic DNA from *A. hypochondriacus* cv. 'Plainsman.' Genomic DNA was partially digested with a mixture of seven blunt-end restriction endonucleases (*Rsa*I, *Hae*III, *Bsr*B1, *Pvu*II, *Stu*I, *Sca*I, and *Eco*RV). Size-separated DNA fragments ranging from 300 to 750 base pairs were ligated with adapters and enriched for each specific microsatellite motifs using biotinylated capture molecules (CPG Inc., Lincoln Park, NJ). The captured fragments were then

amplified and digested with *Hin*dIII to remove the adaptors and clone the fragments into pUC19. The resulting plasmids were subsequently transformed into competent *E. coli* DH5α cells by electroporation.

Microsatellite identification and classification. Enriched libraries were plated on S-gal (Sigma, St. Louis, MO) agar media supplemented with 100 μg mL^{-1} of ampicillin for blue/white selection of recombinant clones. A total of 1457 recombinant clones were sent to the Arizona Genomics Institute (Tucson, AZ) for plasmid DNA isolation and bi-directional sequencing using M13 primers (Forward: 5'-GTA AAA CGA CGG CCA GT; Reverse: 5'-CAG GAA ACA GCT ATG AC) and standard ABI Prism Taq dye terminator cycle sequencing methodology. The computer program Contig Express (InforMax, Inc., Frederick, MD) was used to determine consensus sequences, eliminate redundant clones, and identify microsatellites.

Microsatellites were classified as either simple or compound and perfect or imperfect based on the classification system given by Weber (1990) with modifications described by Mason et al. (2005). A perfect microsatellite was defined as a stretch of repeats without interruption. A microsatellite was considered imperfect if it consisted of more than one stretch of unbroken repeats where the terminal repeat was at least three full repeats for dinucleotide motifs and at least two full motifs for trinucleotides. Internal repeats had to be at least 1.5 repeats in length and interruptions between internal repeats could not exceed the equivalent of 1.5 motif lengths. A simple microsatellite was defined as consisting of only one motif. A compound microsatellite was defined as stretches of

multiple perfect or imperfect repeats with interruptions of no more than three consecutive base pairs and internal repeats of at least two repeats in length.

Primer design. Primers flanking each unique microsatellite were designed using the web-based computer program Primer3 version 2.0 (Rozen and Skaletsky 2000) according to the program's default parameters, with the following exceptions: preferred product size range equal to 150-200 base pairs; melting temperature differences in forward and reverse primers of no more than 1°C; and max poly-X (maximum allowable length of a mononucleotide repeat) of three. Oligonucleotide primers were synthesized by Integrated DNA Technologies, Inc. (Iowa City, IA). Primer pairs were assigned names based on their repeat motif (e.g. AHAAT035, where AH = *A. hypochondriacus*, AAT = motif type, 035 = clone ID).

Bacterial Artificial Chromosome (BAC) -derived microsatellites. BAC-end sequence (BES) microsatellites were identified using the web-based computer program Tandem Repeats Finder (Benson 1999) and 728 amaranth BAC-end sequences. The sequences were obtained from clones of an *A. hypochondriacus* (cv. 'Plainsman') BAC library developed by Maughan et al. (unpublished). Only sequences with total repeat lengths greater than 20 base pairs (n=10 for dinucleotides; n=7 for trinucleotides, etc.) were selected for primer design using the program Primer3 version 2.0 (Rozen and Skaletsky 2000) as described previously.

Microsatellite amplification. Amplification of microsatellite loci were carried out in 10 µl PCR reactions using HotStarTaq Master Mix (Qiagen Inc., Germantown, MD) and 30 ng genomic DNA according to the manufacturer's recommendation. The thermocycling profile was: 95°C for 15 min followed by 31 cycles of 94°C for 60 s, 56°C for 30 s, 72°C for 60 s, with a final extension at 72°C for 10 min. PCR products were separated on 3% Metaphor agarose gels (Cambrex Bio Science Inc., East Rutherford, NJ), run in 0.5X TBE at 120V for 5 hours, and visualized using ethidium bromide staining with UV transillumination. Microsatellite alleles using this protocol were effectively resolved with a resolution of at least four base pairs, as evidenced by the molecular ladders run on each gel.

Data analysis. The number of alleles and the informativeness for each microsatellite locus was determined by calculating heterozygosity (*H*). For a multiallele system, heterozygosity values can be estimated using the following equation:

$$\mathbf{H = 1 - \sum_{i=1}^{k} P_i^2}$$

where P_i is the frequency of the i^{th} allele and k is the number of alleles (Nei 1978).

Phylogenetic analysis of marker data was performed using the distance (neighbor-joining) method and the full heuristic search option (TBR branch swapping, random addition with ten replications) of the computer program PAUP*4.0 (beta version 4.0b10; Swofford 2002). Robustness of the topology of the cladogram was evaluated by bootstrap analysis (10,000 replicates) of the data set. Only groups with frequencies >50% were retained.

Marker segregation was analyzed for conformation to Mendelian ratios expected in an F_2 population using a chi-squared test, with two and one degrees of freedom for codominant and dominant markers, respectively. Linkage groups were constructed with a minimum LOD score of 3.0 using the default mapping parameters (LOD>1.0, recombination threshold = 0.4, ripple value = 1, jump threshold = 5, Kosambi mapping function) of the computer program JoinMap, version 3.0 (Van Ooijen and Voorrips 2001).

Statistical Models. Statistical analysis of factors contributing to the polymorphism of individual microsatellite markers, measured as the observed number of alleles (ONA) per locus, was performed using stepwise forward selection via the computer program NCSS97 (Hintze 1997). Microsatellites were classified according to i) complexity (simple/compound); ii) type (perfect/imperfect); iii) motif (AAC, AAT, AC, other); iv) total complete repeats (TOTAL); v) longest uninterrupted stretch of tandem repeat (excluding partial repeats) (MAX); vi) total length of repeat including non-repeat bases (LENGTH); vii) non-repeat and half-repeat bases (NON); viii) size of motif (dinucleotide, trinucleotide, etc.); ix) number of terminal repeats (TER); x) number of microsatellites amplified per primer pair (NML); and xi) expected PCR product size (PRO). The numerical classifiers TOTAL, MAX, LENGTH, NON, and PRO were measured in base pairs. Motif, complexity, and type were determined based on the repeat with the largest MAX. Motifs observed less than five times were grouped together in the category "other."

RESULTS AND DISCUSSION

Library analysis. To determine the success of the enrichment process for each of the four microsatellite libraries developed, a small number of clones from each library were sequenced and scanned for microsatellites motifs. The initial estimates of enrichment were 67%, 22%, 50%, and 78% for the AC, AG, AAC and AAT libraries, respectively. Due to the low enrichment estimate of the AG library, it was excluded from further analysis. A total of 1457 clones were sequenced from the remaining libraries, of which 487 clones corresponded to the AAT library, 482 to the AAC library and 488 to the AC library. Of the total clones sequenced, only 31 (2%) of clones failed to produce high quality reads (Phred quality > 20). A high rate of redundancy was observed among the sequences with 938 (64%) sequences being redundant with at least one other sequence in the collection. This redundancy was likely due to the enrichment process that utilized an amplification process after the affinity capture and prior to cloning (Jones et al. 2002). Thus after accounting for redundancy and including 29 microsatellites identified in the BAC-end (BES) library, a total of 382 unique microsatellite-containing sequences remained, including 201 (57%), 69 (20%), and 83 (23%) from the AAT, AAC and AC libraries, respectively. The high frequency of AAT microsatellite sequences derived from the AAT library may either be an artifact of the enrichment process or it may be that AAT repeats are characteristically more frequent in the amaranth genome than are the other types of repeat motifs. Evidence for the latter case is seen when searching non-enriched amaranth DNA sequence (e.g., BES) for repeat motifs. Of the 29 microsatellite repeats identified from the 728 amaranth BAC-end sequences (563 kb) the most common microsatellites identified were all AT-rich, with the most frequently observed motifs

being AT and AAT. Elevated numbers of AT-rich microsatellites have also been observed in many of the other species of the Amaranthaceae family, including sugar beet (Mörchen et al. 1996), spinach (Groben and Wricke 1998) and quinoa (Mason et al. 2005) as well as in several unrelated plant species. Morgante et al. (2002) reported that AT repeats are particularly frequent in non-genic regions of *Arabidopsis thaliana* L., soybean (*Glycine max* L.) and maize (*Zea mays* L.). Indeed, in a recent large-scale development of microsatellites from BAC-end sequences in soybean, the AT motif was the most common of all motif classes observed, while the AAT motif was the most common trinucleotide repeat observed (Shultz et al. 2007).

Marker characterization in the grain amaranths. Of the 382 microsatellite loci identified, we successfully designed flanking primer pairs for 319, including 157, 66, 76 and 20 corresponding to the AAT, AAC, AC, and BES libraries, respectively (Table 2). Flanking primers could not be designed for the remaining 63 microsatellite sequences due to T_m constraints and/or the lack of flanking sequence. All primer pairs were initially screened on an exploratory panel of eight diverse amaranth lines (Table 1). Primer pairs that were monomorphic or failed to amplify on this panel were eliminated from further consideration. A total of 179 microsatellites produced strong amplification products that showed simple and polymorphic banding patterns on the subset of grain amaranth accessions in the exploratory panel. Among these were 97, 30, 39, and 13 markers from the AAT, AAC, AC and BES libraries, respectively (Figure 1). Interestingly, 19 of the polymorphic microsatellites identified in this study amplified two distinct polymorphic bands which appear to represent two independent loci. The amplification of duplicate

loci from a single microsatellite marker has been reported for several polyploid plant species (Röder et al. 1998; Han et al. 2004) where it was suggested that each locus represented orthologous loci derived from the independent ancestral genomes of the polyploid. Such loci in the amaranths, once confirmed via segregation analysis, should prove to be valuable tools in elucidating the paleo-polyploidy event that lead to the evolution of allotetraploid amaranths (Pal et al. 1982, Greizerstein and Poggio 1994, 1995).

To characterize the informativeness of these markers we screened all 179 markers on a larger and more diverse panel of 35 grain and wild *Amaranthus* accessions (Table 1; Figure 2). Marker informativeness was quantified by calculating the observed number of alleles (ONA) amplified per marker and by calculating the heterozygosity (*H*) value associated with each marker (Table 2). ONA and *H*-values were calculated for each grain species separately, for the grain species combined and for the *A. hybridus* accessions alone (Table 3). Limiting the data set to the three grain species (n=27), a total of 731 alleles were observed with an average of four alleles per locus and a range of two to eight alleles observed per locus. Using the *H*-value calculated for the grain species and the thresholds given by Ott (1992) where a marker is considered polymorphic if $H \geq 0.1$ and highly polymorphic if $H \geq 0.7$, all 179 markers were considered polymorphic, with 59 (33%) of the microsatellite loci being highly polymorphic ($H \geq 0.7$). Heterozygosity (*H*) values ranged from 0.14 to 0.83 with an average *H*-value of 0.62 per locus and are similar to those obtained from microsatellite development studies in cultivated relatives of amaranth, including sugar beet (0.61; Rae et al. 2000) and quinoa (0.57; Mason et al. 2005).

When the three grain species were analyzed separately, 129, 123, and 136 microsatellites were polymorphic among *A. hypochondriacus, A. cruentus* and *A. caudatus* accessions respectively (Table 3). *A. caudatus* showed the highest total number of polymorphic microsatellite markers and the highest number of alleles observed (371), while *A. cruentus* showed the lowest genetic diversity of the grain species with only 123 polymorphic markers and only 327 total alleles observed (Table 3). The lower degree of genetic diversity observed in *A. cruentus* is consistent with observations using other types of genetic markers including, RFLP, isozyme, and AFLP (Chan and Sun 1997; Xu and Sun 2001). Chan and Sun (1997) suggested that the decrease in genetic diversity observed in *A. cruentus* may be a result of the domestication process where only a small subset of the wild population was initially subjected to artificial selection for specific agronomic characteristics followed by inbreeding to produce true breeding types. In the case of *A. cruentus*, this domestication process coupled with a limited and uniform cultivation range (Central America) may have further reduced the level of intraspecies variation. Conversely, the varied topography (high plateaus and mountain valleys) and niche cultivation zones characterized by extreme abiotic stresses (drought, frequent frost and saline soils) of Andes may account for the increased genetic diversity seen in *A. caudatus*.

Statistically important factors affecting microsatellite polymorphism. To evaluate the factors that influence the informativeness of a potential marker in amaranth, we used a stepwise-forward selection model and found that the factors MOTIF (AAC, AAT, AC, etc.) and MAX (base pair length of the longest uninterrupted tandem repeat) were the

most significant predictors (P<0.01) of marker polymorphism, measured as ONA. The model explained 32% of the variation of ONA and a T-test analysis showed that AAT repeats have a significantly higher (*P*<0.0001) ONA than other types of repeats, especially when the tandem repeat length is greater than 20 base pairs (Figure 3). These observations correlate well with those observed by others, including Moriguchi et al. (2003) who observed that microsatellites with high tandem repeat numbers have higher polymorphism (ONA) and Mason et al. (2005) who reported that a definite change in the percentage of polymorphic versus monomorphic marker occurs when the tandem repeat length is greater than 20 base pairs. The high rate of polymorphism for the AAT motif in comparison to the other repeat types has also been observed in other plant species. For example, in wheat (*Triticum aestivum* L.), Song et al. (2002, 2005) reported that among trinucleotide repeats, the AAT motif is the most polymorphic and even suggested that it may be superior to dinucleotide motifs. These observations and the relative abundance of the AAT motif observed in the amaranth genome (see above) suggest that future development of microsatellite markers with high polymorphic content should focus on AAT repeats with a tandem repeat greater than 20 base pairs.

Genetic diversity within weedy *Amaranthus* species. *A. hybridus*, a putative wild progenitor species of the grain amaranths, showed the most genetic diversity of all the species included in the complete screening panel. One-hundred and sixty microsatellite markers were polymorphic and 472 total alleles were observed. The fact that >99% of the microsatellite markers (developed from *A. hypochondriacus*) amplified in *A. hybridus* is notable in that it confirms the close ancestry between the grain amaranths and *A.*

hybridus at the DNA level. The higher genetic diversity observed among the *A. hybridus* accessions is consistent with an expectation that a wild progenitor species should be more diverse than a derived domesticated species, due to genetic drift and selection (Hilu 1995).

In addition to being the putative progenitor of the grain amaranths, *A. hybridus* (smooth pigweed) along with several other members of the *Amaranthus* genus including, *A. retroflexus* L. (redroot pigweed) and *A. powellii* (Powell amaranth) are particularly notorious weeds (Wassom and Tranel 2005). Various studies have already demonstrated the utility of molecular markers for correcting taxonomic misclassifications among the weedy species within the *Amaranthus* genus (Wetzel et al. 1999, Wassom and Tranel 2005); however, taxonomic problems still exist, especially for closely related species, and highly polymorphic markers are needed to resolve these taxonomic questions. Such markers would also be beneficial in intraspecies population studies and for establishing the first genetic maps in these species. To determine the transferability and utility of these microsatellite markers in the related weedy species, we evaluated the level of amplification for the 179 polymorphic microsatellite markers in three additional weedy species (*A. hybridus*, *A. powellii* and *A. retroflexus*). As previously noted, 177 (>99%) of the markers amplified in the *A. hybridus* accessions, while 158 (88%) and 141 (78%) of the microsatellite markers amplified in the *A. powellii* and *A. retroflexus* accessions, respectively. Between the two *A. powellii* accessions included in the large screening panel, 97 (52%) markers were polymorphic (Table 2). The high transferability observed in this study demonstrates the utility of these markers as new molecular tools for use across the *Amaranthus* genus.

Evolutionary origins of the grain amaranth species. Neighbor-joining analysis reveals that *A. caudatus*, *A. cruentus*, and *A. hypochondriacus* are monophyletic, while *A. hybridus* is polyphyletic (Figure 4). Sauer (1967, 1976) proposed two hypotheses for the evolutionary origins of the grain amaranths. The first hypothesis is based on geography and suggests that all three grain amaranths evolved independently: *A. caudatus* from *A. quitensis* in the Andean region of South America; *A. cruentus* from *A. hybridus* in Central America; and *A. hypochondriacus* from *A. powellii* in Mexico. The second hypothesis is based on morphological features and proposes that *A. hybridus* gave rise to *A. cruentus*, that introgression of *A. cruentus* and *A. powellii* gave rise to *A. hypochondriacus* and that introgession of *A. cruentus* and *A. quitensis* produced *A. caudatus*.

Our results support *A. hybridus* as the progenitor species of all three grain amaranths, but suggest an alternative hypothesis to explain their origins: multiple, independent domestication events from geographically diverse populations of *A. hybridus*, specifically a domestication event in Central Mexico corresponding to *A. hypochondriacus*, a domestication event in Southern Mexico-Northern Central America corresponding to *A. cruentus*, and a South American Andean domestication event corresponding to *A. caudatus*. The progenitor status of *A. hybridus* is supported by: the observation that *A. hybridus* forms hybrids with all of the other species in the complex (Pal and Khoshoo 1974); the polyphyletic placement of *A. hybridus* accessions within the dendrogram (Figure 4); the failure of either of other two proposed progenitors (*A. retroflexus* and *A. powellii*) to group with the grain amaranths (Figure 4); and the extreme genetic diversity of *A. hybridus* (Table 2).

We also note that in the neighbor-joining tree, accessions of *A. caudatus* and *A. cruentus* split into monophyletic subclades corresponding to New and Old World accessions, whereas New World *A. hypochondriacus* is paraphyletic with respect to Old World accessions. The paraphyletic result is not unexpected considering the biogeography of the grain amaranths, which are all native to the New World and were spread to Asia, Europe, and Africa during post-Colonial American times (Sauer 1967). With increased sampling, we expect that New World accessions will be paraphyletic and that Old World accessions will contain a subset of the genetic diversity found in New World populations.

A larger investigation with a wider sampling of *Amaranthus* species, including numerous accessions of *A. hybridus*, needs to be conducted to further evaluate our proposed alternative hypothesis of the origins of the grain amaranths, as well as the relationship between New and Old World accessions of the grain species. The transferability and highly polymorphic nature of microsatellite markers across the genus make them ideal for such an investigation.

