2011-06-24

Assessment of Genetic Diversity Among Peruvian Amaranth (*Amaranthus caudatus* L.) Germplasm Using SNP Markers

Felix Ruben Jimenez Rondan
Brigham Young University - Provo

Assessment of Genetic Diversity Among Peruvian Amaranth (*Amaranthus caudatus* L.) Germplasm Using SNP Markers

Felix R. Jimenez

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Eric N. Jellen, Chair
Peter J. Maughan
Joshua A. Udall

Department of Plant and Wildlife Sciences

Brigham Young University

August 2011

ABSTRACT

Assessment of Genetic Diversity Among Peruvian
Amaranth (*Amaranthus caudatus* L.)
Germplasm Using SNP Markers

Felix R. Jimenez
Department of Plant and Wildlife Sciences, BYU
Master of Science

Abstract Amaranth (*Amaranthus caudatus* L.) is an important pseudocereal in the Andes. The seed has excellent nutritional value (high in protein, essential amino acids, and minerals) and ample capacity for growth in diverse, harsh Andean subsistence-production conditions such as water deficiency, salt stress, and soil mineral nutrient deficiency. The objective of this study was to characterize and quantify the genetic diversity among a series of 178 mostly Peruvian amaranth genotypes using 96 biallelic single-nucleotide polymorphism (SNP) markers. A total of 96 co-dominant, biallelic markers were developed using KASPar™ assays on a 96.96 Fluidigm EP1 array system. The 178 amaranth genotypes included white-opaque and white-translucent (vitreous) putative *A. caudatus* seed types, along with black-seeded *A. hybridus* and brown-seeded types, all isolated from among 48 accessions of the CICA-UNSAAC collection. Variation among and within samples and accessions was compared for empirically determined clusters (northern, north-central, south-central, and southern Peruvian Andes). Variation was highest within samples for all groups, but only in white-seeded amaranth was the p-value significant (17.43). The greatest variation among samples was found in the vitreous-seeded group (99.35). The highest average observed heterozygosity within-groups (H_o) was 0.19 in the brown-seeded group, and expected within-group heterozygosity (H_e) was highest in the vitreous-seeded group (0.359). Cluster analysis (UPGMA), PCA and PCO results partitioned the amaranth accessions into six discrete clusters. Clusters did not manifest obvious structure among accessions, which indicates that genetic diversity has been conserved across a broad region of the Peruvian Andes. The diversity characterization pointed to a center of origin and domestication of *A. caudatus* in the Ayacucho-Cusco region of southern Peru.

Keywords *Amaranthus caudatus* L., genetic diversity, SNP markers, Peruvian amaranth, population structure.

ACKNOWLEDGEMENTS

I want to express my deepest gratitude and appreciation to my dear wife Leonilda, my son Ruben, and my beautiful daughter Zoe who supported me through rough times in the last two years as I have pursued my academic dreams. My gratitude also goes to my parents— Santiago and Escolastica—who worked hard to support me so I could be educated. I would also like to thank Dr. Rick Jellen for opening the doors of opportunity by providing unconditional support to me to attend school. I am grateful to each of these individuals for constantly encouraging me to climb to a level I had thought would be impossible to reach.

I also wish to convey my deep appreciation to Ingeniero Aquilino Alvarez (professor at the University of San Antonio Abad in Cusco-Peru) who contributed the genetic material (amaranth accessions) so this research study could take place. I also want to thank Dr. Jeff Maughan for encouraging my use of the Fluidigm instrument that so greatly facilitated this study. In addition, my heartfelt thanks go to Mr. and Mrs. Lilywhite for their invaluable financial support, which came at a critical time after my sponsor withdrew his support at the end of my first semester.

To all the members of my committee —Drs. Rick Jellen, Jeff Maughan, Joshua Udall, and Ing. Aquilino Alvarez — who contributed to help me create a presentable final document: thank you. For their continual support and encouragement in finishing my research study, I'd like to say thank you to my friends in the Genetics and Biotechnology program, especially to Scott Smith for his guidance and expertise with the use of the Fluidigm machine. I am also grateful to David Elzinga for his assistance and expertise with the statistical analysis and to Carolyn Hanson, the department secretary, for her continuous support and encouragement to overcome the countless challenges I faced during my studies at BYU. My gratitude goes out to

my friends Karlin Nelson, Nick Ames, Orrin Probst, Kendra Patton, and Sharisa Nay for their unconditional support and Christlike assistance in several areas of my thesis.

To my sister Laura Jimenez for her unconditional support, guidance, and feedback throughout the whole process of this research study; to my in-laws; and to my family, for their continual support and trust: thank you all. My heart overflows with appreciation for Mr. Luis Espinoza who, through the Ezra Taft Benson Institute, encouraged me and placed interest in this research study.

TABLE OF CONTENTS

LITERATURE REVIEW

INTRODUCTION

The South American Andes is a major center of origin, domestication, and dispersion of high-protein pseudocereal crops. This group includes one of the amaranths known as 'kiwicha' (*Amaranthus caudatus* L.), quinoa (*Chenopodium quinoa* Willd.), and kanihua (*C. pallidicaule* Aellen; Vavilov, 1926; Hawkes, 1999). Kiwicha is an annual diploid (putative paleo-tetraploid, $2n = 2x = 32$) of nutritional importance for intermediate-altitude (<2800 meters elevation) subsistence farm families due primarily to its high lysine content. Amaranth was independently domesticated in ancient Mesoamerica as *A. cruentus* L. and *A. hypochondriacus* L. Kiwicha was re-domesticated and cultivated for thousands of years by Pre-Inca civilizations and dispersed throughout the central and northern Andes by the Inca Empire (Sauer, 1976; Coons, 1977; Costea et al. 2006). Currently, although there are several opinions about the center of domestication of this crop, the actual center remains unknown (Bermejo, 1994; Sauer, 1967). After the Spanish Conquest, amaranth growth and consumption was restricted and stigmatized by cultural and religious opposition that arose due to its inclusion in native sacrificial ceremonies (O'Brien, 1983; Sauer, 1976; Iturbide and Gispert, 1994).