Sequence homology. The nucleotide and amino acid similarities of the 319 unique microsatellite sequences for which flanking primers were designed were compared with sequences in the GenBank database through BLASTN and BLASTX searches using the program blast2GO (Conesa et al. 2005). A total of 34 sequences (11%) had significant similarities (E value ≤ 1.0E-7) to GenBank entries (Table 4). Of these, seven were significant based on their nucleotide sequences, while 21 were significant at the amino acid level. Six sequences had high homology to GenBank entries at both levels. Most

clones showed homology to sequences from relatives, such as *Amaranthus* spp., *Chenopodium quinoa*, and *Beta vulgaris*, or other plant species, including *Arabidopsis thaliana* L., *Medicago truncatula* Gaertn., and *Vitis vinifera* L. At the nucleotide level, many sequences had hits to genomic regions of other plants. Homology to proteins involved in DNA binding, transcription and repair as well as to transporter proteins, transcription factors, and metabolic proteins were also observed. At the amino acid level, seven sequences had significant hits to repetitive elements. Microsatellites have been observed to associate with repetitive elements in other studies. For example, AT repeats were associated with miniature inverted transposable elements (MITES) in rice (Temnykh 2001), while microsatellites frequently associate with retrotransposons in the barley (*Hordeum vulgare* L.) genome (Ramsay et al. 1999).

Mendelian inheritance of microsatellite markers. To evaluate the utility of these markers for future linkage map construction, we investigated the inheritance of the microsatellite loci in a segregating F_2 cross ('PI 482049' X 'PI 477914'; see Materials and Methods). In total, 40 (22%) loci were polymorphic between these parents and screened on 92 individual of the F_2 population. Thirty-four (85%) of the markers were scored co-dominantly (1:2:1), while the remaining six (15%) were scored in a dominant fashion (3:1). Three (AHAAT143, AHAAT144 and AHAC008) of the loci deviated significantly ($P<0.01$) from their expected Mendelian pattern of inheritance based on chi-squared analysis. All three distorted markers had a significantly higher than expected frequency of the paternal allele ($P<0.0001$) and may reflect marker loci linked to genes affecting gametic or zygotic viability (Xu et al. 1997). Linkage analysis, performed using

the program JoinMap (Van Ooijen and Voorrips 2001), identified nine linkage groups consisting of 21 linked microsatellite loci spanning 108 cM (Figure 5). While most linkage groups consisted of only two linked markers, two groups with three and four linked markers were observed. The largest linkage group spanned 29 cM. While this is only an exploratory linkage analysis, it does highlight two important observations regarding the grain amaranths: 1) although the grain amaranths are allotetraploids they show normal amphidiploid inheritance and are thus amenable to linkage map construction and marker assisted breeding and 2) the level of intraspecies polymorphism is limiting (e.g., in this cross only 22% of the markers were polymorphic), suggesting that either additional markers will be needed to develop saturated intraspecies maps or interspecies populations will be needed to augment the level of polymorphism within a single cross. One such interspecific mapping population, currently being constructed at the University of Illinois by Pat Tranel (personal communication), is between 'Plainsman' (*A. hypochondriacus*) and '21605-16' (*A. hybridus*). Preliminary analysis of our dataset shows that nearly a three-fold increase in the total number of markers (119 microsatellite markers) should be segregating in this population as compared to the intraspecific cross described above.

CONCLUSIONS

We report the first large-scale development of microsatellite markers for the grain amaranths. One-hundred and seventy-nine markers were developed and characterized. Thirty-seven of these markers were shown to segregate in a normal Mendelian fashion. These microsatellite markers will be useful for evaluating the relationships among the

grain amaranths and their relatives and are an ideal resource for use in marker-assisted breeding programs, germplasm analysis and varietal identification.

REFERENCES

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27:573–580

Berger A, Gremaud G, Baumgertner M, Rein D, Monnard I, Kratky E, Geiger W, Burri J, Dionisi F, Allan M, Lambelet P (2003) Cholesterol-lowering properties of amaranth grain and oil in hamsters. Int J Vitam Nutr Res 73:39–47

Breene WM (1991) Food uses of grain amaranth. Cereal Foods World 36:426–430

Brenner DM, Baltensperger DD, Kulakow PA, Lehmann JW, Myers RL, Slabbert MM, Sleugh BB (2000) Genetic resources and breeding of *Amaranthus*. Plant Breed Rev 19:227–285

Bressani R (1989) The proteins of grain amaranth. Food Rev Int 5:13–38

Bressani R, Gonzalez JM, Zuniga J, Bruener M, Elias LG (1987) Yield, selected chemical composition and nutritive value of 14 selections of amaranth grain representing four species. J Sci Food Agric 38:347–356

Bressani R, Sanchez-Marroquin A, Morales E (1992) Chemical composition of grain amaranth cultivars and effects of processing on their nutritional quality. Food Rev Int 8:23–49

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

Conesa A, Gotz S, Garcia-Gomex JM, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676

Ellwood SR, D'Souza NK, Kamphuis, Burgess TI, Nair RM, Oliver RP (2006) SSR analysis of the *Medicago truncatula* SARDI core collection reveals substantial diversity and unusual genotype dispersal throughout the Mediterranean basin. Theor Appl Genet 112:977–983

Fukunaga K, Hill J, Vigouroux Y, Matsuoka Y, Sanchez G J, Kenju L, Buckler ES, Doebley J (2005) Genetic diversity and population structure of teosinte. Genet 169:2241–2254

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

Greizerstein EJ, Poggio L (1995) Meiotic studies of spontaneous hybrids of *Amaranthus*: genome analysis. Plant Breed 114:448–450

Groben R, Wrick G (1998) Occurrence of microsatellites in spinach sequences from computer databases and development of polymorphic SSR markers. Plant Breed 117:271–274

Han Z-G, Guo W-Z, Song X-L, Zhang T-Z (2004) Genetic mapping of EST-derived microsatellites from the diploid *Gossypium arboreum* in allotetraploid cotton. Mol Gen Genomics 272:308–327

He C, Poysa V, Yu K (2002) Development and characterization of simple sequence repeat (SSR) markers in *Lycopersicon esculentum* cultivars. Theor Appl Genet 106: 363–373

Hilu KW (1995) Evolution of finger millet: evidence from random amplified polymorphic DNA. Genome 38:232–238

Hintze JL (1997) NCSS 97, Statistical system for windows. Number Cruncher Statistical Systems, Kaysville, Utah

Iturbide GA, Gispert M (1994) Grain amaranths (*Amaranthus* spp.). In: Hernandez-Bermejo JE, Leon J (eds) Neglected Crops: 1492 from a different perspective. FAO, Rome, pp 93–101

Jones KC, Levine KF, Banks JD (2002) Characterization of 11 polymorphic tetra-nucleotide microsatellites for forensic applications in California elk (*Cervus elaphus canadensis*). Mol Ecol Notes 2:425–427

Kadereit G, Borsch T, Weising K, Freitag H (2003) Phylogeny of Amaranthaceae and Chenopodiaceae and the evolution of C4 photosynthesis. Int J Plant Sci 164:959–986

Martirosyan DM, Miroshnichenko LA, Kulakova SN, Pogojeva AV, Zoloedov VI (2007) Amaranth oil application for coronary heart disease and hypertension. Lipids Health Dis 6:1

Mason SL, Stevens MR, Jellen EN, Bonifacio A, Fairbanks DJ, Coleman CE, McCarty R, Rassmussen A, Maughan PJ (2005) Development and use of microsatellite markers for germplasm characterization in quinoa (*Chenopodium quinoa* Willd.). Crop Sci 45:1618–1630

Mörchen M, Cuguen J, Michaelis G, Hänni C, Saumitoue-Laprade P (1996) Abundance and length polymorphism of microsatellite repeats in *Beta vulgaris* L. Theor Appl Genet 92:326–333

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

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nature Genetics 30:194–200

Moriguchi Y, Iwata H, Ujino-Ihara T, Yoshimura K, Taira H, Tsumura Y (2003) Development and characterization of microsatellite markers for *Cryptomeria japonica* D. Don. Theor Appl Genet 106:751–758

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genet 89:583–590

Ni J, Colowit PM, Mackill DJ (2002) Evaluation of genetic diversity in rice subspecies using microsatellite markers. Crop Sci 42:601–607

Ott J (1992) Strategies for characterizing highly polymorphic markers in human gene mapping. Am J Hum Genet 51:283–290

Pal M, Khoshoo TN (1974) Grain amaranths. In: Hutchinson JB (ed) Evolutionary studies in world crops: diversity and change in the Indian subcontinent. Cambridge University Press, UK, pp 129–137

Pal M, Pandey RM, Khoshoo TN (1982) Evolution and improvement of cultivated amaranths. VIII. Induced autotetraploidy in grain types. Z Pflanzenzuch 78:135–148

Pedersen B, Kalinowski LS, Eggum BO (1987) The nutritive value of amaranth grain (*Amaranthus caudatus*). 1. Protein and minerals of raw and processed grain. Qualitas Plantarum 36:309–324

Rae SJ, Aldam C, Dominguez I, Hoebrechts M, Barnes SR, Edwards KJ (2000) Development and incorporation of microsatellite markers into linkage map of sugar beet (*Beta vulgaris* spp.). Theor Appl Genet 100:1240–1248

Ramsay L, Macaulay M, Cardle L, Morgante M, Ivanissevich S, Maestri E, Powell W, Waugh R (1999) Intimate associations of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. Plant J 17:415–425

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

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S and Misener S (eds) Bioinformatics methods and protocols: methods in molecular biology. Human Press, Totowa, NJ, pp 365–386

Sambrook J, Fritsch EE, Maniatis T (1989) Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cld Spring Harbor, NY

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

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

Sauer JD (1993) Amaranthaceae—amaranth family. In: Historical geography of crop plants: a select roster. CRC, Boca Raton, Florida, USA pp 9–14

Shultz JL, Kazi S, Bashir R, Afzal JA, Lightfoot DA (2007) The development of BAC-end sequence-based microsatellite markers and placement in the physical and genetic maps of soybean. Theor Appl Genet 114:1081–1090

Song QJ, Fickus EW, Cegan PB (2002) Characterization of trinucleotide SSR motifs in wheat. Theor Appl Genet 104:286–293

Song QJ, Shi JR, Singh S, Fickus EW, Costa JM, Lewis J, Gill BS, Ward R, Cregan PB (2005) Development and mapping of microsatellite (SSR) markers in wheat. Theor Appl Genet 110:550–560

Sun M, Chen H, Leung FC (1999) Low-Cot DNA sequences for fingerprinting analysis of germplasm diversity and relationships in *Amaranthus*. Theor Appl Genet 99:464–472

Swofford, DL (2002) PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods), Version 4.0b10. Sinauer, Sunderland, Massachusetts

Tautz D, Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res 12:4127–4138

Temnykh S, Declerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S (2001) Computational and experimental analysis of microsatellites in rice (*Oriza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res 11:1441–1452

Todd JJ, Vodkin LO (1996) Duplications that suppress and deletions that restore expression from a chalcone synthase multigene family. Plant Cell 8:687–699

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

Tucker JB (1986) Amaranth: the once and future crop. BioScience 36:9–13

Van Ooijen JW, Voorrips RE (2001) JoinMap 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, The Netherlands

Wassom JJ, Tranel PJ (2005) Amplified fragment length polymorphism-based genetic relationship among weedy *Amaranthus* species. J Hered 96:410–416

Weber JL (1990) Informativeness of human (dC-dA)n · (dG-dT)n polymorphisms. Genomics 7: 524–530

Wetzel DK, Horak MJ, Skinner DZ (1999) Use of PCR-based molecular markers to identify weedy *Amaranthus* species. Weed Sci 47:518–523

Xu Y, Zhu L, Xiao J, Huang N, McCouch SR (1997) Chromosomal regions associated with segregation distortion of molecular markers in F2, backcross and doubled haploid, and recombinant inbred population in rice (*Oryza sativa* L.). Mol Gen Genet 253:535–545

Xu F, Sun M (2001) Comparative analysis of phylogenetic relationships of grain amaranths and their wild relatives (*Amaranthus*; Amaranthaceae) using internal transcribed spacer, amplified fragment length polymorphism, and double-primer fluorescent intersimple sequence repeat markers. Mol Phylogenet Evol 21:372–387

١

CHAPTER 1: TABLES AND FIGURES

Table 1. *Amaranthus* accessions used in the microsatellite assays. The microsatellite preliminary screening panel consisted of samples 3, 4, 8, 13, 15, 20, 27, and 32.

	Name	Geographical Location	Species
1	PI 274279	India[a]	*A. hypochondriacus* L.
2	PI 337611	Uganda[a]	*A. hypochondriacus*
3	PI 477915[b]	India[a]	*A. hypochondriacus*
4	PI 477916	Mexico	*A. hypochondriacus*
5	PI 477917[b]	Mexico	*A. hypochondriacus*
6	PI 511731	Mexico	*A. hypochondriacus*
7	PI 540446	Pakistan[a]	*A. hypochondriacus*
8	PI 558499[bc]	United States, NE	*A. hypochondriacus*
9	PI 619259	Nepal[a]	*A. hypochondriacus*
10	PI 633589	Mexico	*A. hypochondriacus*
11	Ames 5171	Mexico, Morelos	*A. cruentus* L.
12	Ames 5310	Mexico, Sonora	*A. cruentus*
13	PI 482049[b]	Zimbabwe[a]	*A. cruentus*
14	PI 477913	Mexico	*A. cruentus*
15	PI 477914[b]	Mexico	*A. cruentus*
16	PI 566897	United States, AZ	*A. cruentus*
17	PI 628784	Mexico, Puebla	*A. cruentus*
18	PI 628793	Zaire, Shaba[a]	*A. cruentus*
19	Ames 5127	United States, CA	*A. caudatus* L.
20	Ames 15129[b]	Bolivia	*A. caudatus*
21	PI 166045	India[a]	*A. caudatus*
22	PI 175039	India[a]	*A. caudatus*
23	PI 490440	Peru	*A. caudatus*
24	PI 490604	Bolivia	*A. caudatus*
25	PI 490609	Ecuador	*A. caudatus*
26	PI 568132	Bolivia	*A. caudatus*
27	PI 634914[b]	Pakistan[a]	*A. caudatus*
28	PI 500249	Zambia[a]	*A. hybridus* L.
29	PI 605351	Greece[a]	*A. hybridus*
30	PI 632247	United States	*A. hybridus*
31	Ames 23369	Brazil, Goias	*A. hybridus*
32	21605-16[b]	United States, NC	*A. hybridus*
33	Ames 22592	Mongolia[a]	*A. retroflexus* L.
34	PI 572261	Germany[a]	*A. powellii* S. Wats.
35	PI 604671	United States, WA	*A. powellii*

[a]Originated in the Americas although collected in the Old World according to Sauer (1967)
[b]Included in preliminary microsatellite screening panel
[c]cv. 'Plainsman'

Table 2. Amaranth microsatellite marker name, primary motif, complexity, type, primer sequences, expected PCR product size (PRO), observed number of alleles (ONA), and heterozygosity value (*H*), and cross species amplification (CSA).

Marker Name	Primary motif	Complexity	Type	Forward primer (5'-3')	Reverse primer (5'-3')	PRO[a]	ONA	*H*	CSA
AHAAC001	$(CAA)_{13}$	Simple	Perfect	gggtatgaatgtttcgggaat	cttgtcttaacatcccatactgttg	191	4	0.68	R, A, B*
AHAAC005	$(CAA)_8$	Simple	Perfect	ccatcattgcaccaacagaa	tgtcgacgttgtagttgttcg	172	3	0.46	A, B*
AHAAC006	$(GTT)_3(GCT)_2(GTT)_2$	Compound	Perfect	ttgaagattcgatgccacac	cgttcaagaaagcatgtgga	174	2	0.50	R, A, B
AHAAC007[d]	$(CAA)_2A(CAA)CAGAAACAT(CAA)_5$	Simple	Imperfect	caaagccagcaacatcaaga	tctgccacgacaagaaacc	188	2	0.48	R, A, B
AHAAC011	$(GTT)_8$	Simple	Perfect	ccgtctgtgctgtattgagg	ggccacttgggtttattcct	174	7	0.78	–
AHAAC012	$(CAA)_4$	Simple	Perfect	aacaatgaaccagagacaacaa	ttatgttcttgaagtttcccaaa	197	4	0.72	R, A, B
AHAAC014	$(AAC)_{12}$	Simple	Perfect	acccgaatttcctccagtct	ttgttattgttattgttatggctatgg	157	7	0.81	–
AHAAC016	$(AAC)_8$	Simple	Perfect	cctcaacaaatagcagaaatatcaa	tttaaccctaaacatttcccaaa	168	2	0.38	R, A, B
AHAAC017	$(GGTGTTGTT)_2$	Simple	Perfect	aagggcatctatgggacactt	tgatgcaaattgtggatgct	265	4	0.68	R, A, B*
AHAAC018[d]	$(ACA)_2TAT(AAC)_4$	Simple	Imperfect	cccagcagcataagcatt	cttcaactggtaagtggtttctg	280	4	0.53	R, A, B*
AHAAC019	$(AAC)_{11}GAC(AAC)_3$	Simple	Imperfect	tgaccgagccctagagtatga	gttcccgctggagttgatta	177	3	0.61	R, A, B*
AHAAC021	$(CAA)_9$	Simple	Perfect	gagttatggccgaatttcca	ttggtgttgttcaacatttgg	156	5	0.75	R, A, B*
AHAAC025	$(CAA)_7$	Simple	Perfect	cacaccaaccacaccaagaa	gttgtggcacctgtttcctc	213	3	0.60	R, A, B
AHAAC026	$(CAA)_{10}T(AAT)_2A(AAT)AT(AAT)_9$	Compound	Imperfect	ggttgagtgtccttgccttt	ttcaaccacaaggccattag	224	6	0.78	B
AHAAC030	$(CAA)_7TAA(CAA)AAA(CAA)_7C$	Simple	Imperfect	atactaagagcaaggcata	ctcatataggtattctgattatt	210	6	0.80	A, B*
AHAAC031	$(GAA)_9$	Simple	Perfect	ccagaagggtacgaccaaga	aaatgtcgctaaatatatcccactaaa	196	4	0.73	B
AHAAC035	$(GTT)_3GAT(GAGTTTGATGCTTGT)(GTT)G(GTT)CT(GTT)_3(GAGTTTGATGCTTGT)CCTGATTAT(GTT)_4$	Compound	Imperfect	actatactcataggtctcatacg	tcatcattatcaacaacaac	189	3	0.66	R, B
AHAAC036	$(CCACAC)_4$	Simple	Perfect	cctatcttcgcaccagaacc	tcacttatgggtcgggtttc	158	4	0.66	R, B
AHAAC037	$(CAA)_2CAG(CAA)_7$	Simple	Imperfect	cagcaacaacaattgcaaca	tgagattgattgttgtaacctcct	194	4	0.60	R, A, B
AHAAC038	$(TTG)_{10}A(TTG)_5$	Simple	Imperfect	tccaagtagattgattggttatgg	actagaatcgggcagctgaa	150	6	0.76	R, A, B*
AHAAC046	$(AAT)_3(AAC)_8(AGT)_2$	Compound	Perfect	ctgcgtaagcgtgatagtcg	gagacactgggtgaagagtgc	179	4	0.67	R, A, B*
AHAAC048	$(GTT)_3CTT(GTT)_{10}$	Simple	Imperfect	aaacaacatgctttcttgtctaaa	cacagaaaccacagcatcac	227	3	0.51	R, A, B*
AHAAC049	$(TTG)_9$	Simple	Perfect	tcaggagatcaacaattcctctt	tcccaatgtgaaggaaggag	198	4	0.38	R, A, B*
AHAAC050	$(TTG)_7(CTA)_2(TTG)_6$	Compound	Perfect	gcctctttaaatgatgacacttg	aaacgttacatcaatacaatctaacaa	227	5	0.57	B
AHAAC051	$(AAC)_7ATC(AAC)_2$	Simple	Imperfect	cgctagagatgcctgtttcc	acattgcgtacaaccgacct	180	2	0.26	R, A, B*
AHAAC052	$(AGC)_7(AAC)_{10}$	Compound	Perfect	tcaccatcacctaccaccaa	tgtacacaaagggtcccaca	274	3	0.30	R, A, B*
AHAAC053	$(AAC)_{10}$	Simple	Perfect	aaagtgcaacaccagcaatg	cggtttgaatttgatgttgg	160	4	0.64	R, A, B
AHAAC056	$(GTT)_7$	Simple	Perfect	ctcactcttgttgcggcata	gcacgcaacagtaatgacatc	193	4	0.56	R, A, B*
AHAAC062	$(GTT)_2ATT(GTT)_{12}$	Simple	Imperfect	tttgaatttgatgttggtctgc	gcatccactttcaaaggatca	172	4	0.73	R, A, B*
AHAAC065	$TT(GTT)_4$	Simple	Perfect	aaccctaaacattgtcaaccact	tttgggagttgaagaccaaga	310	4	0.72	R, A, B*

Marker Name	Primary motif	Complexity	Type	Forward primer (5'-3')	Reverse primer (5'-3')	PRO[a]	ONA	*H*	CSA
AHAAC066	$(TTG)_6T(CTT)_{13}$	Simple	Imperfect	agggttggattgttgttgttg	tgcagaaggagtggaatgag	188	5	0.69	R, A, B*
AHAAT001	$(TAT)_{21}$	Simple	Perfect	aagggtaaactttgaggctgttt	ttgtatctcggcttcccact	149	6	0.83	R, A, B
AHAAT003	$(ATT)_6ACT(ATT)_5$	Simple	Imperfect	tgcgtctcagaccaaattgt	ttcaacaacgtcagctcgtt	164	6	0.74	A, B
AHAAT004	$(AAT)_{13}$	Simple	Perfect	tagaaggtgtgcccgtcttt	aactccaaccattaggccatt	250	3	0.47	R
AHAAT005	$(AAT)_{17}$	Simple	Perfect	cattctagcagcagcaccaa	cccaccaacttctttgtaaagtct	193	5	0.53	R, A
AHAAT007	$(ATT)_{16}ATC(ATT)_9$	Simple	Imperfect	tcgtgtttggtaatttgaaggtt	tcacactgatgcactcgaca	234	4	0.74	A
AHAAT008	$(AAT)_{16}$	Simple	Perfect	tggagcacatgatcttcaaca	caggctctccagcagtaggt	200	6	0.78	R, A
AHAAT010	$(TAT)_{11}$	Simple	Perfect	ttgtcgatgttcttcctctctaa	ccagcctgcatcttatcaca	184	5	0.62	R, A, B*
AHAAT011	$(AAT)_3AAC(AAT)_7$	Simple	Imperfect	cgttgagtgcgcttcaataa	aacggttggtgttaagttgacc	197	4	0.65	R, A, B*
AHAAT012	$(AAT)_{14}$	Simple	Perfect	ggttggtagaattcagtgttggt	ccaactgacgttacctccaaa	180	6	0.68	R, A, B*
AHAAT013	$(TTA)_3CAT(TTA)_7$	Simple	Imperfect	ctcacgacatcacagggaaa	aatgacacgtggagtgtgga	162	7	0.82	R, A, B*
AHAAT014	$(ATT)_{23}$	Simple	Perfect	tcagaccagaacagaccagatt	cagcttgttgctcgactcc	186	4	0.64	A
AHAAT015	$(AAT)_2AAA(AAT)_8$	Simple	Imperfect	ggtcgagtttctttgccttg	tgtctttcgttgcggttgta	166	4	0.67	R, A
AHAAT016	$(TAA)_{12}TAG(TAA)_2(TGA)_9$	Compound	Imperfect	gcattctgaattgttcgaggt	acgagatgacccttcctcct	185	6	0.74	–
AHAAT017	$(AAT)_{14}$	Simple	Perfect	caaacacatacttagtctttcctgttt	tttcacgtgccattgattatt	194	4	0.66	R, A, B*
AHAAT018	$(TAT)_9$	Simple	Perfect	gaaggagaggaagcccaaac	agacagcctctcacaagaatttg	159	4	0.68	R, A, B*
AHAAT019	$(ATA)_{13}$	Simple	Perfect	tctttacaaagacgacctcaaca	ttggagaaggaaggtgggta	179	7	0.71	R, A, B*
AHAAT020	$(TAT)_4CAT(TCT)_4(TAT)17TTTTT(TTTTTA)_4$	Compound	Imperfect	cgctaggcccatcagttatt	ggagtcggttaccttaggcttt	264	3	0.53	B
AHAAT024	$(TAT)_{24}$	Simple	Perfect	agatgagcacccacaccttt	cccaccaacttcgtgtaaagt	192	4	0.65	R, A, B*
AHAAT025	$(ATA)_5GTA(ATA)_{12}$	Simple	Imperfect	ttcaacaacgtcagctcgtt	tgcgtctcagaccaaattgt	163	3	0.53	R, A
AHAAT027	$(ATT)_{12}$	Simple	Perfect	cggtaacggtttgctatgtg	tcattgtgctaatggctaatca	195	5	0.74	R, A, B*
AHAAT029	$(TAA)_{12}$	Simple	Perfect	ctcgaatttgaagccgtgta	acacaaataagtgtggctaataaaca	191	5	0.68	R, A, B*
AHAAT030	$(ATT)_{11}$	Simple	Perfect	ccagatgccagatgtgctta	ccaaacaaggtcgatttcaga	189	6	0.75	R, A, B*
AHAAT031	$(ATT)_{14}$	Simple	Perfect	ttaacaaggttgcggaatga	ttggatgaaatgctctccaac	159	6	0.66	R, A, B*
AHAAT033	$(ATT)_{17}GTT(ATC)_{15}(ATT)_9$	Compound	Imperfect	aacatagcactcctaccagttagtg	tttattgccttggtgtattatgttt	242	6	0.63	B
AHAAT035	$(TTA)_9$	Simple	Perfect	tgtggtagaatttgatagggtgtg	tgggagtactagactaggagttgtctt	193	3	0.41	R, A, B*
AHAAT036	$(ATT)_{22}$	Simple	Perfect	atgagttgccgacgttcttc	aattaaatgtgtggattatgctatgg	160	6	0.77	–
AHAAT037	$(TAT)_{10}$	Simple	Perfect	ttggagtagttccaaatgcttg	gaagatggtgggtggacatt	190	4	0.53	R, A, B*
AHAAT038	$(AAT)_{34}$	Simple	Perfect	tgaccatctctttctcttctgactt	ccgaaacttcacaccaatctc	200	8	0.82	R, A, B*
AHAAT039	$(ATT)_{14}$	Simple	Perfect	ctcatacacacacatttccttctatt	gcatttggtatgtgttgagagag	193	6	0.55	R, A, B*
AHAAT042	$(ATT)_{13}$	Simple	Perfect	aaccactaggccatcaattcttt	ggcattcaagaagatctgcaa	200	4	0.67	A
AHAAT043	$(AAT)_{14}$	Simple	Perfect	caatgggtgatagttgggatt	caactgctcctatccctggt	172	4	0.47	R, A, B*

Marker Name	Primary motif	Complexity	Type	Forward primer (5'-3')	Reverse primer (5'-3')	PRO[a]	ONA	*H*	CSA
AHAAT044	$(TAT)_{13}$	Simple	Perfect	aaatgagagttcgttatccttatcttt	tcaataggtcacagggcataa	200	6	0.58	–
AHAAT045	$(TAA)_{10}TAG(TAA)_3$	Simple	Imperfect	gggtgtggatgcttacttagttg	aaccctatgtaaaccatgtatcttca	196	5	0.70	–
AHAAT046	$(TAT)_{20}$	Simple	Perfect	ttatcaaccttcactccattcttt	cattataatttgcacaacccaaa	192	5	0.80	R, A, B*
AHAAT047	$(ATT)_{10}(GTT)_{13}(ATT)_8$	Compound	Perfect	tgttagatccacaaataactcgaaa	cccttggtttattattatgaatttgtt	184	6	0.78	R, A, B*
AHAAT048	$(TAT)_8$	Simple	Perfect	gagcaatcgagtacaggaggtt	aagacgatctcatgtaagaatttgtg	191	NT[e]	NT[e]	–
AHAAT050	$(TTA)_{11}$	Simple	Perfect	ggtttgtgggtcttggtttg	agggccgttcctgtaacatt	199	5	0.70	R, B
AHAAT051	$(AAC)_7AGC(AAT)AGT(AAT)_{20}$	Compound	Imperfect	tgtaacactgcgctacaaatca	ccctcagagtttccttcacc	188	8	0.84	R, A, B*
AHAAT053	$(TTA)_{19}T(TTA)_2$	Simple	Imperfect	attccaaccactaggccatt	gcttgtaagcctcggctatt	230	5	0.78	R, A, B*
AHAAT056	$(TAT)_{22}$	Simple	Perfect	tgtgtgggaatgggtttagg	gctgaagttgaagccacctc	156	5	0.69	R, A, B*
AHAAT059	$(AAT)_{10}$	Simple	Perfect	tgtatctgtctgggcttgtca	cagggttaccacgtggattt	181	4	0.69	–
AHAAT060	$(ATT)_3AT(ATT)_{22}AT(ATT)_5$	Simple	Imperfect	ttgtacattttactttctcttcaactt	aacaacttcctgtaaagtcttcttc	450	3	0.59	R, A, B*
AHAAT061	$(TAA)_{14}T$	Simple	Perfect	tgcatttaggaatattatttgacca	cttataacataacgctcctaccactt	186	4	0.68	A
AHAAT062	$(ATA)_{10}(ATG)(ATA)_2(ATG)_4$	Compound	Perfect	ggcccagattgtatctcgact	ttgagggcgatcaacatttc	176	5	0.69	R, A, B*
AHAAT063	$(TTA)_{25}$	Simple	Perfect	tcggaaattagtcggaggttt	cgatgacaattatgtaacccaatg	191	7	0.77	R, A, B*
AHAAT064	$(AAT)_{20}$	Simple	Perfect	ctcagcttgttgctcgactc	ccaacgggtgatagttgaga	156	7	0.82	R, B
AHAAT065	$(TTA)_{14}$	Simple	Perfect	aactccaaccattaggccatt	ggtttcctaacccagccttt	195	2	0.43	–
AHAAT071	$(AAT)_{21}AAA(AAT)_6GAT(AAT)_4GAT(AAT)_{23}$	Simple	Imperfect	tgatgtggtatatggtgaattatcaa	gagttgatgagccacgtcac	300	6	0.78	R, A, B
AHAAT076	$(TAA)_{10}$	Simple	Perfect	acggtcatgtagagtttgacca	ctgagcaaggttcacccatt	189	3	0.63	R, A, B*
AHAAT077	$(ATA)_{13}$	Simple	Perfect	aaccctcaaacaaacactttca	ttcaaattatgctaaatatcggtgt	170	5	0.65	R, A, B*
AHAAT078	$(AAT)_6GAT(AAT)_{22}$	Simple	Imperfect	ttcctacattgatatggataatgc	cattattagtcgtatttgtgtttcatt	300	6	0.75	R, A, B*
AHAAT079	$(ATT)_{11}$	Simple	Perfect	tggttggtcgggaataaatc	atgatgacgtggcaaatgaa	288	4	0.68	A
AHAAT080	$(TTA)_{15}$	Simple	Perfect	cagcattcattgacgcgtat	ctgaaaccgcaaaccttgat	167	5	0.72	R, A, B
AHAAT087	$(ATT)_{21}$	Simple	Perfect	gatgaaggccatcaacaggt	ccatgaaatagaatgcggttta	160	8	0.79	R, A, B*
AHAAT089	$(ATT)_{12}$	Simple	Perfect	atggtttactccatgcacca	aaactaattaataatggcatggtcttt	184	4	0.50	R, A, B*
AHAAT090	$(TTA)_{13}$	Simple	Perfect	tggtattattgttgagaaataagcaaa	gtgggtctcgtttccacact	198	6	0.79	R, B
AHAAT093	$(ATT)_{28}$	Simple	Perfect	gaaggaacctagcggacaag	gggctaccattgagaatttcc	300	5	0.75	R, A, B*
AHAAT095	$(ATT)_{13}$	Simple	Perfect	aaaggtgagcccaatcaaac	aggtcaaacataccctaatacgact	199	6	0.74	R, A, B*
AHAAT096	$(ATT)_{12}$	Simple	Perfect	ttcaattggcatcaaatcca	taggcccatgtcaccatctt	298	3	0.63	R, A, B*
AHAAT097	$(AAT)_{23}$	Simple	Perfect	cagcctggacctttgtcttg	gtgtgttcatccacctcgtg	193	3	0.67	A, B
AHAAT098	$(CAA)_6(TAA)_9$	Compound	Perfect	acagcaacagcaacagcaac	tcattcaaactctcacccaca	167	7	0.79	R, A, B
AHAAT101	$(AAT)_{11}$	Simple	Perfect	ggttctttgcacaattggttt	tgaatagacacaagcatcacaaa	200	3	0.63	–
AHAAT103	$(ATT)_{17}$	Simple	Perfect	cataacgcttctaccactctgtgt	tttgaccctgtctttgtgtttg	192	6	0.80	B

Marker Name	Primary motif	Complexity	Type	Forward primer (5'-3')	Reverse primer (5'-3')	PRO[a]	ONA	*H*	CSA
AHAAT105	$(TAA)_{24}$	Simple	Perfect	aagtgtaacgctgtccgtga	gaagagagatggaggctcaaa	244	4	0.53	R, A, B*
AHAAT106	$(TAA)_{21}$	Simple	Perfect	caacaaagcatcgagttgaca	ttgatgagccgtctactttcc	180	4	0.64	–
AHAAT107	$(TAA)_3AA(TAA)_{12}TGA(TAA)_3$	Simple	Imperfect	cccaccaacttcttgtaaagtttc	cagatcagaccagcccagat	191	6	0.70	R, A, B*
AHAAT108	$(AAT)_9(AGT)_4(AAT)_4$	Compound	Perfect	caaatcatatcacagaccgtatcag	ggtttggtatgatctgctttgg	241	4	0.64	R, A, B*
AHAAT110	$(AAT)_{10}AA$	Simple	Perfect	aaacgtttactcttcgttctctaca	gctgtcgtgtttggttacttg	200	4	0.66	R, A, B*
AHAAT113	$(AAT)_{13}AAC(AAT)_3$	Simple	Imperfect	agggagacacttggctatctg	cgtggctaaatgctttagtactct	200	5	0.66	R, A, B*
AHAAT114	$(TTA)_8(TTG)_8(TTA)_{12}$	Compound	Perfect	tgggtggatttgtgtatttctt	gggccacgataattatacagc	274	4	0.69	A, B
AHAAT115	$(CAA)_6(TAA)_{11}$	Compound	Perfect	tgtgaacttaatgtggtaatcgaaa	ccctttagtgtaaactccattatagca	191	2	0.49	R, B
AHAAT116	$(TAT)_{22}(TAC)_5(TAT)_5TAA(TAT)_6$	Compound	Imperfect	gagagaaggaggagcgtttc	gtcttcttctatgaaatgaattacga	188	6	0.70	–
AHAAT118	$(AAT)_{14}$	Simple	Perfect	ccttcatagaaagtgggctcct	tcgaaaggcttcaaggtgac	194	6	0.73	R, A, B*
AHAAT120	$(ATT)_{24}$	Simple	Perfect	tggcgttgcctacgactta	gggcttgttcccttgtatcc	243	4	0.67	R, A, B
AHAAT123	$(ATT)_2(ATG)_2(ATT)_8$	Compound	Perfect	gccgacattcaaattgcttt	tggtgcttaccatgtagaaacg	194	6	0.65	A, B*
AHAAT125	$(TAA)_{15}$	Simple	Perfect	atctcagggttaccacgtgaa	ttccaatacccaactaccacct	212	7	0.81	A
AHAAT126	$(TAA)_{14}$	Simple	Perfect	acacgaattgcaactttaccttt	tttggaaatccaccagaagc	180	5	0.75	R, A
AHAAT129	$(ATT)_{14}AGTAT(ATT)_9$	Simple	Imperfect	aaagcaccaaaccctaaacc	cgcggatttcacagatacc	166	7	0.82	A, B*
AHAAT131	$(ATT)_{16}$	Simple	Perfect	cctttgaaagttaggattcaagat	agaaccctcagcactcttca	184	6	0.76	R, A
AHAAT132	$(TAA)_{23}$	Simple	Perfect	gcgccacacatgataggtaa	gggtgccacactagaagagg	198	7	0.74	R, A, B*
AHAAT133	$(AAT)_{18}$	Simple	Perfect	acgttcctgccacttgagat	ttgctttctttcctgttcttatttc	184	6	0.66	R, A, B
AHAAT134	$(ATT)_{13}$	Simple	Perfect	atggtttactccatgcacca	aaactaattaataatggcatggtcttt	187	4	0.52	B
AHAAT137	$(ATT)_{13}$	Simple	Perfect	tgggaatatatttacccttgatctg	tcgttgtgtttgttttctgctg	168	4	0.58	R, A, B*
AHAAT141	$(ATT)_{12}$	Simple	Perfect	aagcatgccaaagagtgtttc	ttgccgccacttcattcta	191	4	0.62	R, A, B*
AHAAT142	$(ATT)_{15}$	Simple	Perfect	acgttggaatacccacttctc	aaagagaccagagtaactgtaataccc	234	6	0.71	R, A, B*
AHAAT143	$(ATT)_{13}$	Simple	Perfect	gggctgtctgcttaaatcca	cgggtcagctcaattaggac	233	4	0.67	R, A, B*
AHAAT144	$(ATT)_{17}$	Simple	Perfect	gaggaactgacctccgagtg	tcgagtaattgttcctccttagttt	222	4	0.63	R, A, B*
AHAAT145	$(ATA)_{25}$	Simple	Perfect	ctgtttgtggcagctgtttg	atgccatggcggagtaagt	370	4	0.64	R, A, B
AHAAT147	$(AAT)_6(AGT)_9A(AAT)_{14}$	Compound	Imperfect	aaccagaattatccggatttc	agaataggtagtttctcacaatttctc	194	5	0.67	R, A, B*
AHAAT148	$(AAT)_{13}$	Simple	Perfect	acacctcgccgacatttaac	tgagatttcgggctttactca	192	5	0.70	R, A, B*
AHAAT151	$(TAA)_{14}$	Simple	Perfect	cgtctttacacaagacgatctca	tgggtatgtataatgggagacg	158	6	0.82	–
AHAAT154	$(ATT)_{21}$	Simple	Perfect	tggcaatggtttagtcatcag	tggtttgatgaaatgaaatacga	272	6	0.74	–
AHAAT156	$(ATT)_{14}$	Simple	Perfect	gcatctttactcacagccaatg	tccaaatgtaatcactctatttgctc	152	3	0.59	–
AHAAT157	$(TAT)_{15}$	Simple	Perfect	ccattctatgacatttggtccat	ttattcacagcgaacgagca	199	4	0.65	–
AHAAT158	$(AAT)_{12}$	Simple	Perfect	ttgccgccacttcattcta	aagcatgccaaagagtgtttc	191	4	0.62	R, A, B*