There is a resurgence in amaranth cultivation because its value for human nutrition has been widely recognized (Tucker, 1986; Bressani et al., 1992). For example, amaranth leaves are used as a mineral-rich vegetable like spinach; the dried seed are used as a cereal (grains can be processed without losing protein quality (Pendersen et al., 1987); fresh roots, leaves and dried grains are used in traditional medicine (amaranth is used for stomach flu, diarrhea, and gastroenteritis); and it is a popular ornamental plant due to its brilliant crimson, magenta, or purple inflorescences (Espitia, 1992). In the Andes of Peru, Bolivia, and Ecuador, amaranth is

mostly cultivated for on-farm consumption, though the surplus harvest is often sold or exchanged for other crops in rural fairs. Trade and exchange of amaranth grains has contributed to the dynamic movement of the gene pool (Tumba, 1993).

Amaranth, along with some other Amaranthaceae relatives like quinoa, is problematic taxonomically because morphological descriptors for pigmentation, flower and leaf morphology, and other characters are notoriously plastic. This phenotypic plasticity is due to environmental influences; exacerbating this problem is the presence of considerable morphological variation within cultivated populations (Sauer, 1967; Espitia, 1992). Amaranth is known to experience outcrossing at a rate up to or exceeding 10% with *A. quitensis* Kunth and *A. hybridus* L. weeds cohabiting cultivated fields and other disturbed areas nearby (A. Alvarez, personal communication). However, some amaranth accessions conserved in gene banks are well characterized phenotypically, and efforts to develop homogeneous cultivars have met with some limited success (A. Alvarez and L. Gomez, personal communication).

Today, molecular markers are being used for genetic characterization of amaranth germplasm and help to recognize redundancy and intra-morphotype variation. Also, molecular markers were used to differentiate genotypes under environmental conditions that confounded their phenotypes (Gonzales, 1997; Costea et al. 2006). Among molecular markers, those most frequently used for genotyping Andean crops were simple sequence repeats (SSRs). Simple sequence repeats are short tandem repeats, usually from one to four bases in length and have conserved flanking sequences. They were proposed as markers by Tautz (1989), who correctly predicted that they would be ubiquitous in eukaryotic genomes and highly polymorphic due to a tendency for DNA polymerase to errantly introduce copy number variation at these sites. Several studies described the use of SSRs to reveal polymorphism in Andean crops and grains

like amaranth (Mallory et al. 2008), kanihua (Vargas et al. 2011), and quinoa (Fuentes et al. 2009; Jarvis et al. 2008; Mason et al. 2005; Maughan et al. 2004; Pissard *et al.* 2006). The primary drawbacks of SSR markers are that they may be difficult to score manually and they are relatively expensive for high-throughput genotyping applications.

In 1998 SNPs (single nucleotide polymorphisms) were introduced as a new marker technology by the Human Genome Project (Kadarmideen et al., 2006). Since then, SNPs have become increasingly important genetic markers for studying the evolutionary history of populations (Gupta et al., 2005). Rafalski et al. (2002) indicated the discovery of SNPs and insertions/deletions, which are the basis of most differences between alleles, has been simplified by developments in second-generation DNA sequencing technology. Following SNP identification, primers can be produced to run automated SNP-detection assays, for example KASPar™ (KBioscience Ltd., Hoddesdon, UK), on platforms like the Fluidigm EP-1 96.96 array imager (Fluidigm Corp, South San Francisco, CA).

Maughan et al. (2009) employed a relatively low-cost, whole-genome reduction strategy to develop an expansive SNP resource for *A. caudatus*. The process involved restriction digestion, magnetic bead-mediated fragment removal, and MID-barcoding of four genotypes to effect 50-fold genome reduction, which in turn was followed by Roche 454-GS FLX pyrosequencing (Branford, CT). Maughan et al. (2011) demonstrated the utility of this marker platform by creating an extensive SNP-based linkage map for *A. caudatus* and by using these markers to identify phylogenetic relationships and measure diversity in a test panel of 41 genotypes that included *A. caudatus*, *A. cruentus*, *A. hybridus* L., *A. hypochondriacus*, *A. powellii* S. Wats., *A. retroflexus* L., and *A. tuberculatus* (Moq.) Sauer. Maughan et al. (2011), after finding that *A. hybridus* samples clustered with all three New World domesticates while the

other wild New World species did not, suggested *A. hybridus* is most likely the progenitor of all three cultivated New World amaranth species.

The development of molecular markers and their validation are significant initial steps to (a) recognize diagnostic genetic characteristics within *in situ* and *ex situ* collections; (b) increase the diversity of core collections of amaranth in the center of domestication; (c) scrutinize and examine genetic modifications or aberrations (natural and artificial) that are present in gene banks (e.g., mixes, hybridizations, dispersion, inter-varietal cross pollination, etc.); and (d) identify phylogenetic relationships within collections that include wild and weedy species along with domesticates. The massive genetic erosion that is thought to have occurred in amaranth since the Conquest has stimulated efforts to create germplasm collections to conserve variation in the Andean region (Izquierdo 1998; Gonzales 2004; Ruiz-Tapia 2003). Thus, most universities in Peru have their own collections (*in situ* and *ex situ*). For example, 1600 accessions of amaranth are maintained by the National University of San Antonio Abad of Cuzco (UNSAAC) and the Center for Andean Crops Research (CICA), while approximately 800 accessions are conserved by the National University of the Altiplano (UNA, Puno, Peru) and by other institutions like the International Potato Center (CIP, Lima, Peru).