Marker Name	Primary motif	Complexity	Type	Forward primer (5'-3')	Reverse primer (5'-3')	PRO[a]	ONA	*H*	CSA
AHAC002	$(GT)_9$	Simple	Perfect	tatgatgggctgatggcata	agcgaactatccggaaacaa	165	5	0.72	R, A, B*
AHAC003	$(AC)_5TC(AC)_8$	Simple	Imperfect	caagcacgacaggaatttca	tggccattggttaagtgtgt	122	5	0.69	R, A, B*
AHAC007	$(TG)_{11}$	Simple	Perfect	tttggaactattgagatttagttgaaa	cgacaggaatttcaccctct	174	3	0.60	R, A, B
AHAC008	$(GT)_8GA(GT)_5$	Simple	Imperfect	ggtatggcattggttaagtgtg	cctttaagcgctgcaaaca	153	5	0.76	R, A, B
AHAC010	$(AC)_6AA(AC)_{16}AA(AC)_{17}AA(AC)_5(AT)_7$	Compound	Imperfect	ggctcccaagtctcagtgtg	acgtgtgtgagtggaatgtga	282	6	0.78	R, B
AHAC011	$(AC)_8(TCAC)_2(AC)_3(TCAC)_4(AC)_4$	Compound	Perfect	cgactacacaatttctcatagttgg	agggtgatttgttgacctcac	229	3	0.60	R, A, B
AHAC013	$(CA)_{13}AA(CA)TA(CA)_3$	Simple	Imperfect	gcacaaccgtccagactcta	ttaatgggtggttaagtgtgtctt	190	3	0.63	R, A, B*
AHAC014	$(GT)_5AT(GT)_4(GC)_3CT(GT)_5AT(GT)_5$	Compound	Imperfect	acaggtcgtatgtatgcttgtct	ttgatttgagtttagcaatgaca	157	4	0.70	R, A, B
AHAC015	$(AT)_9(GT)_{10}(AT)(GT)_5(GC)_5$	Compound	Perfect	tcgattctcattgcattaaactaca	gcactgaagggcacttgg	196	6	0.71	R, A, B*
AHAC016	$(AC)_{15}(AT)_7$	Compound	Perfect	agataaacattcaattcctttacaca	tttggtcacttgtgaatagtgtttat	187	4	0.64	R, A, B
AHAC019	$(GA)_{23}$	Simple	Perfect	ccacttggctgttcttctaattg	agccacaagaggcaagtacc	200	3	0.66	R, A, B*
AHAC020	$(CA)_4AA(CA)CT(CA)CT(CA)_{11}$	Simple	Imperfect	acttctaatctccatgcttt	aactgtgtcttaatgtgtgtg	150	3	0.44	R
AHAC021	$(TG)_{11}$	Simple	Perfect	tggctgatgccttcaagata	ctagcctcccaacactctgg	200	3	0.45	R, A, B*
AHAC026	$(GT)_8AA(GT)_{10}CT(GT)_2CA(GT)_3TT(GT)_7$	Simple	Imperfect	acgtgcaccaaagcgtaaa	caccgtccttggttgaggta	163	6	0.76	R, A, B*
AHAC028	$(GT)_{12}$	Simple	Perfect	cagaagataatgtgaaagacctatcg	tataccataagcataacaccacacc	150	3	0.61	R, A, B*
AHAC030	$(GT)_{11}$	Simple	Perfect	accaacaatggaggcacagt	tttggtgggtaaactaaattcca	160	3	0.53	R, A
AHAC032	$(GT)_{18}$	Simple	Perfect	ttggcagtacgactgttgct	gcctttccagagccacttc	154	4	0.67	R, A, B
AHAC035	$(CA)_4AC(CA)_{10}$	Simple	Imperfect	cgagcacaacctttcagacc	ggatgttcattcaatctcaaagtg	188	3	0.38	R, A, B*
AHAC036	$(AC)_{11}(AT)_7$	Compound	Perfect	tttgcaacacattgacaatttaata	tggtaaagtgtgttctcatactaacaa	317	3	0.60	R, A
AHAC038	$(TG)_{10}$	Simple	Perfect	ggctcaagtgtcttaggcttg	cgctcatgaatctgagaaacac	197	3	0.55	R, A, B*
AHAC041	$(CA)_{17}(AT)_5$	Compound	Perfect	gtcgttgtacgtgggaatga	cttgaaccaagagcctaacca	197	4	0.64	B
AHAC042	$(CA)_{12}$	Simple	Perfect	gcacaaccgtccagactcta	ttaatgggtggttaagtgtgtctt	188	5	0.55	R, A, B*
AHAC044	$(AT)_{10}(GT)_{49}$	Compound	Perfect	aaatcagtacgtaattcctgatgtaac	ctgtcctagtcgccaatatca	258	4	0.74	R
AHAC045	$(AC)_{11}$	Simple	Perfect	gcacgacaacccaatacaca	aatagtgcgttgatggcaca	156	5	0.73	R, A, B*
AHAC046	$(TG)_8TA(TG)_2$	Simple	Imperfect	tttgtgtgcatcattcaagtca	ctcatgaactcaatacccaatcc	157	2	0.41	R, A, B
AHAC048	$(TG)_9$	Simple	Perfect	aaacacgttctaatacgtgatcatt	ttacgatggcatcaacaagaa	188	3	0.54	R, A, B
AHAC049	$(AC)_{10}$	Simple	Perfect	ctgcaaacagccaatcagag	tgctcatattgaccggaactt	188	4	0.61	R, A, B
AHAC050	$(AT)_3(GT)_7(AT)(GT)_3$	Compound	Perfect	ccatttatgttgcgggtca	ttcataagtcatggagtaacaacca	180	5	0.49	R, A, B*
AHAC055	$(CA)_6CG(CA)TG(CA)_7$	Simple	Imperfect	tccaagatacctcataaacacttg	ttgccacgatatgtatatctcaa	193	3	0.32	R, A, B*
AHAC057	$(AC)_5ATGA(AC)_5$	Simple	Imperfect	aggctcattgcaaacataagagt	cgtttgtccatcatgcattc	169	3	0.47	R, A, B*
AHAC058	$(AC)_7A(AC)A(AG)_3(AC)_3A$	Compound	Imperfect	tgcatacttgctttgcagga	cgttgctttagcaagctggt	173	2	0.12	R, A, B
AHAC060	$(CA)_8TA(CA)_5$	Simple	Imperfect	agaattaggagataaatgtgctctaaa	tgattccaaattctgatttcaa	200	3	0.47	R, A, B

Marker Name	Primary motif	Complexity	Type	Forward primer (5'-3')	Reverse primer (5'-3')	PRO[a]	ONA	*H*	CSA
AHAC062	$(AC)_{11}(AGACACAC)_2(AC)_5(AGACACAC)_4(AC)_4$	Compound	Perfect	ggctcccaagtcacagtgtt	tcatctttatcgttgattcgtttc	189	8	0.75	R, A, B*
AHAC064	$(AT)_{11}(GT)_{26}$	Compound	Perfect	gggcttccaagtttcagtgt	tccttatctcaagcatcaccttt	183	8	0.76	R, A, B*
AHAC066	$(AT)_4(GT)_{13}$	Compound	Perfect	aacttgtaattatgtgtacgatgaaga	gcacatacatatacagacacacaaac	163	6	0.67	R, A, B*
AHAC067	$(CA)_4A(CA)_{17}$	Simple	Imperfect	cataaactatatgttagaacagc	taaagttgctctactttcct	165	5	0.71	R, A, B*
AHAC070	$(GT)_8G(GT)_2$	Simple	Imperfect	tgggtggttaaatgtgcctta	gcacaaccttccagacccta	177	2	0.44	A, B*
AHATG001	$(ATG)_6$	Simple	Perfect	ttggatttcttgcaggttcc	tgagtgtgcgggagaggtat	186	3	0.49	R, A, B
AHGA002	$(CT)_{13}$	Simple	Perfect	tctctttcatctcaaactaactcact	ttgtacttgaatccacaacttcg	238	3	0.55	R, A, B
AHGA004	$(CT)_{20}TT(CT)_7CC(CA)_{14}$	Compound	Imperfect	cctgcgtgtacttatgaaggtg	tccaagatgcttcatttcca	177	4	0.64	R, A, B*
BAHAAC066[d]	$(GGT)12GG(GGT)_8$	Simple	Imperfect	gggaaaggaaatcataatcca	gctgcaactgttgtatatttgg	199	4	0.60	R
BAHAAT160[d]	$(TTA)_8$	Simple	Perfect	ggcatcaaactgctccattt	aattctgacctttcctttcatca	188	5	0.75	R, B
BAHAAT162[d]	$(ATA)_{13}$	Simple	Perfect	agggcaagttcatttgaagag	aagaaaccaaactatacaagagacaaa	164	7	0.78	R, A, B*
BAHAAT163[d]	$(AAT)_7$	Simple	Perfect	tttgtctcttgtatagtttggtttctt	acataatacaccgaggcaagg	150	4	0.70	–
BAHAT002[d]	$(AT)_{10}$	Simple	Perfect	cagatccacgtgctattgatg	tttgagggtttcatgttgga	197	5	0.69	R, A, B
BAHAT003[d]	$(TA)_{22}$	Simple	Perfect	tagccaaattggttctcacg	cagcagcttctgttggtgtt	170	8	0.73	R, A, B*
BAHAT004[d]	$T(AT)_{10}$	Simple	Perfect	tgagatgctcatcactatcaagtt	aaagagtccatattcagatgacatt	231	4	0.68	R, B
BAHAT005[d]	$(TA)_{29}$	Simple	Perfect	aatgtcatctgaatatggactcttt	ccttcttattgaatttctacccata	243	5	0.75	–
BAHAT006[d]	$(TA)_{10}$	Simple	Perfect	tggttgatggtatgttggttg	aacgggtagtgaaccaaagc	188	3	0.58	R, A, B*
BAHAT007[d]	$(AT)_{29}$	Simple	Perfect	cgttagcaactcagcatcacc	caagaaatagcgcgacaaca	201	9	0.76	R, A, B*
BAHATTTT001[d]	$(ATTTT)GTTTT(ATTTT)_3$	Simple	Imperfect	ctaacatgggaaagggctga	aaattacccgcaacctgttaaa	150	3	0.58	R, A, B*
BAHGAA003[d]	$(GAA)_7$	Simple	Perfect	gatttgaaagagacatggattgg	ctttcatttcatttagcaattagca	195	5	0.67	R, A, B*
BAHTTGGG001[d]	$(TTGGG)_5(TCGGG)_2$	Compound	Perfect	gtgaaggagaccggctgtt	aaacttgtttcggtagcacga	278	3	0.57	R, A, B

[a]Reported in base pairs

[b]R = successful amplification in *A. retroflexus* 'Ames 22592'; A = successful amplification in *A. powellii* accession 'PI 572261'; B = successful amplification in *A. powellii* accession 'PI 604671'

'*' = polymorphism detected among the *A. powellii* accessions (A, B)

[c]Not a true microsatellite

[d]BAC-end sequence derived microsatellites

[e]NT = Microsatellites missing marker data and therefore not included in the analysis

Table 3. Analysis of marker results by species and by all three grain types combined, including total number of polymorphic microsatellites, observed number of allele (ONA) range and average, total alleles observed, heterozygosity value (*H*) range and average, and total highly polymorphic microsatellites.

	All grains[a]	***A. hypochondriacus***	***A. cruentus***	***A. caudatus***	***A. hybridus***
Sample Size	28	10	9	9	5
Polymorphic microsatellites	179	129	123	136	160
ONA range	2 to 8	2 to 5	2 to 6	2 to 6	2 to 5
Average ONA	4	3	3	3	3
Total Alleles	731	344	327	371	472
***H* Range**	0.14 to 0.83	0.18 to 0.74	0.12 to 0.78	0.10 to .77	0.18 to .80
Average *H*	0.62	0.49	0.49	0.50	0.56
Highly Polymorphic[b]	59	4	6	8	29

[a]Includes *A. hypochondriacus*, *A. cruentus*, and *A. caudatus* accessions
[b]$H \geq 0.7$

Table 4. Significant protein and DNA sequence homologies (with GenBank) to microsatellite-containing clones used to assess diversity among 36 *Amaranthus* accessions.

Microsatellite ID	E-value	Nucleotide homology	E-value	Protein homology	Organism matched	GenBank accession #
AHAAC006	1.0E-18	whole genome shotgun sequence	1.0E-18	AtGRF8 (GROWTH-REGULATING FACTOR 8)	*Vitis vinifera* *Arabidopsis thaliana*	AM470232.1 NP_194146.1
AHAAC007			1.0E-25	probable transposase - soybean transposon mariner element Soymar1-related	*Arabidopsis thaliana*	ABD32675.1
AHAAC013			1.0E-17	CDC45 (CELL DIVISION CYCLE 45)	*Arabidopsis thaliana*	NP_189146.1
AHAAC025	1.0E-07	clone mth2-25a12, complete sequence	1.0E-15	transcription factor	*Medicago truncatula* *Arabidopsis thaliana*	AC152057.11 NP_188034.1
AHAAC039	1.0E-10	whole genome shotgun sequence, contig VV78X090243.7, clone ENTAV 115	1.0E-67	protein binding	*Vitis vinifera* *Arabidopsis thaliana*	AM469471.1 NP_196819.1
AHAAC054			1.0E-11	RNA-directed DNA polymerase (Reverse transcriptase)	*Medicago truncatula*	ABD33261.1
AHAAC058			1.0E-13	Histone-fold	*Medicago truncatula*	ABE92437.1
AHAAT014	1.0E-25	AHAF000132*Amaranthus* hypochondriacus betaine aldehyde dehydrogenase (ahybadh4) gene, complete cds			*Amaranthus hypochondriacus*	AF000132.1
AHAAT017	1.0E-13	Medicago truncatula clone mth2-12a23, complete sequence	1.0E-32	carbohydrate transporter/ sugar porter/ transporter	*Medicago truncatula* *Arabidopsis thaliana*	AC146558.22 NP_177937.1
AHAAT032	1.0E-48	Chenopodium quinoa clone QAAT050 SSR marker sequence			*Chenopodium quinoa*	DQ462137.1
AHAAT034			1.0E-11	hypothetical protein MtrDRAFT_AC140030g15v1	*Medicago truncatula*	ABE80451.1
AHAAT038			1.0E-10	hypothetical protein MtrDRAFT_AC149038g20v1	*Medicago truncatula*	ABD32857.1
AHAAT039			1.0E-12	IMP dehydrogenase/GMP reductase	*Medicago truncatula*	ABE93135.1
AHAAT042			1.0E-16	non-LTR retroelement reverse transcriptase-like protein	*Arabidopsis thaliana*	BAB08270.1
AHAAT045	1.0E-65	Beta vulgaris subsp. vulgaris mitochondrial genomic DNA, complete sequence			*Beta vulgaris*	BA000009.3
AHAAT052			1.0E-25	putative non-LTR retroelement reverse transcriptase	*Arabidopsis thaliana*	AAC63678.1
AHAAT062	1.0E-68	Chenopodium quinoa clone QAAT050 SSR marker sequence	1.0E-21	RNA-directed DNA polymerase (Reverse transcriptase)	*Chenopodium quinoa* *Arabidopsis thaliana*	DQ462137.1 NP_565142.1
AHAAT077			1.0E-10	RNA binding / nucleic acid binding	*Arabidopsis thaliana*	NP_189242.1
AHAAT087	1.0E-11	Chenopodium quinoa clone QAAT024 SSR marker sequence			*Chenopodium quinoa*	DQ462136.1
AHAAT089			1.0E-24	catalytic/ hydrolase	*Arabidopsis thaliana*	NP_179939.1

Microsatellite ID	E-value	Nucleotide homology	E-value	Protein homology	Organism matched	GenBank accession #
AHAAT106			1.0E-25	putative non-LTR retroelement reverse transcriptase	*Arabidopsis thaliana*	AAC63678.1
AHAAT119	1.0E-07	Glycine tomentella clone gtt1-310B2, complete sequence			*Glycine tomentella*	AC195450.5
AHAAT136			1.0E-13	RNA-directed DNA polymerase (Reverse transcriptase)	*Medicago truncatula*	ABE89954.1
AHAAT142			1.0E-15	putative non-LTR retroelement reverse transcriptase	*Arabidopsis thaliana*	AAC67331.1
AHAAT149	1.0E-46	*Amaranthus* caudatus agglutinin gene, complete cds			*Amaranthus caudatus*	AF401479.1
AHAC017			1.0E-23	Polynucleotidyl transferase, Ribonuclease H fold	*Medicago truncatula*	ABD33245.1
AHAC018			1.0E-24	PREDICTED: similar to LINE-1 reverse transcriptase homolog	*Canis familiaris*	XP_851237.1
AHAC022			1.0E-14	unnamed protein product	*Homo sapiens*	BAC85286.1
AHAC028	1.0E-10	SOL133751Spinacia oleracea mRNA for chloroplast ribosome recycling factor			*Spinacia oleracea*	AJ133751.1
AHAC030			1.0E-29	Integrase core domain containing protein	*Solanum demissum*	ABI34329.1
AHAC036			1.0E-18	hypothetical protein MtrDRAFT_AC148764g10v1	*Medicago truncatula*	ABE88260.1
AHAC037			1.0E-22	PREDICTED: similar to LINE-1 reverse transcriptase homolog	*Canis familiaris*	XP_535099.2
AHAC068			1.0E-35	PREDICTED: similar to LINE-1 reverse transcriptase homolog isoform 1	*Canis familiaris*	XP_537276.2
AHGA004	1.0E-18	Vitis vinifera, whole genome shotgun sequence, contig VV78X102968.6, clone ENTAV 115	1.0E-16	AChain A, X-Ray Crystal Structure Of Leacx1, An Acyl-Coa Oxidase From Lycopersicon Esculentum (Tomato)	*Vitis vinifera* *Lycopersicon esculentum*	AM475004.1 2FON

Figure 1. Distribution of the 319 microsatellite primer pairs developed in this study according to library and classified as polymorphic, monomorphic, or having poor amplification. AAT, AAC, AT libraries were enriched for microsatellites, whereas the BES library represents microsatellites identified in amaranth BAC-end sequences.

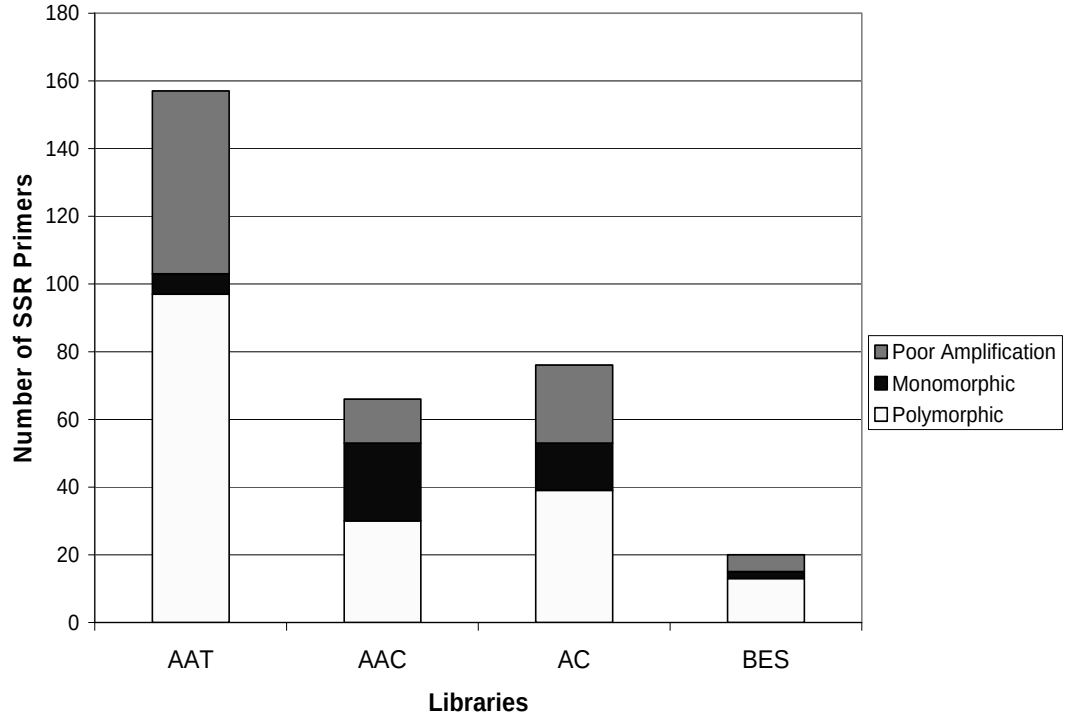


Figure 2. Gel showing polymorphism of microsatellite marker AHAAC030 across all 35 individuals, including individuals from all 3 grain species and from *A. hybridus*, *A. retroflexus* and *A. powellii*. R = *A. retroflexus* 'Ames 22592'; A = *A. powellii* accession 'PI 572261'; B = *A. powellii* accession 'PI 604671'

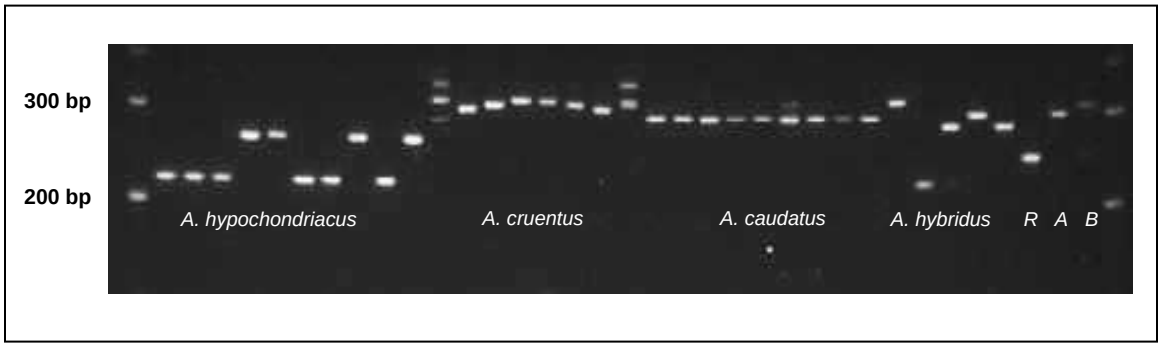


Figure 3a. Distribution of microsatellite markers classified as either polymorphic or monomorphic. Polymorphic markers become prevalent as compared to monomorphic markers when the length of the repeat is greater than 20 bp.

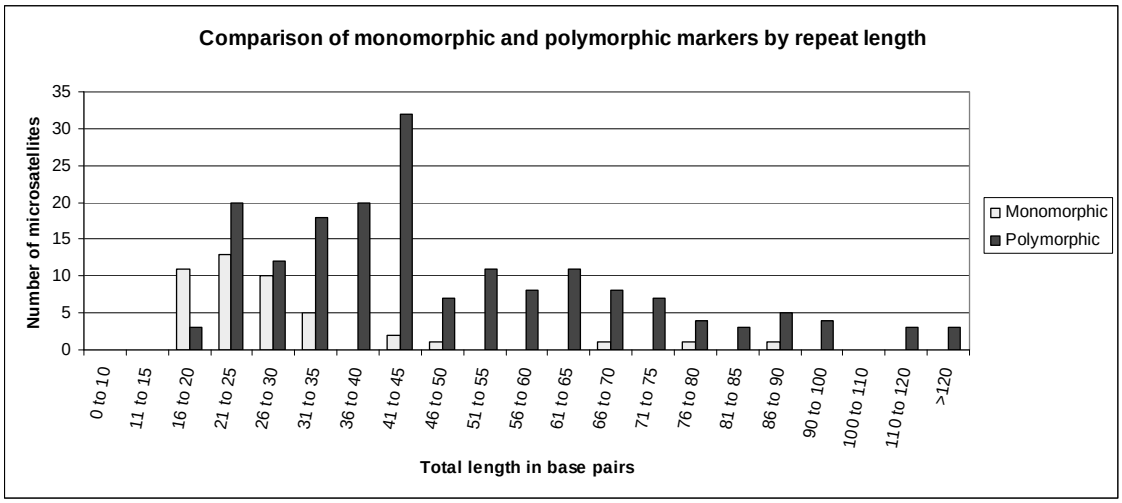


Figure 3b. Distribution of polymorphic microsatellite markers color-coded according to heterozygosity (*H*) vlaue. Again, polymorphic markers are observed frequently for repeat lengths greater than 20 bp.

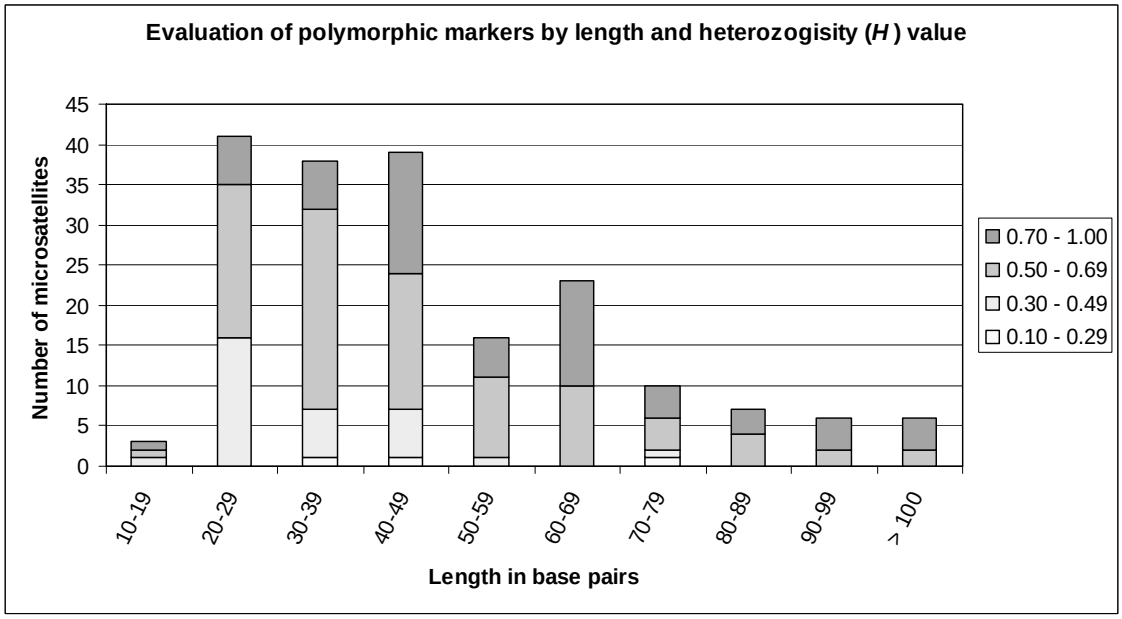


Figure 4. Neighbor-joining analysis of amaranth accessions based on microsatellite data set. Bootstrap support values are given above branches. Individuals in the tree are identified by their abbreviated species (A. hypo = *A. hypochondriacus*, A. caud = *A. caudatus*, A. hybr = *A. hybridus*, A. crue = *A. cruentus*, A. retro = *A. retroflexus*, A. pow = *A. powellii*), panel number and geographic location.

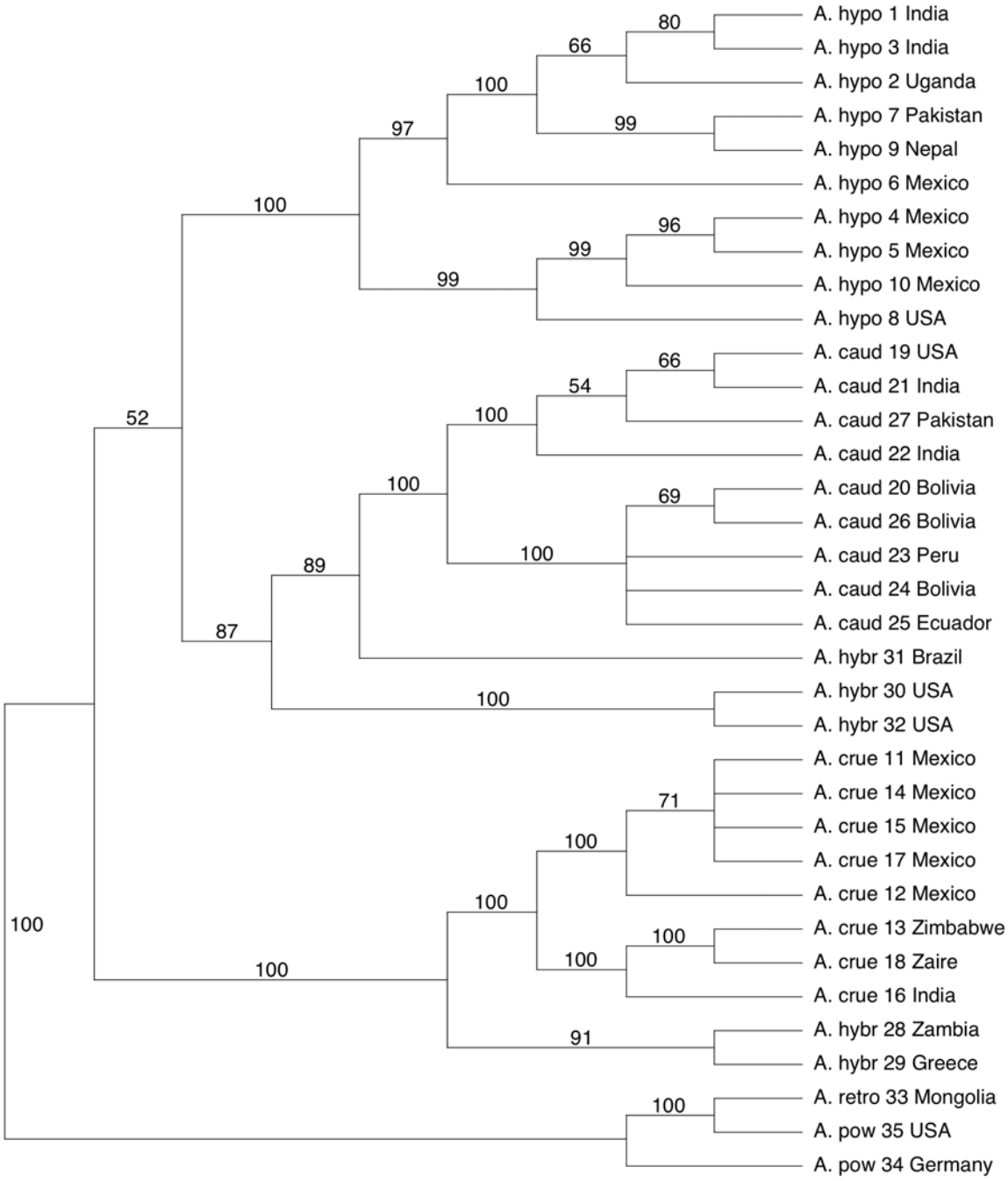
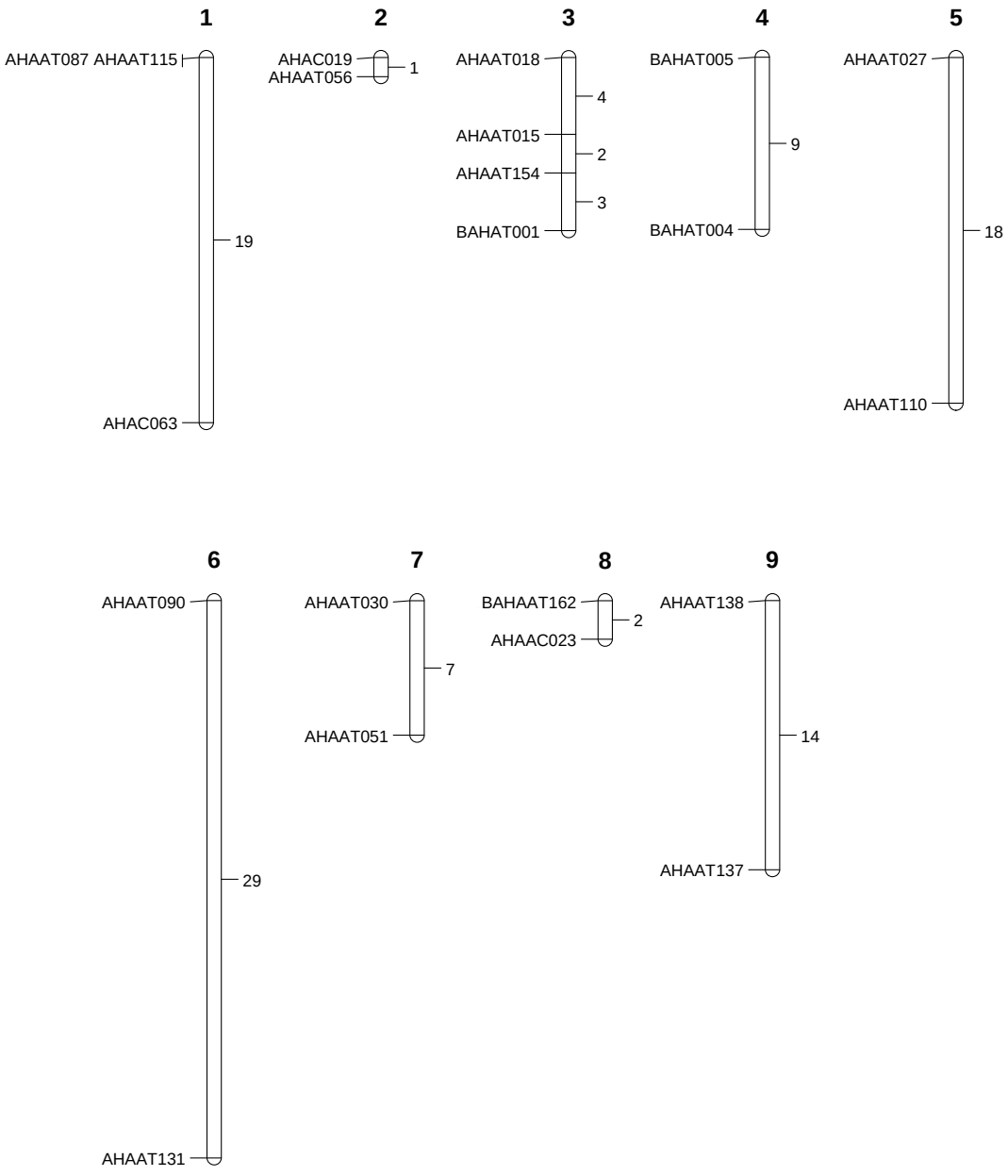


Figure 5. Twenty-one SSR markers incorporated into linkage groups constructed using JoinMap (Van Ooijen and Voorrips 2001) and spanning 108 cM.



CHAPTER 2: LITERATURE REVIEW

INTRODUCTION

Recently, the pseudocereal crop known as amaranth has received attention because it contains many of the nutrients essential for the human diet (Bressani et al. 1992; Tucker 1986). It is further praised for its tenacity under harsh growing conditions and its variety of uses (Gupta and Gudu 1991, Breene 1991). Thus, some believe the study and improvement of amaranth may be useful as an option for alleviating hunger in developing nations, especially in overpopulated and undernourished areas (Pal and Khoshoo 1974, Sauer 1993).

The grain amaranths (*Amaranthus hypochondriacus* L., *A. caudatus*, and *A. caudatus*) are indigenous to Mexico, Central and South America. Thus, the improvement of this understudied crop may benefit both the health and economy of the impoverished regions to which it is indigenous. Improvement of amaranth may be accomplished through supplementing careful plant breeding with research geared toward understanding of the genetic makeup of the plant. Microsatellite markers are an especially useful genetic marker system in that they are highly reproducible, informative, locus specific, multiallelic and codominant, in addition to being inexpensive and easy to use once developed.