The primary aims of this study were to (a) assess and quantify genetic diversity among 48 mostly Peruvian amaranth accessions; (b) determine and compare allele frequencies among accessions that cover a wide geographical range to establish the centers of origin, domestication, and dispersion of this crop; (c) compare genetic variation within and among accessions, and investigate the genetic basis for using seed color as a defining taxonomic character; and (d) validate SNP markers that might be useful to accelerate breeding for higher yield, disease and pest resistances, and tolerance of abiotic stresses such as drought.

MATERIALS AND METHODS

Germplasm and DNA extraction

Forty-eight Peruvian amaranth accessions were included in the analysis and evaluation. These materials were obtained from the CICA-UNSAAC collection in Cuzco, Peru (Table 1). This genetic material was collected by A. Alvarez and his students and has been maintained *ex situ* by a three-year rotating regeneration process to conserve variation and retain seed viability. These accessions are representative of overall Peruvian amaranth variability, following a longitudinal/latitudinal gradient from northwest to southeast along the Andean cordillera. Accessions were also collected from isolated valleys and are meant to represent the elevational range of the crop within Peru (A. Alvarez, personal communication). Two accessions from Ecuador and two from Bolivia were also included in the study. Since most populations contained variation for seed color, having black, brown, vitreous-white (hereafter 'vitreous'), and white-opaque (hereafter 'white') types, four samples per accession were included in the study.

All amaranth samples were grown in greenhouses at Brigham Young University in Provo, Utah. Amaranth seeds were placed in 12cm pots using Sunshine Mix II (Sun Grow, Bellevue, WA), supplemented with approximately 15 pellets of Osmacote Smart-Release® fertilizer (Scotts, Marysville, OH). Plants were maintained at 25^{o}C with a photoperiod of 12 hours. DNA was extracted from 30 mg of young leaves of each sample plant. The accession and seed-color identities of each sample were recorded. The extraction procedure was described by Sambrook et al. (1989) with modification by Todd and Vodkin (1996). The DNA samples were quantified in a Nanodrop ND 1000 Spectrophotometer (NanoDrop technologies Inc., Montchanin, DE).

SNP primer selection

A total of 11,038 potential SNP genetic markers for amaranth were previously developed and reported by Maughan et al. (2009, 2011). All of these genetic markers were published and deposited in the GenBank database under batch number 2009A (GenBank: ss161123993 to ss161151650; build B131). From this marker set, we selected a subset of 96 highly polymorphic markers distributed across the 16 chromosomes (Table 2; Maughan et al. 2011). The SNP assays were previously determined to be polymorphic between parents of an amaranth mapping population (Maughan et al. 2009, 2011). The primers were synthesized by Integrated DNA Technologies (Coralville, IA).

Single nucleotide polymorphism (SNP) analysis

For genotyping, 5 µl of master mix was prepared, consisting of 2.25 µl of gDNA (20 µg/µl), 2.5 µl of 2xKASPar ™ reagent (KBioscience Ltd., Hoddesdon, UK), and 0.25 µl of 20X GT loading buffer (Fluidigm Corp., South San Francisco, CA). In addition, 4 µl of KASPar™ assay was mixed, including 0.50 µl of KASPar™ primer mix (12 µM of two forward allele specific primers, 7.5 µM common reverse primer, and 11.5 µl of DNase-free water), 2X assay loading reagent (Fluidigm Corp., South San Francisco, CA) and 1.44 µL DNase-free water for each assay inlet. The KASPar assay primers mix and samples were loaded into the 96.96 nanofluidic chip. The PCR-based assay mix on the chip was then mixed and amplified using an IFC Controller HX and FC1 thermal cycler (Fluidigm Corp., South San Francisco, CA) using the manufacturer's protocol. The PCR thermal cycling parameters were 70ºC for 30 min and 25ºC for 10 min using HotStar Taq polymerase (Qiagen Corp., Valencia, CA), with initial activation at 94ºC for 15 min. This cycle was followed by touchdown amplification of 10 cycles at 94ºC for 20 sec, 65ºC for 1

min with a decreasing cycle of 0.8ºC; a second touchdown of 26 cycles at 94ºC per 20 sec, 57ºC per 1 min; and a hold cycle at 20ºC for 30 sec. Fluorescent images depicting amplification on the array were acquired on the EP-1™ System (Fluidigm Corp., South San Francisco, CA) and the data were analyzed with Fluidigm SNP genotyping Analysis Software®.

Data Analysis

The SNP assays generated biallelic data, reflecting the diploid structure of the *A. caudatus* genome. Heterozygosity across all loci was used as a genetic variability indicator and was measured by determining the proportion of individuals that were homozygous or heterozygous for each particular SNP marker. For each locus the observed heterozygosity (H_o) was calculated as follows: H_0 = [number of heterozygotes at a locus] x [total number of individuals surveyed]. The expected heterozygosity (H_e) for a given marker was calculated by relating the observed allele frequencies to the expected frequency of heterozygotes based on Hardy-Weinberg Equilibrium. For instance, for a single locus with two alleles X and Y, whose frequencies are p(X) and q(Y), where p+q=1, the frequencies of the three possible genotypes are given by p2+2pq+q2=1.

Gene diversity and differentiation statistics

Allele frequencies and genetic diversity were calculated for each population based on within-population data. The parameters calculated were as follows: number of alleles per locus ($A_{x,y}$), observed heterozygosity (H_o), and expected heterozygosity (H_e), as per Brown and Weir (1983). Gene differentiation parameters were measured by Nei's genetic diversity statistic (Nei, 1973; Nei and Chesser, 1993). Total gene diversity (H_t) was separated into three sub-components, D_{gt},