Therefore, the development and characterization of this set of microsatellite markers will be useful for studying the genome of the grain amaranths. Diversity analysis using SSRs will shed new light on the currently debated hypotheses of the origin of the grain amaranth species. These markers will be utilized in amaranth linkage mapping, breeding, trait analysis, etc. They may also prove useful to scientists studying closely related weedy and wild *Amarnathus* species (*Amaranthus* spp. L.). An in-depth literature review of the grain amaranths and an introduction to microsatellite markers follows.

GRAIN AMARANTHS

History

Early History. Amaranth is among the oldest of America's crops (Sauer 1950, 1967). Archaeological evidence suggests that grain amaranths have been cultivated in Mexico since as early as 5000 B.C.E. The earliest known record of the pale-seeded, cultivated grain form of *A. cruentus* was discovered in Tehuacan, Puebla, Mexico and is dated at about 4000 B.C.E. However, because *A. cruentus* is not native to this region, it is likely that it was introduced after being cultivated elsewhere earlier (Sauer 1976).

Pre-Columbian History. Amaranth was cultivated for centuries by the indigenous people who resided in present-day Latin America. Amaranth was sacred to the Aztecs and was also cultivated by the Mayas, Incas and other Pre-Columbian civilizations. In addition to cultivating it for food uses, these peoples also bred the plant for its rich color, which was used as dyes in religious rites. As a grain crop, the indigenous peoples prepared and cultivated amaranth in a manner similar to that of maize (*Zea mays* L.). The seeds were ground into flour and used to make tortillas, tamales, etc. (Iturbide and Gispert 1994). It was also popped, parched or made into gruel (Sauer 1967). *Huatli*—as the Aztecs collectively termed amaranth and its relative quinoa (*Chenopodium quinoa* Willd.)—was important in Aztec ritual worship as well as in their diets. The Spaniards recorded that Moctezuma II required a tribute of approximately 200,000 bushels per year of *huatli*, an amount nearly equal to the maize tribute (about 280,000 bushels) (Sauer 1950, 1967, 1993).

As part of their daily diets, the Aztecs prepared a drink called *atole* by mixing amaranth flour with water. A dough, known as *tzoalli*, was made from amaranth flour and

honey extracted from the maguey (*Agave americana* L.*)* plant. According to the month of the year, the dough was made into small pyramids or was formed into the shapes of Aztec deities, such as their god of war Huitzlipochtli. The dough idols were worshipped and then broken into pieces and eaten by the participants. The Spanish Catholics who witnessed these practices considered them a satanic facsimile of the Christian Eucharist (Iturbide and Gispert 1994). Thus, because of the deeply rooted use of *huatli* in Aztec religious practices, the Spaniards, who viewed the grain as a symbol of paganism, suppressed its cultivation (Sauer 1976, 1993). In addition to religious reasons, the reduction in amaranth cultivation in post-Columbian Mesoamerica and South America is also attributed to its replacement by Old World grains and to a possible dislike for its flavor (Iturbide and Gispert 1994).

Amaranth outside of the Americas. After the Spanish conquest, amaranth cultivation dwindled in the Americas. However, the Spaniards introduced the crop into Europe where, by the 18th century, all three grains were widely distributed. A dark-seeded form of *A. cruentus* was introduced into Africa during the 1800s. In West Africa, it became an important vegetable crop. It was also found and grown in gardens in various parts of Asia. Once introduced into the Old World, the common ornamental form of *A. caudatus* became popular and is now cosmopolitan; one popular ornamental form is known by the common name "love-lies-bleeding." Both dark- and pale-seeded *A. hypochondriacus* were introduced into Europe. Unlike its relatives, *A. hypochondriacus* varieties continued to be cultivated as a grain outside of the Americas in India and a few places in China. The Indians and Chinese adopted the light-seeded types of *A. hypochondriacus* during the 1700s. These have now become widespread throughout India where they are sometimes sewn with *A. caudatus*. In

the middle of the 20th century, *A. hypochondriacus* was introduced in Kenya (Sauer 1993). Although amaranth grain use has spread since the Spanish Conquest, amaranth has remained essentially neglected.

Amaranth in the Modern Era. Lehmann (1996) notes two milestones in amaranth's history in the modern era that revived amaranth from its dormancy: 1) Sauer's (1950) rediscovery of the sacred Aztec crop and 2) the discovery of its advantageous protein content by Dowton (1972). Following these discoveries, the cause of amaranth was taken up by Robert Rodale. Rodale, a well-to-do American publisher and philanthropist, who advocated preventive medicine and organic gardening, was intrigued by amaranth's favorable properties and inadvertently became its "germplasm collector, 'public' breeder and promoter" (Lehmann 1996). Through his magazine *Prevention and Organic Gardening* and other publications, Rodale disseminated information on amaranth to production farmers as well as backyard gardeners (Lehmann 1996). This introduction of amaranth as an American crop is singular when contrasted with more common procedures for the introduction of new crops by land-grant universities or via the plant introduction system (Lehmann 1996).

Through his Rodale Research Institute, Rodale procured 1,200 accessions for an amaranth germplasm (Kauffman and Reider 1986) and developed breeding techniques along with advanced breeding lines which he donated to public institutions. The Institute was also instrumental in orchestrating several major research projects on amaranth in various countries, including Mexico, Guatemala, Peru, Thailand and Kenya and in convening national meetings for amaranth farmers, researchers, and advocates (Lehmann 1996).

Following the impetus Rodale provided, amaranth production began to rise in the US in the late 1970s and increased throughout the 1980s. In the 1990s, however, production rates of amaranth fluctuated (Brenner et al. 2000). The 1980s and 1990s have seen much improvement in the understanding of grain amaranths, especially in the areas of its nutritive value and production requirements. In addition, improved cultivars have been released, although much room for improvement remains (Brenner et al. 2000, Brenner 2002).

Taxonomy

The grain amaranths belong to the family Amaranthaceae, which contains 169 genera and approximately 2400 species. The most abundant species in the Amaranthaceae family are herbs that colonize shorelines and other open habitats. A few of the genera are cultivated as ornamentals such as *Celosia*, *Iresine*, and *Gomphrena*, known by the common names cockscomb, bloodleaf, and globe amaranth respectively. The grain amaranths are found within the genus *Amaranthus*. Other relatives in the Amaranthanaceae family that are cultivated as crops are from the group of plants formerly known as the family Chenopodiaceae—such as beets, sugar beets (*Beta vulgaris* L.) and quinoa.

Amaranthus genus. There are 60 to 75 species in the genus *Amaranthus*, sixty of which are native to the Americas while another 15 are indigenous to Africa, Asia, Australia and Africa. These are found mainly in the world's temperate and tropical climates (Sauer 1967). The genus is generally separated into three subgenera: 1) *Albersia*, 2) *Acnida*, the dioecious amaranths and 3) *Amaranthus*, which includes the *A. hybridus* complex. The *A. hybridus* complex consists of the three grain types and their 3 putative progenitors. The majority of

the species are wild or weedy. *Amaranthus* species grow best in desert washes, lakeshores, marshes, ocean beaches, and stream banks. Their seeds are naturally dispersed to these habitats by migratory birds that feed on them (Sauer 1967, 1993). *A. hybridus* is also known as smooth pigweed and is considered a particular notorious weed, along with several other members of the *Amaranthus* genus such as waterhemp (*A. tuberculatus* (Moq.) Sauer), redroot pigweed (*A. retroflexus*), and Powell amaranth (*A. powellii* S. Wats) (Wassom and Tranel 2005). Many of these species are rapidly evolving herbicide resistance (Patzoldt et al. 2006).

Vegetable species. No clear separation between vegetable and grain species exists, because the leaves of young grain varieties may be used as potherbs. *A. cruentus*, as previously noted, is cultivated as both a vegetable and a grain. Its relatives *A. tricolor*, *A. dubius*, and *A. lividus* are also grown as vegetables (Stallknecht and Schulz-Schaeffer 1993). Species are mostly grown as potherbs in India, the East Indias, Southeast Asia, and the Far East. In English, they are known by the common names Chinese spinach, Malabar spinach and tampala (Sauer 1993).

Grain species. As previously noted, the three amaranths principally grown as grains consist of *A. cruentus*, *A. caudatus*, and *A. hypochondriacus*. In the literature, *A. edulis*, which is grown in the northern Andes of Argentina, is also sometimes cited as a grain species. However, *A. edulis* may be more appropriately considered *A. caudatus* spp. *Mantegazzianus* —a mutant of *A. caudatus* with the phenotype of club-shaped inflorescent branches and determinant growth habit (Sauer 1976). The wild putative progenitor species of the grains

include *A. powellli*, *A. hybridus* and *A. quitensis* (Sauer 1950, 1967, 1976). *A. hybridus* is also sometimes cultivated as a grain. Some of the wild relatives of the grains are fairly tall with large inflorescences; however the cultivated species are taller and more robust, with enormous inflorescences. Unlike true cereals, grain amaranths were selected for their high seed production rather than for increased seed size (Sauer 1993). Table 1 shows each grain species' proposed region of origin next to what is thought to be its most closely related progenitor.

Taxonomic problems. Because the species within the *Amaranthus* genus are very closely related, the literature shows that misclassifications among the grains as well as their weedy and wild relatives occur frequently. Genetic and environmental factors, as well as ambiguous or atypical morphology in some accessions, can lead to classification errors among these closely related species. Several studies have demonstrated the utility of molecular markers such as RAPDs and AFLPs for correcting these types of errors in grains and weeds (see Transue et al. 1994, Chan and Sun 1997, Wetzel et al. 1999, Sun et al. 1999, Xu and Sun 2001, Wassom and Tranel 2005).

Botanical Description

While grain amaranth species may be difficult to distinguish from one another on the basis of morphology, the features they share in common separate them from other amaranths. The vegetable amaranths have smooth leaves and exhibit an indeterminate growth habit. The grain amaranths are annuals and have a main stem axis with a large branched inflorescence at the apex (Stallknecht and Schulz-Schaeffer 1993). The grain species usually range from 0.4

to 3.0m in height. The grain amaranths are dicotyledonous, and, therefore, are not true cereals.

Leaves, Inflorescences and Flowers. Grain amaranth leaves are petiolate and oval to ovulate-oblong and lanceolate in shape with acute apices. The inflorescence is a dichasial cyme with unisexual flowers, which develop in a variety of colors, including red, purple, orange, or gold (Iturbide and Gispert 1994, Tapia 1994). The first flower of each of the numerous cymes is staminate followed by an indefinite number of pistillate flowers, frequently over a hundred (Pal and Khoshoo 1974, Sauer 1993). Some pistillate flowers on the cyme develop early before the staminate flower has opened, while others become receptive following the abscission of the male flower. However, because cymes at different developmental stages are present on each indeterminate inflorescence branch, self-pollination is more likely than outcrossing, although both types of fertilization are possible (Sauer 1976).

Fruit. Unlike other cereals, grain amaranths have retained the dehiscent fruits of their wild progenitors (Sauer 1993). The fruits are pyxides, meaning that they house their seeds in circumscissile capsules, which are subtended by colorful bracts and sepals (Tapia 1994). The top half of the papery utricle surrounding each seed acts as a lid-like section, which pops off at the equator of the utricle to reveal the enclosed seed. Thus, although the majority of seeds remain in the densely packed inflorescences, some seeds are lost during the harvest (Sauer 1993). However, in recent years non-shattering grain amaranth populations have been developed (Brenner 2002).

Seeds. The seeds of the grain amaranths are lens-shaped and approximately 1 to 1.5mm in diameter. The seeds come in a variety of colors, ranging from white to yellow to red to black (Iturbide and Gispert 1994, Tapia 1994, Sauer 1993). These colors are governed by simple Mendelian recessive alleles. All three grains produce both dark- and light-colored seed. Although the dark grains, which are dominant to the light-colored grains, are edible and were eaten by prehistoric hunter-gathers, the lighter grains have been selected for due to their improved flavor and popping. Furthermore, the pale color also seems to be linked to a loss of dormancy in the seed (Sauer 1976, 1993).

The seeds exhibit epigeal germination, in which the cotyledons emerge above ground as in common beans (*Phaseolus vulgaris* L.). Seedlings emerge three to four days following sowing and after about two and half months the panicle appears and flowering occurs. The seeds maintain viability for over five years at ambient temperature and <5% humidity (Iturbide and Gispert 1994).

Cultivation Practices

Traditionally, amaranth is planted either by direct sowing or by sowing in seed beds and transplanting to irrigated land (Iturbide and Gispert 1994, Tapia 1994). Amaranth is often sewn together with maize or as a border. A mixture of amaranths is usually sown in order to ensure a harvest. Seeds are sown in pre-prepared ground under dry conditions in furrows spaced 80cm apart and fed with a constant stream of water. The plants are harvested before they are fully mature in order to prevent seed fall. When the lower leaves show signs of yellowing, the plants are cut about 20cm above the soil. The sheaves are usually allowed to dry on the ground above the furrows. To remove the seeds, the sheaves are placed on sheets

on top of the ground and beaten with sticks. The chaff is removed from the grain by sifting or winnowing (Tapia 1994).

Nutrients

Amaranth seeds as a grain have been praised for their nutrient content. Amaranths are 50 to 60% starch, with higher protein (15 to 16%) and more fat (7 to 8%) than most cereals (Breene 1991). They also have nutritionally significant levels of vitamins A and C, as well as a higher mineral content than wheat (Becker et al. 1981). Amaranths also have high dietary fiber content reported to be about 8% for pale-seeded types, while the black-seeded grain types may have twice that (Pedersen et al. 1987).

Starches. In amaranth, 78 to 100% of the starch content is found in the branched-chain amylopectin form, while the remaining 0 to 22% of starch content is in the amylose or unbranched form (Tomita et al. 1981, Okuno and Sakaguki 1984). Overall, amaranth's starch composition shows a low gelatinization temperature and good stability during freezing and thawing (Yanez et al. 1986). Amaranth starch is observed in granules that are approximately 1-3 μm in diameter (Irving and Becker 1985)—much smaller than most other commercial cereals. Rice (*Oryza sativa* L.), for example, has starch granules of about 3 to 8μm, while potato's (*Solanum tuberosum* L.) are 100μm in diameter. It has thus been suggested that the small granule size might make amaranth starch useful as a food thickener, a dusting powder in foods and cosmetics, a laundry starch, etc (Yanez et al. 1986).

Proteins. The protein content of the grains has been extensively studied. Amaranth is one of a handful of plants whose protein content approaches animal protein quality on the basis of bioavailability and amino acid content (Bressani 1989). Other examples of plants with essential amino acid patterns that come close to satisfying the needs of the human diet include soybean, high-quality protein maize and quinoa (Table 2; Bressani 1989). Crude protein content from pale-seeded grain types has been reported to range from 12.5 to 22.5%, with an average of about 15% (Becker et al. 1981, Saunders and Becker 1984, Teutonics and Knorr 1985, Correa et al. 1986, Bressani et al. 1987, Pedersen et al. 1987, Bressani 1989). Furthermore, amaranth is relatively rich in the essential amino acid lysine, which is usually limiting in other cereal crops (Table 2). Lysine content ranges from 0.73 to 0.84% of amaranth's total protein content (Bressani et al. 1987). Seed storage proteins from amaranth have been introduced successfully through transgenics into other crop species. Species such as potato and maize that have been modified to express amaranth seed proteins show improved amino acid composition (Chakraborty et al. 2000, Sinagawa-Garcia et al. 2004).

Oils. The 7-8% oil content found in amaranth seeds may be too low and expensive to compete with other oils commercially available, although it is similar in content to corn and cotton seed oils (Bressani et al. 1987). Table 3 shows the ranges of fatty-acids observed for the oil content based on Breene's (1991) summary of various studies (Fernando and Bean 1984, 1985; Saunders and Becker 1984; Lorenz and Hwang 1985; Sanchez-Marroquin et al. 1986; Lyon and Becker 1987, Bressani et al. 1987). The saturated/ unsaturated fatty acid ratio has been observed to range from 0.29 to 0.43; this ratio is favorable from a nutritional standpoint because unsaturated fatty acids are predominant in amaranth oil (Breene 1991).

High levels of tocopherols (vitamin E) and tocotrienols have been reported in amaranth oil as well (Lehmann et al. 1994).

Amaranth oil has been noted for its relatively high concentration of squalene (7-8%) (Bressani et al.1987). Squalene is a lucrative ingredient used in cosmetics, skin penetrants, lubricants and is a precursor to cholesterol. The traditional source of squalene for commercial use is liver oil extracted from threatened sea animals such as whales (*Physeter macrocephalus*) and sharks (*Centrophorus squamosus*). Therefore, there is interest in other potential alternative sources. The use of amaranth oil as a squalene source may further its commercialization (Brenner et al. 2000). Recent studies have also shown that amaranth oil may be effective in reducing cholesterol levels in mammals, including humans (Berger et al. 2003, Martirosyan et al 2007).

Antinutrients. Unlike its relative quinoa, amaranth does not contain high amounts of bitter saponins that must be washed away before consumption (Tapia 1994). Low levels of saponin—around 0.1%of total seed dry weight—that have been observed for *A. cruentus* showed low toxicity in animal tests (Oleszek et al. 1999). Furthermore, amaranth grain shows low levels of some other antinutrients. For example, Lorenz and Wright (1984) studied the tannin and phytate content of *A. hypochondriacus, A. cruentus, A. hybridus* and some interspecific crosses and found that tannins were localized in seed coat and were present at 0.04-0.12%, while phytates dispersed throughout the kernel were observed at 0.5-0.6%. However, amaranth seeds and leaves are known to accumulate high levels of trypsin inhibitors as well as α-amylase inhibitors (Sanchez-Hernadez et al. 2004). These

antinutritional inhibitors are well documented and the DNA as well as protein sequences are available for some (Valdes-Rodriguez et al. 1993, 1999).

Food Uses

Traditional Uses. These nutritious crops continue to be used in Latin America much like they were during pre-Columbian times although to a much lesser extent. In Mexico, the preparation of the sacred Aztec dough, *tzoalli*, by mixing amaranth flour with maguey honey led to the current use of amaranth for preparing *alegria*, a sweet snack. However, because the use of amaranth flour was discouraged, the modern process has been altered so that *alegria* is currently made with popped amaranth seeds instead of flour (Iturbide and Gispert 1994). The popped form is also used in cereals. The seed is milled into flour to make a variety of foods, while the leaves are used as a vegetable, particularly in soups. The stems are useful as animal feed (Iturbide and Gispert 1994).