H_s, and D_{gt}, where $H_t = H_s+D_{gs}+D_{gt}$; H_s is within population gene diversity; D_{gs} is a component of diversity resulting from population division; and D_{gt} is a component of diversity resulting from geographical division. Since we based our analyses on biallelic SNP genotyping, $H_t = 2\bar{a}(1-\bar{a})$ and $D_{st}= \sigma^2 x$, where $\bar{a}$ and $\sigma^2 x$ are the mean and variance of the frequency of alleles among the total set of genotypes (178 accessions; Nei, 1973). Levels of differentiation within geographical groups were represented by $G_{st} = [(H_t-H_s)/H_t]$, where H_t and H_s were calculated separately based on measurements of allele frequency from within-accession data (up to four samples from each accession but of the same seed type). Geographical groupings were based on provincial origin and within four broad latitudinal groupings (Table 3). If there were only two alleles at a locus, G_{st} became identical to F_{st}. Absolute differentiation (D_m) was calculated as follows: $D_m = {}^{s}/_{s-1}\,(H_t - H_s)$, where: 's' is the number of accessions for group; H_t is total gene diversity; and H_s is within-gene diversity (Nei, 1977). The parameter D_m is independent of the genetic diversity within sub populations. The coefficient of absolute gene differentiation G'_{st}, was determined using the formula $G'_{st} = D_m/(H_t + D_m)$.

AMOVA and multivariate analyses

Analysis of molecular variance (AMOVA; Excoffier et al, 1992) was performed to describe the genetic variation among and within accessions (groups) and to quantify the contribution of each accession to the population genetic structure at latitudinal level using the Arlequin program (Excoffier et al, 2005). Within-population variation is a calculation using average of gene diversity and percentage of polymorphic loci values. Average gene diversity is the probability that two samples are different at a homologous locus, or the percentage of polymorphic loci between two samples, and is therefore indicative of allelic richness within a population (Nei,

1987). The AMOVAs were calculated as hierarchical analyses of variance. The total variances were thus partitioned into several covariance components: for example, inter-sample (among all samples individually, regardless of seed morphology); intra-accession or among-sample (same or different seed morphologies within a single accession); and among accessions for the same seed morphology (i.e., among all white-seeded samples on an accession-by-accession basis).

Fixation Index (F_x) was calculated using the covariance components of the AMOVA. Fixation index values measured the effects of inbreeding in a population, thus providing an indication of the magnitude of genetic divergence among populations (Hartl and Clark, 1997).

Three different multivariate analyses were performed to group amaranth samples: principal component analysis (PCA), principal coordinate analysis (PCO), and UPGMA-based cluster analysis. To perform multivariate analyses raw SNP data were used to generate a matrix. Principal component and cluster analyses was performed from the SNP similarity matrix using JMP 9 software (SAS Institute, Cary, NC). Principal coordinate analysis (PCO) was performed using PAST v2.05 software (Hammer et al., 2001). Dendrograms were created based in the dissimilarities matrices and UPGMA (unweighted pair group method with arithmetic mean) to demonstrate the organization and structure of the amaranth diversity.

RESULTS

A total of 48 *A. caudatus* accessions from nine Peruvian departments (Ancash - 4, Apurimac - 4, Arequipa - 1, Ayacucho - 9, Cajamarca - 4, Cusco - 17, Junin - 4, La Libertad - 1, and Lima - 1), two Bolivian departments (Beni -1 and Tarija - 1), and two Ecuadorian provinces (Carchi - 1 and

Chimborazo - 1) were analyzed using 96 codominant SNP markers. From the Fluidigm genotyping platform, 94% of the data were clearly distinguishable and therefore scored either manually or using the automated genotype-calling software provided by the manufacturer.

Initial data analysis, in which we pooled different-colored seed samples on an accession-by-accession basis, did not resolve accessions into latitudinal or altitudinal groupings using PCA, PCO, or cluster analyses. Consequently, we reevaluated the data by partitioning instead on a per-sample basis using seed morphology (color and opacity) as the distinguishing criterion. We had previously separated out seed from four different types (white, black, brown, and vitreous) if they were present in a 1 gram sample from each accession; thus, a total of 178 individual samples were analyzed to evaluate the genetic diversity and genetic relationships among the accessions (Fig 1; Table 1).

Genetic diversity measurements for four latitudinal groups of amaranth (empirical groupings) are given in Table 3. Comparisons among means reveals the spatial distribution of genetic diversity in this set of Peruvian amaranth samples. For example, H_o in white-seeded amaranth appeared to show a clinal trend from north to south, with higher values in the south. Observed heterozygosity followed the same trend in the brown- and vitreous-seeded groups. In contrast, the black-seeded samples showed increasing heterozygosity along a south-to-north latitudinal distribution. These observations were confirmed by the G'_{ST} and H_T values for black- and white-seeded amaranths (Table 3).

The results of gene diversity measurements and AMOVA are presented in Table 3. The SNP genotyping detected a high level of polymorphism among and within Peruvian amaranth accessions. Across the 96 SNP loci, 45.75 were polymorphic on average among the four black-

seeded, *A. hybridus*-type latitudinal groups, with an average gene diversity of 0.2045 (+/-0.10). Average number of polymorphic loci dropped to 36.75 among the white-seeded *A. caudatus*-type amaranth groups, with an average of gene diversity measure of 0.1776 (+/-0.09). Similarly, in the brown-seeded amaranths the average number of polymorphic loci was 23.75, with 0.2362 (+/-0.13) as the average gene diversity. The vitreous-seeded amaranth groups had the lowest average number of polymorphic loci at 18.25, but with the highest within-group average gene diversity at 0.3787 (+/-0.22).

The genetic diversity of amaranth within and among samples and seed-color groups is shown on the scatter diagrams obtained from the PCA, the two most influential components of which explained 27% and 13.24% of the total variance (Fig. 2). Principal coordinate analysis (PCO) identified two highly significant components, which explained 26% and 9.48% of the genetic variation, respectively. These two analyses together highlighted definitive genetic compositional characteristics of white- (*A. caudatus*), black- (*A. hybridus*), brown- (*A. caudatus*-*A. hybridus* hybrids), and vitreous-seeded amaranths.