Commercial Uses. Industrial food uses of amaranth are similar to its traditional uses in Latin America (see Breene 1991 for review). Amaranth seed is packaged and sold as a whole grain or is milled into whole, high-bran and low-bran flour. Amaranth grain can be difficult to mill, however, due the grain's unusual morphological characteristics and small seed size. Many studies have focused on the beneficial use of amaranth as a replacement in wheat and corn flour. Malted flours are also produced from amaranth seeds that are allowed to germinate to produce "malt," which is then dried and ground into flour. It has been noted that processing amaranth in this manner resulted in an increase of 25 to 30% in true protein content likely due to the decrease in total fat and carbohydrates during the malting process. Another common cereal preparation method applied to amaranth is extrusion. Extrusion

involves exposing the food product to intense pressure and heat within an apparatus known as an extruder and cooking it in such a way that the product acquires a particular desired shape or greater uniformity. Amaranth prepared in this manner has been used as an ingredient in beverages, baby formula, atole, croutons, snacks, breakfast cereals, and as a textured vegetable protein (Breene 1991).

Adaptations

The grain amaranths exhibit C_4 photosynthesis. Thus, they grow rapidly in bright sunlight, high temperatures, and low moisture conditions. Other cultivated crops that exhibit C_4 photosynthesis include maize, sorghum (*Sorghum* spp. L.) and sugarcane (*Saccharum officinarum* L.). Amaranth is better adapted to semiarid environments than these plants, however, because it can make osmotic adjustments that allow it to tolerate dry conditions without wilting or drying (Tucker 1986). Amaranths can also tolerate a variety of unfavorable soil conditions such as high salinity, acidity, or alkalinity (Tucker 1986). Grain amaranths have also been reported to adapt readily to new environments, including some that are inhospitable to traditional grain crops (Gupta and Gudu 1991).

Pathology

Damping Off. Common and potential pathogens and insect problems of amaranths have been reviewed (see Weber et al. 1990, Wilson 1990). A frequent problem in young amaranths known as "damping off" is caused by soil-dwelling fungi such as *Pythium* spp. and *Rhizoctonia* spp. These fungi infect roots and developing stems of amaranths. The soft, water-soaked stems of infected plants cannot support the seedlings, which subsequently

lodge and die. Plants are vulnerable to damping off from the germination stage until they reach 4 to 6 inches in height, at which stage their stem thickens and damping off ceases to be a serious problem. Some amaranth species are resistant to damping off (Sealy et al. 1988). Cold, wet soils; early planting; high populations; and excessive nitrogen in nitrate forms contribute to damping off in grain amaranths. Damping off may be avoided if the crop is planted in warm, dry soils without excessive nitrates (Wilson 1990).

Crown and Root Rot. Various insects tunnel into the stems of grain amaranths. This tunneling by insects allows fungi and bacteria to enter a plant's stem and cause rotting, which may lead to crown and root rot. Symptoms—soft, mushy crown tissue; browning; or lodging—are not usually visible until the flowering stage. In the presence of insects, the fungi *Pythium*, *Rhizoctonia* and *Fusarium* are the most common culprits of rot diseases (Wilson 1990). Fusarium wilt caused by the fungus *Fusarium oxysporum* is considered a serious threat to the production of *A. hybridus* as a crop in South Africa (Chen and Swart, 2002). In the absence of insects, *Rhizoctonia solani* has been observed to cause basal stem canker and girdling (Wilson 1990).

Leaf diseases. Leaf diseases in grain amaranths are not as prevalent in the United States as they are in warmer, wetter climates such as India and South Africa (Wilson 1990). One disease known to infect Indian varieties of *A. hybridus* is alternaria leaf spot caused by *Alternaria* spp. (Mondal et al. 2002). White rust caused by *Albugo bliti*, which is a common plague of Indian vegetable amaranths, was observed for the first time in the United States infecting seabeach amaranth (*Amaranthus pumilus* Raf.) (Keinath et al. 2003). Thus, it could

be a potential disease in grain crops in the US as well. Varalakshmi and Celiachalam (2002) found that some types of vegetable amaranths show resistance to *Albugo bliti*.

Insect Pests

In addition to pathogens, a variety of insects have been observed feeding on amaranth, including Lygus bug (*Lygus lineolaris* Palisot de Beauvois), cowpea aphid (*Aphis craccavora* Koch), fall armyworm (*Spodoptera frugiperda* J.E. Smith), cabbage looper (*Trichoplusia ni* Hubner), corn earworm (*Heliothis zea* Boddie), striped blister beetle (*Epicauta vittat* Fab.), weevil (*Conotrachelus seniculus* LeConte), and spinach flee beetle (*Disonycha xanthomelas* Dalman) (Wilson 1989).

Lygus Bug. The most significant of these insect pests for amaranth is the Lygus bug, specifically the tarnished plant bug. Lygus bug is a common pest throughout the world. In the US alone it feeds on at least 328 known hosts (Young 1986). It feeds on developing amaranth flower tissues, which leads to a variety of problems including deformation or abscission of the fruit, necrosis and localized wilting (Gupta et al. 1980, Khattat and Stewart 1975, Strong 1968, Tingey and Pillemer 1977). These problems can lead to significantly decreased seed yield if the pest is present in high enough quantities (Wilson and Olsen 1990).

Amaranth Curculio. Another major concern for US populations of amaranth is the amaranth weevil, because it infests wild populations of amaranth in most areas of the US. The adults oviposit at the base of the stem, causing stem breakage. The tunneling and feeding of the larvae weakens the plant's root system and causes decreased nutrient uptake while increasing

the risk of lodging. Holes in the plant left by tunneling provide entryways for pathogens, especially root rots (Weber et al. 1990).

Evolution

Currently, there are two competing hypotheses for the origin of the grain amaranths. Both hypotheses were proposed by Sauer (1950, 1967, 1976) with one based on geography and the other based on morphological features. The first hypothesis suggests that all three grain amaranths evolved independently—*A. caudatus* from *A. quitensis* in the Andean region of South America, *A. cruentus* from *A. hybridus* in Central America, and *A. hypochondriacus* from *A. powellii* in Mexico. The second is that all three grains are descended mainly from *A. hybridus*. More specifically, *A. hybridus* gave rise to *A. cruentus* and introgression to this grain from *A. powellii* and *A. quitensis* produced *A. hypochondriacus* and *A. caudatus* respectively (see Table 1). In the latter scenario, however, the origin of *A. caudatus* is somewhat puzzling, because a large region exists between the Andes and northern Central America where *A. cruentus* is not cultivated as a grain. If the second hypothesis is correct, perhaps *A. cruentus* moved through this region, though it did not remain there as a cultivated crop, to the Andean region in order to combine with *A. quitensis* to give rise to *A. caudatus* (Sauer 1993).

Hybrid Fertility and Chromsome Number Studies

In addition to Sauer's morphological studies, other studies on hybrid fertility, chromosome number and molecular markers have attempted to further the understanding of the evolutionary relationships among the grain amaranths. Most of the three grains and their

putative progenitors are thought to be paleo-allotetraploids (2n=2x=32). However, chromosome counts of both 32 and 34 are reported in the literature for *A. cruentus* and *A. powellii*. Brenner et al. (2000) summarized data from several cytological studies conducted by Pal et al. (1982) and Greizerstein and Poggio (1992, 1994, 1995) (See Table 4).

Based on chromosome number Pal and Khoshoo (1972) hypothesized that *A. powellii*, rather than *A. hybridus*, is the more likely to be the most closely related putative progenitor *A. cruentus*. However, this assumption has been refuted by molecular studies (Sun et al. 1999, Xu and Sun 2001). Based on crosses among amaranth species, multiple studies have described *A. hybridus* as being capable of forming hybrids with all of the other grain types and putative progenitors (Pal and Khoshoo 1974), which supports Sauer's hypothesis of one common ancestor for all of the grains. Pal and Khoshoo (1972) also suggested that *A. caudatus* and *A. hypochondriacus* were strongly differentiated genetically and that significant genetic exchange between them was unlikely based on the deformity and high sterility rate they observed among hybrids. However, crosses among the three grain types conducted by Gupta and Gudu (1991) suggested that *A. caudatus* and *A. hypochondriacus* were more closely related than previously suggested by Pal and Khoshoo (1972), because F_1 hybrids between the two species were easily obtained and most were phenotypically normal, although some plants exhibited low pollen fertility. Gupta and Gudu (1991) suggested that environmental factors were likely responsible for the discrepancies between the two studies. In the Gupta and Gudu study (1991), crosses between *A. caudatus* and *A. cruentus* were more difficult, although hybrid seedlings were obtained. None survived to maturity, dying shortly after germination, however. Similar results were observed for crosses between A. *cruentus* and *A. hypochondriacus*. Thus, Gupta and Gudu

(1991) concluded that *A. caudatus* and *A. hypochondriacus* were the most closely related grain species on the basis of hybrid development.

Molecular Marker Studies

Studies using molecular markers, including analyses with isozymes (Chan and Sun 1997), RAPDs (Transue et al. 1994, Mandal and Das 2002, Chan and Sun 1997), AFLPs (Xu and Sun 2001), and low-Cot DNA probes generated from highly and moderately repetitive DNA (Sun et al. 1999) have attempted to further clarify the question of the origin of the grain amaranths.

Isozymes and RAPDs. Transue et al. (1994) analyzed 282 polymorphic RAPD markers in order to classify 70 amaranths whose species had previously been undetermined. Transue et al. (1994) phylogenetic analysis using these markers indicated that all three grains can be unambiguously classified molecularly although some overlap in their morphology does exist. Their data also supported Sauer's hypothesis of a single common ancestor for all the grains. Chan and Sun (1997) revealed similar relationships for the amaranth grains based on RAPD and isozyme data. They also observed that *A. cruentus* had the lowest level of variation based on isozyme and RAPD data. The most recent study using RAPDs was conducted by Mandal and Das (2002). Curiously, these three studies suggested that *A. hypochondriacus* and *A. caudatus* are more similar to each other than either is to *A. cruentus* although they are geographically separated by *A. cruentus*. (Transue et al. 1994, Chan and Sun 1997, Mandal and Das 2002).

Restriction-site variation. Lanoue et al. (1996) analyzed 28 *Amaranthus* species for restriction-site variation within nuclear ITS1 and ITS2 regions as well as in chloroplast DNA. A low level of variation based on restriction-site analysis was observed and resulted in poorly resolved trees. However, Lanoue and colleagues (1996) did observe a close relationship between *A. cruentus* and its putative progenitor *A. hybridus* as well as a close relationship between *A. caudatus* and *A. quitensis*. Contradictory to the RAPD studies, they also found that among the grains *A. cruentus* is more closely related to *A. caudatus* than either is to *A. hypochondriacus*.

Low-Cot DNA fingerprinting. Sun et al. (1999) developed probes from highly repetitive sequence. The method used to develop the probes relied on the faster annealing rate of repetitive sequence over unique sequences. The probes were developed from *A. tricolor* genomic DNA and consisted of microsatellites, minisatellites, ribosomal RNA genes (rDNA), interspersed retrotransposons or retrotransposons-like sequences, and other unidentified "junk" DNA. They were termed "low-Cot" probes because their Cot value, which is based on initial concentration of ssDNA (Co) and annealing temperature (t), was low for these repetitive sequences.

Sun and colleagues (1999) found that more highly conserved sequence such as rDNA was most useful for resolving interspecific relationships among more distantly related species, whereas probes with microsatellite sequence were best for resolving intraspecific relationships as previously noted. Phylogenetic analysis based on low-Cot data showed that the grain amaranths clustered close to *A. hybridus*, with *A. cruentus* being particularly closely related to *A. hybridus*. The next closest putative progenitor species was *A. quitensis*, while *A.*

powellii was clearly the most distant species included in the analysis. It was also observed that low-Cot DNA probes did not cluster *A. caudatus* accessions well. For example, one *A. caudatus* accession in the anlaysis clustered with A. quitensis, while another accession grouped with *A. hypochondriacus*. However, the overall analysis still yielded relationships consistent with Sauer's hypothesis of the evolution of *A. caudatus* through introgression to *A. cruentus* from *A. quitensis*.

Sun et al. (1999) also noted that a probe consisting of microsatellite and minisatellite sequence showed the highest polymorphism—75% on average across six species of grain amaranths and their close relatives—among probes designed from the various types of repetitive sequences in the study. Sun and colleagues (1999), therefore, suggest that microsatellites are the most suitable type of marker for characterizing intraspecific amaranth accessions. Thus, SSR markers may be even more helpful in taxonomic classification than other markers such as AFLPs and RAPDs.

ITS Sequence, AFLPs, and ISSRs. Xu and Sun (2001) studied variation in a panel of 30 *Amaranthus* accessions based on ITS Sequence, AFLPs, and ISSRs (double primer intersimple sequence repeats). The panel included four accessions from each of the grain and putative progenitor species—also known as the *A. hybridus* species complex—and 3 accessions from *A. tricolor*. Although they are more cost effective, Xu and Sun (2001) found that ISSRs alone could not separate *A. caudatus* accessions from *A. quitensis* accessions, and thus concluded that AFLPs were more effective for phylogenetic analysis than ISSRs. Consistent with Lanoue et al. (1996), Xu and Sun (2001) found little sequence diversity based on ITS region within the *A. hybridus* species complex. Only A. *powellii* was

sufficiently divergent at the ITS to be resolved from the other species in the *A. hybridus* complex.

Phylogenetic analysis based on combined data from AFLPs and ISSRs supported a close relationship between *A. caudatus* and A. quitensis. Xu and Sun (2001) generated multiple trees using multiple methods for construction: UPGMA, parsimony, and neighbor-joining. They concluded that the neighbor-joining method yielded the best trees based on their dataset, as it was the method with the most consistent topology between the AFLP and ISSR datasets and most coherent with the current morphology-based intra- and interspecifc classifications.

Although the studies based on RAPD data showed that *A. hypochondriacus* was more closely related to *A. caudatus* than to *A. cruentus*, Xu and Sun's (2001) data showed conflicting relationships depending on which method of tree construction was used. Xu and Sun's (2001) neighbor-joining analysis (based on combined ISSR and AFLP data) indicated that *A. caudatus* and *A. hypochondriacus* were the two most closely related species. However, their strict consensus tree of 21 equally most parsimonious trees using the same data set showed approximately equal branching among the three grain types. As previously noted, Sun et al.'s (1999) analysis supported this same relationship.

In summary, the commonalities among the conclusions of these phylogeny studies are i) that the grain amaranths are monophyletic in origin, ii) *A. hybridus* is the most closely related ancestor to the grain amaranths, iii) *A. powellii* is the most divergent of the putative progenitors, and iv) new methods are needed to resolve other relationships that remain ambiguous within the complex. The discrepancies identified among the studies are likely attributable to the dissimilar marker systems used, differences in accessions and sample sizes

within each study's panel, as well as to the methodology used for constructing phylogenetic trees. A major problem in particular with previous studies is that the sampling has been almost exclusively from the grain species, while very few accessions from the weedy putative progenitors were included. These studies also highlight the need for a more robust set of molecular markers. The use of microsatellite markers will likely meet this need.

MICROSATELLITES

Microsatellite markers are short tandem repeats of nucleotides that are usually one to four base pairs in length, although motifs as long as seven or eight bp in length may be classified as microsatellites. Microsatellites are thus distinguished from minisatellite (10 to 30bp long motifs) and satellite (>30bp long motifs) DNA. Unlike their larger counterparts that tend to be found mainly in telomeric regions of chromosomes, microsatellites are relatively evenly dispersed throughout eukaryotic genomes (Ellegren 2004), making them useful for obtaining a glimpse of an organism's entire genome when complete sequencing is not an option. Microsatellites are also known as Short Tandem Repeats (STRs) or Simple Sequence Repeats (SSRs). Microsatellites are thought to be generated by the slippage of one DNA strand during replication of double-stranded DNA or by unequal crossing over during meiosis (Ellegren 2004).

Applications in Plant Genetics Research

Microsatellite markers are useful in diversity studies, in identifying specific cultivars, and in studies involving comparisons among species. The most common microsatellite motif among plant genomes is AT, in contrast to animal genomes in which AC/TG repeats are the

most frequently observed. The most common trinucleotide motif in plants is ATT (Morgante and Olivieri 1993). Most repeats are found in noncoding sequences. If they occur in coding sequences, they are usually found in untranslated regions.

The benefits of microsatellite marker systems have already been demonstrated in a number of crops. For example, microsatellites developed from peaches (*Prunus persica*) were used to identify similarities between peaches and sweet cherry (*P. avium*) genotypes (Wunsch and Hormaza 2002). In addition to their value in comparing multiple species, microsatellites are also useful for evaluating diversity within a single crop species as shown by Ni et al. (2002) who used SSRs to show differences between two very closely related rice (*Oryza sativa* L.) subspecies: *indica* and *japonica*. Microsatellites are also used to understand individual traits within a species. For example, researchers used SSRs to study genes conferring resistance to Septoria tritici blotch in wheat (*Triticum aestivum* L.) (Adhikari et al. 2004). SSRs have also been used to help researchers understand the evolutionary development of organisms. Olsen and Schaal (2001) used SSRs to study the development of Cassava (*Manihot esculenta* subsp. *esculenta*).

Advantages in Plant Breeding

The use of microsatellites, as well as other molecular markers, in plant breeding has advantages over conventional breeding methods in that it can significantly decrease the time required to breed an improved cultivar (Yousef and Juvik 2001). Microsatellite markers, for example, can be used at any stage of a plant's development. They do not vary with the environment, and can be used to detect heterozygosity that may not be apparent in the plant's

phenotype. They are also useful in finding genes that contribute to polygenic traits (Prasad et al. 2003).

Microsatellites as a Marker System

Microsatellites are a model genetic marker system because they are reproducible and relatively easy to use in comparison to other types of genetic markers. For this reason, "microsatellites quickly became the marker of choice in genome mapping, and subsequently also in population genetics studies and related areas" (Ellegren 2004) after 1989 when they were first developed. Microsatellites are initially expensive to develop due to sequencing costs. Once developed, however, microsatellite markers are inexpensive and relatively easy for breeders to use, because they require only a small amount of DNA and relatively little technique expertise. Furthermore, they are codominant, locus specific, and highly polymorphic (Morgante and Olivieri 1993). Microsatellite markers are especially useful for identifying multiple alleles (He et al. 2002).

Microsatellites in Related Species

The closest relatives for which microsatellite markers have been developed are the group of organisms formerly known as the Chenopods, most notably quinoa and sugar beet. Over 400 SSR markers have been developed for quinoa, while a much smaller number have been developed in *B. vulgaris* (Rae et al. 2000, Cureton et al. 2002, Mason et al. 2005, Jarvis 2006). Data from these studies are summarized in Table 5.

Microsatellite markers developed for these related crop species have already proven useful in mapping and trait analysis studies (Rae et al. 2000, Maughan et al. 2004, Ricks

2005, Jarvis 2006). The development of SSRs for the grain amaranths in this study was accomplished according to the methodology used for SSR development in its relative quinoa (Mason et al. 2005, Jarvis 2006).