The dendrogram generated from the NJ analysis resolved into three main groups (Fig. 3). Brown-, vitreous-, and white-seeded types, with an occasionally interspersed black-seeded sample, composed the first group, which was clustered apart from the two other groups. The second group included mainly black-seeded *A. hybridus* with some vitreous- and brown-seeded types that could be hybrids or introgression products. The third group was composed of mostly white-, brown-, and vitreous-seeded samples.

Cluster analysis of the similarity data showed a similar relationship, with the samples falling out into four clusters. The results displayed as a dendrogram in Fig. 4 are clearly

separated into two great clusters at 71% similarity. While one cluster contained all the white-seeded *A. caudatus*-type samples, the other was subdivided into a large subcluster dominated by black-seeded *A. hybridus* types and a small subcluster containing seven of the vitreous-seeded types, mostly from Cusco and Ayacucho. The branches for the two *A. hybridus*-vitreous subclusters were separated at level of 33%. The two subclusters of the other great cluster were separated at a level of 50% and were both characterized by non-differentiated groups of mostly white-, brown-, and vitreous-colored seeds, with black-seeded *A. hybridus* samples interspersed.

A second cluster analysis was performed using all the white-, black-, and the nine vitreous-seeded amaranth samples that clustered with *A. hybridus* in Fig. 4 (Fig. 5). The results showed a separation of white- and black-seeded groups at a level of 55%. Among the white-seeded *A. caudatus* samples were black-seeded types from Apurimac (2), Junin, and Ayacucho. One white-seeded sample from Cajamarca, Caj397W, grouped by itself with a similarity value of 28%. The vitreous-seeded group formed a separate cluster at 27% similarity.

DISCUSSION

Earlier studies evaluating the evolutionary relationships within *Amaranthus* (cultivated and wild relatives) had been made during the past 15 years using lower-resolution marker techniques such as RFLPs (Chan and Sun, 1997), RAPDs (Mandal and Das, 2002), and SSRs (Mallory et al., 2008). These studies generated data that led to various hypotheses about the geographical location of the centers of origin and domestication of *A. caudatus* and established tentative phylogenetic links among Andean accessions. The present study detected high levels of genetic variation within and among *A. caudatus* accessions and verified that, although there is abundant

evidence for interspecific introgression products (brown- and some vitreous-seeded samples in our data set), *A. caudatus* represents a unique gene pool sampled via domestication bottleneck(s) from its more genetically diverse *A. hybridus* progenitor. Additionally, biallelic SNP-based marker analyses uncovered for the first time evidence for a unique but somewhat cryptic gene pool of vitreous-seeded *Amaranthus* domesticates in the Andean region. Thus, this study corroborated previous studies of amaranth characterization using SSRs markers, RAPDs markers, and RFLPs markers (Mallory et al., 2008; Mandal and Das, 2002; Transue et al., 1994; Chan and Sun, 1997; Chan and Sun, 1997; Xu and Sun, 2001).

Our data indicate that weedy amaranth (black-seeded amaranth, *A. hybridus*), perhaps came from Mesoamerica, arrived to the northern Andes and later migrated southward, where the *A. caudatus* domestication occurred. This hypothesis is supported by the high G_{ST} levels present in black seeded amaranth in the North and North-Central regions. The domestication event in the southern regions is born out by the high diversity measures of white-seeded samples from the South. Interregional movement of Andean crops was a prominent economic activity during the Inca Empire period (Sauer, 1950, 1957, 1967). From southern valleys, the cultivation of *A. caudatus* (white seeded) would have spread to nearby valleys and later throughout the Andes, as is confirmed by PCA and PCO analysis.

The amaranth diversity trees (N-J tree and cluster dendrograms) indicated that a wide-ranging admixture of *A. hybridus* and *A. caudatus* has occurred in the Andean region. The results indicated two possibilities: first, there has been extensive interchange of kiwicha genetic material within the Andean region via rural seed fairs *(minkas*)*;* and second, weedy species are capable of interchanging genes with cultivated amaranths - an observation widely accepted as fact by kiwicha farmers and breeders, and owing to the commonality of *A. hybridus* and *A. quitensis*

weeds in and around kiwicha fields (Sauer, 1957). These data were confirmed by PCA (Fig 2), PCO tests (Fig. 6, Table 2), and cluster analysis (Fig. 3, 4). The cluster analysis performed by JMP 9 did not show any geographical stratification among accessions. Indeed, most accessions were admixed through the identified clusters. The general pattern of clustering showed a distinct group of black-seeded *A. hybridus* types, with other *A. hybridus* samples interspersed among *A. caudatus* and brown- and vitreous-seeded putative hybrids (Figs. 3, 4). In addition, analyses revealed the presence of a small, genetically distinct cluster of vitreous-seeded types.

The results obtained by PCA, PCO and cluster analysis shows that spatial distribution of seed color in amaranth is not simple. There is a between- population structure of Peruvian amaranth indicated by high levels of G'_{ST}, but this is not correlated with the distance. The lack of spatial and geographical distribution and structure among non-white-seeded amaranth suggests extensive intercrossing of the amaranth across to the Andean valleys, presumably as a result of crop distribution during the Incas time and currently by trading among farmers or the process of domestication was recently, thus there is not enough time for diversification.

These data are also very interesting with respect to the vitreous-seeded amaranth samples. The PCA and PCO plots (Figs. 5 and 6) indicate two types of vitreous-seeded samples (Figs. 2-6): those genetically intermediate between the white- and black-seeded types – perhaps hybrids - and a second group clustering by itself and of unique genotypic composition (Fig. 6B). Table 3 illustrates that the vitreous-seeded group actually had the lowest average number of polymorphic SNP loci at 18.25, perhaps due to selection for some trait (vitreous seed?) or traits present within this subset of plants in multiple locations. At the same time, this group had the highest within-group average SNP diversity at 0.3787 (+/-0.22), which is not surprising given that there are clearly two types: black-white “hybrids” and the separate group of 10 accessions of unique

genetic composition (Fig. 6B). These results suggest that at least one vitreous-seeded type of *Amaranthus* in the Andes has been, or is currently, under selection, perhaps for genes moving into *A. caudatus* from *A. hybridus*. We have wondered if the gene(s) for vitreous seed morphology is present in *A. hybridus* but is masked by the presence of black pigmentation, and if this trait has been under selection by subsistence farmers for some economic advantage it confers.

This situation illustrates the notion that limited gene flow from weedy to cultivated amaranths in the Andes is advantageous in that the wild species provide a persistent reservoir of genetic variation. Presumably, these wild species harbor tremendous variation for resistance to pests, diseases, and abiotic stresses that could threaten the crop. In addition, these weeds serve as a *de facto* buffer against genetic erosion should farmers continue the pattern of discarding minor traditional crops in favor of maize and "alien" cash crops. At the same time, the presence of wild species is a disadvantage for creating pure-line kiwicha varieties and for accurately cataloging genetic diversity – for example, in putting together a core collection - in *ex situ* collections.

The study of genetic diversity among accessions from different regions is important for gene pool conservation, perpetuation, and survival of *A. caudatus*. The present SNP analysis revealed the presence of abundant genetic diversity among and within Peruvian amaranths and described a powerful tool – biallelic SNP markers – for measuring that genetic variation. These markers will next be applied to the entire UNSAAC-CICA kiwicha collection and used to assemble a core collection representing the preponderance of genetic variation found in *A. caudatus*.

CONCLUSIONS

Overall, the Peruvian kiwicha samples showed a high level of genetic diversity (G_{ST} = 0.494; H_o = 0.185; H_e = 0.332). Considering previous studies and the present results, white- (*A. caudatus*), black- (*A. hybridus*), brown-, and vitreous-seeded (hybrids) Andean amaranths represent a rich, diverse set of gene pools that are nonetheless capable of exchanging genes in the field. Thus, the collection and preservation of Peruvian amaranth (CICA germplasm) are important for future research.

In summary, this study has demonstrated that *A. caudatus* was most likely domesticated and later distributed throughout the Andean valleys from what is now southern Peru. In addition, we provided evidence for a unique gene pool of vitreous-seeded Andean amaranths that apparently arose from hybridization between *A. hybridus* and *A. caudatus*. Future studies will be needed to identify the basis of selection for this seed morphotype, and to investigate the underlying genetic basis of the morphological variation observed in the broader Andean amaranth germplasm, for example the rest of the CICA collection and a larger sample of Bolivian accessions.

REFERENCES

Bermejo, J.E.H., and Leon, J. 1994. Neglected Crops: 1492 from a different perspective. FAO, Rome, Italy.

Bressani, R. 1989. The proteins of grain amaranth. Food Rev Int 5; 13-38.

Brown, A.H.D., and Weir, B. 1983. Measuring genetic variability in plant populations. In: Tanksley S. D. and Orton T.J. (eds) Isozymes in Plant Genetics and Breeding, Part A, pp. 219—239. Elsevier Sciences, Amsterdam.

Chan, K.F., and 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.

Coons, M.P. 1977. The Status of Amaranthus hybridus L. in South America. Ciencia y Naturaleza. 18: 80-87.

Costea, M., Brenner, D., Tardif, F., Tan, Y., and Sun, M. 2006. Delimitation of Amaranthus cruentus L. and Amaranthus caudatus L. using micromorphology and AFLP analysis: an application in germplasm identification. Genet Resour Crop Evol 53: 1625-1633.

Espitia, E. 1992. Amaranth germplasm development and agronomic studies in Mexico. Food Rev Int 8; 71-86.

Excoffier, L., Laval, L.G., and Schneider, S. 2005 Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol Bioinform Online 1:47-50.

Excoffier L., Smouse, P., and Quattro, J. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genet 131:479 -491.

Falush, D., Stephens, M., and Pritchard, J.K, 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genet 164:1567-1587.

Falush, D., Stephens, M., and Pritchard, J.K. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes, doi: 10.1111/j.1471-8286.2007.01758.x

Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle.

Felsenstein, J. 1989. PHYLIP -- Phylogeny Inference Package (Version 3.2). Cladistics 5: 164-166.

Fuentes, F.F., Martinez, E.A., Hinrichsen, P.V., Jellen, E.N., and Maughan, P.J. 2009. Assessment of genetic diversity patterns in Chilean quinoa (*Chenopodium quinoa* Willd.) germplasm using multiplex fluorescent microsatellite markers. Conserv Genet 10:369-377

Gepts, P. 1990. Biochemical evidence bearing on the domestication of Phaseolus (Fabaceae) beans. Econ Bot 44: 28–38.

Gonzáles, R., and Terrazas, F. 2004. Ferias de semillas en el microcentro de diversidad de tubérculos andinos (Tas) Candelaria "componente de la estrategia de monitoreo y apoyo a

los flujos locales de mantenimiento de la diversidad de Tas". En: Libro de Resúmenes del XI Congreso Internacional de Cultivos Andinos. PROINPA – Ministerio de Asuntos Campesinos y Agropecuarios. Cochabamba – Bolivia, 3 -6 de febrero 2004. Pág. A-24

González, J.M., and Bressani, R. 1987. Guide to amaranth grain cultivation: summary of experiences at the INCAP experimental farm. Amaranth Newsl 2:5-7.

Gupta, P.K., Rustgi, S., and Kulwal, P.L. 2005. Linkage disequilibrium and association studies in higher plants: present status and future prospects. Plant Mol. Biol., 57, 461–485.

Green, J.M., Barker, H.A., Marshall, E.J.P., Froud-Willaims, R.J., Peters, N.C.V., Arnold, G.M., Dawson, K., and Karp, A. 2001. Microsatellite analysis of the inbreeding grass weed Barren Brome (Anisantha sterilis) reveals genetic diversity at the between and within farm scales. Mol. Eco., 10, 1035-1045.