CONCLUSIONS

This project constitutes the first development of SSRs among the *Amaranthus* species. Because microsatellites are inexpensive to use and require little technical expertise once developed, they will be useful for breeders in the areas to which the grains are indigenous. Furthermore, analysis with these markers will be a key stepping stone in understanding more about amaranths, particularly the grain amaranths. They will fuel additional research projects such as genome mapping, trait analysis, etc. The markers will also be helpful in evaluating the origin of the grains, which still remains ambiguous.

REFERENCES

Adhikari T, Wallwork H, Goodwin S (2004) Microsatellite markers linked to the *Stb2* and *Stb3* genes for resistance to Septoria tritici blotch in wheat. Crop Sci 44:1403–1411

Becker R (1989)

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

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27:573–80

Berger A, Gremaud G, Baumgertner M, Rein D, Monnard I, Kratky E, Geiger W, Burri J, Dionisi F, Allan M, Lambelet P (2003) Cholesterol-lowering properties of amaranth grain and oil in hamsters. Int J Vitam Nutr Res 73:39–47

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Gene 32:314–31

Breene WM (1991) Food uses of grain amaranth. Cereal Foods World 36:426–430

Brenner DM, Baltensperger DD, Kulakow PA, Lehmann JW, Myers RL, Slabbert MM, Sleugh BB (2000) Genetic resources and breeding of *Amaranthus*. Plant Breeding Reviews 19:227–285

Brenner DM (2002) Non-shattering grain amaranth populations. In: Janick J, Whipkey A (eds) Trends in new crops and new uses. ASHS Press, Alexandria, VA, pp 104–106

Bressani R (1989) The proteins of grain amaranth. Food Rev Int 5:13–38

Bressani R, Gonzalez JM, Zuniga J, Bruener M, Elias LG (1987) Yield, selected chemical composition and nutritive value of 14 selections of amaranth grain representing four species. J Sci Food Agric. 38:347–356

Bressani R, Sanchez-Marroquin A, Morales E (1992) Chemical composition of grain amaranth cultivars and effects of processing on their nutritional quality. Food Rev Int 8:23–49

Chen WQ, Swart WJ (2002) The in vitro phytotoxicity of culture filtrates of *Fusarium oxysporum* to five genotypes of *Amaranthus hybridus*. Euphytica 127:61–67

Chakraborty S, Chakraborty N, Datta S (2000) Increased nutritive value of transgenic potato by expressing a nonallergenic seed albumin gene from *Amaranthus hypochondriacus*. PNAS 97:3724–3729

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

Correa AD, Jokl L, Carlsson R (1986) Amino-acid composition of some *Amaranthus* spp. grain proteins and of its fractions. Arch Latinoam Nutr 36:466–476

Cureton AN, Burns MJ, Ford-Lloyd BV, Newbury HJ (2002) Development of simple sequence repeat (SSR) markers for the assessment of gene flow between sea beet (*Beta vulgaris* ssp. *maritima*) populations. Mol Ecol Notes 2:402–403

Dowton WJS (1972) *Amaranthus edulis*: a high lysine grain amaranth. World Crops 25:20

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5:435–445

Fernando T, Bean G (1984) Fatty acids and sterols of *Amaranthus tricolor* L. Food Chem 15:233–237

Fernando T, Bean G (1985) A comparison of the fatty acids and sterols of weedy and vegetable species of *Amaranthus* spp. J Am Oil Chem Soc 62:89–91

Greizerstein EJ, Poggio L (1992) Estudios citogenetics de seis hibridos interespicificos de *Amaranthus* (Amaranthaceae). Darwiniana 31:159–165

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

Greizerstein EJ, Poggio L (1995)Meiotic studies of spontaneous hybrids of *Amaranthus*: Genome analysis. Plant Breed 114:448–450

Gupta VK, Gudu S (1991) Interspecific hybrids and possible phylogenetic relations in grain amaranths. Euphytica 52:33–38

Gupta RK, Tamaki G, Johanson CA (1980) Lygus bug damage, predator-prey interaction and pest management implications on alfalfa grown for seed. Wash State Univ Coll Agric Res Cent Tech Bull, No. 92

He C, Poysa V, Yu K (2002) Development and characterization of simple sequence repeat (SSR) markers in *Lycopersicon esculentum* cultivars. Theor Appl Genet 106: 363–373

Irving DW, Becker R (1985) Seed structure and composition of potential new crops. Food Microstruct 4:43–54

Iturbide GA, Gispert M (1994) Grain amaranths (*Amaranthus* spp.). In: Hernandez-Bermejo JE, Leon J (eds) Neglected Crops: 1492 from a different perspective. FAO, Rome, pp 93–101

Jarvis (2006) Simple sequence repeat development, polymorphism, and genetic mapping in quinoa (*Chenopodium quinoa* Willd). Thesis. Brigham Young University, Provo, UT

Kauffman CS, Reider C (1986) Rodale germplasm collection. RRC/NC86/2. Rodale Press, Emmaus, PA.

Keinath AP, Strand AE, Hamilton RD (2003) First report of white rust caused by *Albugo bliti* on seabeach amaranth in the United States. Plant Dis 87:602

Khattat AR, Stewart RK (1975) Damage by tarnished plant bug to flowers and setting pods of greenbeans. J Econ Entomol 68:633–635

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

Lehmann JW (1996) Case history of grain amaranth as an alternative crop. Cereal Foods World 41:399–409

Lehmann JW, Putnam DH, Qureshi AA (1994) Vitamin E isomers in grain amaranths (*Amaranthus* spp.). Lipids 29:177–181

Lorenz K, Wright B (1984) Phytate and tannin content of amaranth. Food Chem 14:27–31

Lorenz K, Hwang YS (1985) Lipids in amaranths. Nutr Rep Int 31:83–89

Lyon CK, Becker R (1987) Extraction and refining of oil from amaranth seed. J Am Oil Chem Soc 64:233–236

Mandal N, Das PK (2002) Intra- and interspecific genetic diversity in grain *Amaranthus* using random amplified polymorphic DNA markers. *Plant Tissue Culture* 12:49–56

Martirosyan DM, Miroshnichenko LA, Kulakova SN, Pogojeva AV, Zoloedov VI (2007) Amaranth oil application for coronary heart disease and hypertensions. Lipids Health Dis 6:1

Mason SL, Stevens MR, Jellen EN, Bonifacio A, Fairbanks DJ, Coleman CE, McCarty R, Rassmussen A, Maughan PJ (2005) Development and use of microsatellite markers for germplasm characterization in *Chenopodium quinoa* Willd. Crop Sci 45: 1618–1630

Maughan PJ, Bonifacio A, Jellen EN, Stevens MR, Coleman CE, Ricks M, Mason SL, Jarvis DE, Bardunia BW, Fairbanks DJ (2004) A genetic linkage map of quinoa (*Chenopodium quinoa*) based on AFLP, RAPD, and SSR markers. Theor Appl Genet (2004) 109:1188–1195

Mondal KK, Dihman KC, Argawal DK (2002) Alternaria leaf spot of grain amaranth. Indian Phytopathology 55:248

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

National Academy of Sciences (1984) Amaranth: Modern prospects for an ancient crop. Natl Acad Sci, Washington DC.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genet 89:583–590

Ni J, Colowit PM, Mackill DJ (2002) Evaluation of Genetic Diversity in Rice Subspecies Using Microsatellite Markers. Crop Sci 42:601–607

Okuno K, Sakaguki S (1984) Differentiation of starch property in perisperm of amaranths. JARQ 18:1–5

Oleszek W, Junkuszew M, Stochmal A (1999) Determination and toxicity of saponins from *Amaranthus cruentus* seeds. J Agric Food Chem 47:3685–3687

Olsen KM, Schaal BA (2001) Microsatellite variation in cassava (*Manihot esculenta*, Euphorbiaceae), and its wild relatives: further evidence for a Southern Amazonian origin of domestication. Am J Bot 88:131–142

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

Pal M, Khoshoo TN (1974) Grain amaranths. In: Hutchinson JB (ed) Evolutionary studies in world crops: diversity and change in the Indian subcontinent. Cambridge University Press, UK, pp 129–137

Pal M, Pandey RM, Khoshoo TN (1982) Evolution and improvement of cultivated amaranths VIII. Induced autotetraploidy in grain types. Z. Pflanzenzuch. 78:135–148.

Patzoldt WL, Hager AG, McCormick JS, Tranel PJ (2006) A codon deletion confers resitance to herbicides inhibiting protoporphyrinogen oxidase. PNAS 103:12329-12334

Pedersen B, Kalinowski LS, Eggum BO (1987) The nutritive value of amaranth grain (*Amaranthus caudatus*). 1. Protein and minerals of raw and processed grain. Qualitas Plantarum 36:309–324

Prasad M, Kumar N, Kulwal PL, Roder MS, Balyan HS, Dhaliwal HS, Gupta PK (2003) QTL analysis for grain protein content using SSR markers and validation studies using NILs in bread wheat. Theor Appl Genet 106:659–667

Rae SJ, Aldam C, Dominguez I, Hoebrechts M, Barnes SR, Edwards KJ (2000) Development and incorporation of microsatellite markers into linkage map of sugar beet (*Beta vulgaris* spp.). Theor Appl Genet 100:1240–1248

Ricks MD (2005) Genetic mapping of the bitter saponin production locus (BSP locus) in *Chenopodium quinoa* Willd. Thesis. Brigham Young University, Provo, UT

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S and Misener S, eds. Bioinformatics methods and protocols: Methods in molecular biology. Human Press, Totowa, NJ, pp 365–386

Sanchez-Hernandez C, Martinez-Gallardo N, Guerrero-Rangel A, Valdes-Rodriguez S, Delano-Frier J (2004) Trypsin and α-amylase inhibitors are differentially induced in leaves of amaranth (*Amaranthus hypochondriacus*) in response to biotic and abiotic stress. Physiol Plant 122:254–264

Sanchez-Marroquin A, Del Valle FR, Escobedo M, Avitia R, Maya S, Vega M (1986) Evaluation of whole amaranth (*Amaranthus cruentus*) flour, its air-classified fractions, and blends of these with wheat and oats as possible components for infant formulas. J Food Sci 51:1231–1238

Sambrook J, Fritsch EE, Maniatis T (1989) Molecular cloning. A laboratory manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cld Spring Harbor, NY

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

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

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

Sauer JD (1993) Amaranthaceae—amaranth family. In: Historical Geography of Crop Plants: A Select Roster. CRC, Boca Raton, Florida, USA pp 9–14

Saunders RM, Becker R (1984) *Amaranthus*: A potential food and feed resource. In Pomeranz Y (ed) Advances in cereal science and technology, Vol 6. American Association of Cereal Chemists, St. Paul, Minn.

Sealy RL, Kenerley CM, McWilliams EL (1988) Evaluation of *Amaranthus* accessions for resistance to damping off by *Pythium myriotylum*. Plant Disease 72:985–989

Sinagawa-Garcia SR, Rascon-Cruz Q, Valdez-Ortiz A, Medina-Godoy S, Escobar-Gutierrez A, Paredes-Lopez O (2004) Safety assessment by in vitro digestibility and allergenicity of genetically modified maize with an amaranth 11S globulin. J Agric Food Chem 52:2709–2714

Stallknecht GF, Schulz-Schaeffer JR (1993) Amaranth rediscovered. In: Janick J, Simon JE (eds), New crops. Wiley, New York, pp 211–218

Strong FE (1968) The selective advantage occurring to lygus bugs that cause blating of floral parts. J Econ Entomol 61:315–316

Sun M, Chen H, Leung FC (1999) Low-Cot DNA sequences for fingerprinting analysis of germplasm diversity and relationships in *Amaranthus*. Theor Appl Genet 99:464–472

Tapia M (1994) Neglected crops of the Andean region. In: Hernandez-Bermejo JE, Leon J (eds) Neglected Crops: 1492 from a different perspective. FAO, Rome, pp 144–148

Teutonics RA, Knorr D (1985) Amaranth: composition, properties, and applications of a rediscovered food crop. Food Tech 39:49–61

Tingey WM, Pillemer EA (1977) Lygus bugs: crop resistance and physiological nature of feeding injury. J Econ Entomol 61:315–316

Todd JJ, Vodkin LO (1996) Duplications that suppress and deletions that restore expression from a chalcone synthase multigene family. Plant Cell 8:687–699

Tomita Y, Sugimoto Y, Sakamoto S, Fuwa H (1981) Some properties of starches of grain amaranths and several millets. J Nutr Sci Vitaminol 27:471–484

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

Tucker JB (1986) Amaranth: the once and future crop. BioScience 16:9–13

Valdes-Rodriguez S, Segura-Nieto M, Chagolla-Lopez A, Vargas-Cortina AVy, Martinez-Gallarado N, Blanco-Labra A (1993) Purification, characterization, and complete amino acid sequence of a trypsin inhibitor form amaranth (*Amaranthus hypochondriacus*) seeds. Plant Phys 103:1407–1412

Valdes-Rodriguez S, Blanco-Labra A, Gutierrez-Benicio G, Boradenenko A, Herrera-Estrella A, Simpson J (1999) Cloning and characterization of a trypsin inhibitor cDNA from amaranth (*Amaranthus hypochondriacus*) seeds. Plant Mol Biol 41:15–23

Varalakshmi B, Celiachalam (2002) Screening of vegetable amaranth germplasm against white rust. Plant Dis Res 17:125–126

Wassom JJ, Tranel PJ (2005) Amplified fragment length polymorphism-based genetic relationship among weedy *Amaranthus* species J Hered 96:410–416

Weber JL (1990) Informative value of human (dC-dA)n · (dG-dT)n polymorphisms. Genomics 7: 524–530

Weber LE, Applegate WW, Baltensperger DD, Irwin MD, Lehmann JW, Putnam DH (1990) Amaranth-grain production guide. Rodale Press, Emmaus, PA.

Wetzel DK, Horak MJ, Skinner DZ (1999) Use of PCR-based molecular markers to identify weedy *Amaranthus* species. Weed Sci 47:518–523

Wilson RL (1989) Studies of insects feeding on grain amaranths in the Midwest. J Kan Entomol Soc 62:440–480

Wilson RL, (1990) Insect and disease pests of amaranth. In: Proc. 4th Nat. Amaranth Symposium: Perspectives on Production, Processing and Marketing. University Minnesota Press, St Paul, MN, pp 163–169

Wilson RL, Olson DL (1990) Tarnished plant bug, *Lygus lineolaris* (Palisot de Beauvois) (Hemiptera:Miridae) oviposition site preference on three growth stages of a grain amaranth, *Amaranthus cruentus* L. J Kan Entomol Soc 63:88–91

Wunsch A, Hormaza JI (2002) Molecular characterisation of sweet cherry (*Prunus avium* L.) genotypes using peach (*Prunus persica* L. Batsch) SSR sequences. Heredity 89:56–63

Xu F, Sun M (2001) Comparative analysis of phylogenetic relationships of grain amaranths and their wild relatives (*Amaranthus*; Amaranthaceae) using internal transcribed spacer, amplified fragment length polymorphism, and double-primer fluorescent intersimple sequence repeat markers. Mol Phylogenet Evol 2:372–387

Yanez GA, Messinger JK,Walker CE, Rupnow JH (1986) *Amaranthus hypochondriacus*: starch isolation and partial characterization. Cereal Chem 63:273–276

Young OP (1986) Host plants of the tarnished plant bug *Lygus lineolaris* (HeteropteraL Miridae). Ann Entomol Soc Am 79:747–762

Yousef GG and Juvik JA (2001) Comparison of phenotypic and marker-assisted selection for quantitative traits in sweet corn. Crop Sci 41:645–655

CHAPTER 2: TABLES

Table 1. Grain amaranths and putative progenitor species (summarized from Sauer 1950, 1967, 1976)

Grain Species	Grain Region	Putative Progenitor	Progenitor Description	Progenitor Region
A. hypochondriacus	N.W. and central Mexico	*A. powellii*	pioneer of open habits (desert washes, canyons, etc.)	western Cordillera
A. cruentus	S. Mexico and C. America	*A hybridus*	riverbank pioneer (moister regions)	E. North America, Central America highlands
A. caudatus	Andean region	*A. quitensis*	riverbank pioneer	subtropical South America

Table 2. Comparison of amino acid balance of amaranth with other cereals (reproduced from Becker 1989).

Amino Acid	FAO-WHO standard	Amaranth	Wheat	Brown Rice	Soybean
		--- % of protein ---			
Ile	4.0	3.6	3.3	3.8	4.5
Leu	7.0	5.3	6.7	8.2	7.8
Lys	5.4	5.1	2.9	3.8	6.4
Met/Cys	3.5	4.4	4.0	3.6	2.6
Phe/Tyr	6.1	6.6	7.5	8.6	8.1
Thr	4.0	3.4	2.9	3.9	3.9
Trp	1.0	0.9	1.1	1.2	1.3
Val	5.0	4.2	4.4	5.5	4.8
Limiting Amino Acid	---	Leu	Lys	Lys	Met/Cys
Score	100	75	53	70	74

Table 3. Observed ranges for fatty acid components of oil content summarized in Breene (1991)	
Fatty Acid	**Observed Range (%)**
linoleic	37-62
oleic	19-35
palmitic	12-25
stearic	2-5
linolenic	0.3-2

Table 4. Genome formulas for grain amaranths and some related wild species (reproduced from Brenner et al. 2000)		
Amaranthus species	Genome formulas*	n
caudatus	$A_1A_1B_1B_1$	16
cruentus	$A_2A_2B_2B_2$	17
hypochondriacus	$A_4A_4B_4B_4$	16
mantegazzianus	$A_3A_3B_3B_3$	16
quitensis	AABB	16
hybridus	$A_5A_5B_5B_5$	16
spinosus	A_6A_6CC	17
*Minor differences are expressed in the subscripts for the A and B genomes. The genomes are x = 8 except for B_2 and C, which are x = 9 (Greizerstein and Poggio 1995).		

Table 5. Summary of SSR development studies in the Amaranthaceae family				
	Quinoa--Mason et al. (2005)	**Quinoa--Jarvis (2006)**	**Sea beet--Cureton et al.(2002)**	**Sugar beet--Rae et al. (2000)**
Total individuals in screening panel	34	23	6	12
Total Clones	1472	1276	256	1536
Total SSR primers (% of clones)	397 (31%)	402 (35%)	30 (12%)	114 (7.4%)*
Polymorphic SSRs	208 (51%)	216 (50%)	6 (20%)	57 (50%)
ONA range	2 to 13	2 to 13	3 to 5	3 to 9
average ONA	4	4	3.83	—
H-value range	0.20 to 0.90	0.12 to 0.90	—	0.10 to 0.81
average H	0.57	0.56	—	0.61
*Unlike the other three libraries in this table, this library was screened for SSRs using a probe-hybridization technique instead of sequencing all clones				