Hammer, O., Harper, D.A.T., and Ryan, P.D. 2001. PAST: paleontological statistics software package for education and data analysis. Paleontologica Electronica 4: 9.

Hawkes, J.G. 1999. The evidence for the extent of N.I. Vavilov's new world Andean centres of cultivated plant origins. Genet Resour Crop Evol 46:163–168.

H.O.D. 2011. Notepad++ [computer software]. [http://notepad-plus-plus.org/release/5.8.7].

Hubisz, M.J., Falush, D., Stephens, M., and Pritchard, J.K, 2009. Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour, 9:1322-1332.

Iturbide, G.A., and Gispert, M. 1994. Grain amaranths (Amaranthus spp.). p. 93–101. In J.E. Hernandez-Bermejo and J. Leon (ed.) Neglected crops: 1492 from a different perspective. FAO, Rome.

Izquierdo, J., and Roca, W. 1998. Under-utilized Andean food crops: status and prospects of plant biotechnology for the conservation and sustainable agricultural use of genetic resources. FAO – ISHS.

Izquierdo, J. 1998. biotecnología y seguridad alimentaria: promesas, realidades y desafíos. un estudio crítico. en: III Latin american meeting on plant biotechnology (REDBIO´98), 1-5 June 1998, Havana, Cuba.

Jarvis, D.E., Kopp, O.R., Jellen, E.N., Mallory, M.A., Pattee, J., Bonifacio, A., Coleman, C.E., Stevens, M.R., Fairbanks, D.J., and Maughan, P.J. 2008. Simple sequence repeat marker development and genetic mapping in quinoa (*Chenopodium quinoa* Willd.). J Genet 87:39-51.

Kadarmideen, H.N., von Rohr, P., and Janss, L.L. 2006. From genetical genomics to systems genetics: potential applications in quantitative genomics and animal breeding. Mamm Genome 17:548–564.

Kimuraz, M., and Crow, J.F. 1964. The number of alleles that can be maintained in a finite population. Genet 49: 725-738.

Ladizinsky, G. 1985. Founder effect in crop-plant evolution. Econ Bot 39:191–199.

Liu, K., and Muse, S. 2005. PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinform 21:2128–2129.

Mallory, M.A., Hall, R.V., McNabb, A.R., Pratt, D.B., Jellen, E.N., and Maughan, P.J. 2008. Development and characterization of microsatellite markers for the grain Amaranths. Crop Sci 48:1098–1106.

Mandal, N., and Das, P.K. 2002. Intra- and interspecific genetic diversity in grain Amaranthus using random amplified polymorphic DNA markers. Pl Tissue Cult 12:49–56.

Mason, S.L., Stevens, M.R., Jellen, E.N., Bonifacio, A., Fairbanks, D.J., Coleman, C.E., McCarty, R.R., Rasmussen, A.G., and Maughan, P.J. 2005. Development and use of microsatellite markers for germplasm characterization in quinoa (Chenopodium quinoa Willd.). Crop Sci 45:1618-1630.

Maughan, P.J., Bonifacio, A., Jellen, E.N., Stevens, M.R., Coleman, C.E., Ricks, M., Mason, S.L., Jarvis, D.E., Gardunia, B.W., and Fairbanks, D.J. 2004. A genetic linkage map of quinoa (Chenopodium quinoa Willd.) based on AFLP, RAPD, and SSR markers. Theor Appl Genet 109:1188-1195.

Maughan, P.J., Smith, S.M., Fairbanks, D.J., and Jellen, E.N. 2011. Development, characterization, and linkage mapping of single nucleotide polymorphisms in the grain amaranths (Amaranthus sp.) Pl Genome (in press)

Maughan, P.J., Yourstone,, S., Jellen, E.N., and Udall, J.A. 2009. SNP Discovery Via Genomic Reduction, Barcoding and 454-Pyrosequencing in Amaranth. Pl Genome 2: 260-270.

Morden, C.W., Doebley, J.F. and Schert, K.F. 1989. Allozyme variation in Old World races of Sorghum bicolor (Poaceae). Am J Bot 76: 247–255.

Nei, M., and Chesser, R.K.1983 Estimation of fixation indices and gene diversities. Ann Hum Genet 47: 253–259.

Nei, M, and Li, W.H. 1979. Mathematical-model for studying genetic-variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273.

Nei, M., and Tajima, F. 1983. Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. Genet 105:207–217.

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

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

O'Brien, G.K., and Price, M.L. 1983. Amaranth: Grain & Vegetable Types. ECHO Technical Note.

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

Pissard, S., Max-Audit, I., Skopinski, L., and Vasson, A. 2006. Pyruvate kinase deficiency in France: a 3-year study reveals 27 new mutations. Br J Haematol 133:683-689.

Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genet 155:945-959.

Rafalski, A. 2002. Applications of single nucleotide polymorphism in crop genetics. Curr Opin Biol 5: 94-100.

Rambaut, A., and Drummond, A.J. 2007. Fix Tree [computer software]. [http://tree.bio.ed.ac.uk/software/treeedit/].

Ruiz-Tapia, E. 2003. Historia de las ferias de semilla en el distrito de Pomata, Puno. Informe anual de ejecución del proyecto conservación in situ de los cultivos nativos y sus parientes silvestres. Puno, Perú. 103 p.

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

Sauer, J.D. 1950. The Grain Amaranths: a Survey of their History and Classification. Ann Missouri Bot Garden 37:561-631.

Sauer, J.D. 1957. Recent migration and evolution of the dioecious amaranths. Evol 11:11-31.

Sauer, J.D. 1967. The grain amaranths and their relatives: a revised taxonomic and geographic survey. Ann Missouri Bot Garden 54:103-137.

Shannon, C.E., and Weaver, W. 1949. The mathematical theory of communications. Urbana, Il: University of Illinois Press.

Sokal, R.R., and Rohlf, F.J. 1981. Biometry, 2nd edn. W. H. Freedman and Co., New York.

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

Todd, J.J., and Vodkin, L.O. 1996. Duplications that suppress and deletions that restore expression from a chalcone synthase multigene family. Plant Cell 8:687–699.

Transue, D.K., Fairbanks, D.J., Robinson, L.R., and Andersen, W.R. 1994. Species identification by RAPD analysis of grain amaranth genetic resources. Crop Sci 34:1385–1389.

Tucker, J.B. 1986. Amaranth: the once and future crop. BioSci 36:9–13.

Tumba, Z.J. 1993. Cultivo de la amaranth (*Amaranthus caudatus* L.). Instituto Nacional de Investigación Agraria (Peru). no. 2

Van Ooijen, J.W., and Vorrips, R.E. 2001. JoinMap® Version 3.0, Software for the calculation of genetic linkage map. Plant Research International. Wageningen, The Netherlands.

Vargas, A., Elzinga, D.B., Rojas-Beltran, J.A., Bonifacio, A., Geary, B., Stevens, M.R., Jellen, E.N., and Maughan, P.J. 2011. Development and use of microsatellite markers for genetic diversity analysis of canahua (*Chenopodium pallidicaule* Aellen). Genet Resour Crop Evol (in press)

Vavilov, N. 1926. Studies on the origin of cultivated plants. Bull Applied Bot 16:139-248

Wang, J., Lin, M., Crenshaw, A., Hutchinson, A., Hicks, B., Yeager, M., Berndt, S., Huang, W.Y., Hayes, R.B,, and Chanock, S.J. 2009. High-throughput single nucleotide polymorphism genotyping using nanofluidic Dynamic Arrays. BMC Genom 10:561.

Wegerle, N., and Zeller, F.J. 1995. Grain amaranth (*Amaranthus ssp*.): cultivation, breeding and properties of an old Indian plant.. J Agron Crop Sci 174: 63–72.

Xu, F., and 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.

TABLES AND FIGURES

TABLES

Table 1. Passport data of amaranth (*A. caudatus.)* accessions used in this study.

No.	BYU No.	CICA No.	Department	Province	Country	Longit	Latit	Alt
1	Chi97B	CKC097-4B	Chimborazo	Guano	Ecuador	-1.6063	-78.6413	2919
2	Caj401B	CKC401-4B	Cajamarca	Cajamarca	Peru	-7.1419	-78.5153	2627
3	LaL691B	CKC691-1B	la Libertad	Huamachuco	Peru	-8.1090	-78.0641	3085
4	Anc118B	CKC118-2B	Ancash	Carhuaz	Peru	-9.2814	-77.6464	2688
5	Jun20B	CKC020-3B	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
6	Jun22B	CKC022-3B	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
7	Jun379B	CKC379-4B	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
8	Lim68B	CKC068-2B	Lima	Yauyos	Peru	-12.2272	-75.7672	1179
9	Aya26B	CKC026-1B	Ayacucho	Cangallo	Peru	-13.2186	-74.3606	3598
10	Aya28B	CKC028-4B	Ayacucho	Lucanas	Peru	-13.2186	-74.2017	1802
11	Aya27B	CKC027-2B	Ayacucho	Lucanas	Peru	-13.2186	-74.2017	3876
12	Cus137B	CKC137-2B	Cusco	Urubamba	Peru	-13.3042	-72.1167	3105
13	Apu321B	CKC321-1B	Apurimac	Abancay	Peru	-13.5474	-72.7047	2688
14	Cus16B	CKC016-3B	Cusco	Quispicanchi	Peru	-13.6356	-71.7375	3253
15	Apu70B	CKC070-3B	Apurimac	Andahuylas	Peru	-13.6545	-73.3825	2926
16	Apu71B	CKC071-4B	Apurimac	Cotabambas	Peru	-13.7571	-72.3636	1008
17	Cus141B	CKC141-3B	Cusco	Acomayo	Peru	-13.9167	-71.6811	3376
18	SanL38B	CKC038-3B	San Lorenzo	San Lorenzo	Bolivia	-15.3823	-65.7975	3052
19	Tar36B	CKC036-3B	Tarija	Tarija	Bolivia	-21.5310	-64.7754	2114
20	Caj395B	CKC395-3B	Cajamarca	Cajamarca	Peru	-7.1621	-78.5106	2627
21	Anc44B	CKC044-2B	Ancash	Huaraz	Peru	-9.5333	-77.5333	3052
22	Jun19B	CKC019-4B	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
23	Aya30B	CKC030-2B	Ayacucho	Huamanga	Peru	-13.0581	-74.1344	3287
24	Aya50B	CKC050-3B	Ayacucho	Huamanga	Peru	-13.1901	-74.2205	2746
25	Cus76B	CKC076-4B	Cusco	Urubamba	Peru	-13.3042	-72.1167	3105
26	Cus134B	CKC134-2B	Cusco	Calca	Peru	-13.3235	-71.9618	3105
27	Cus133B	CKC133-4B	Cusco	Urubamba	Peru	-13.3958	-72.0547	3839
28	Cus5B	CKC005-2B	Cusco	Quispicanchi	Peru	-13.5928	-71.7719	3276
29	Apu70B	CKC070-1B	Apurimac	Andahuaylas	Peru	-13.6545	-73.3825	2926
30	Cus67B	CKC067-1B	Cusco	Canchis	Peru	-14.2853	-71.2247	3903
31	Are431B	CKC431-4B	Arequipa	Arequipa	Peru	-16.7236	-71.8585	1822
32	Car195B	CKC195-1B	Carchi	Montufar	Ecuador	0.6037	-77.8330	2897
33	Anc120B	CKC120-3B	Ancash	Huari	Peru	-9.5928	-77.1785	3180
34	Aya25B	CKC025-3B	Ayacucho	Huamanga	Peru	-13.2546	-74.3445	3498
35	Cus154B	CKC154-3B	Cusco	Calca	Peru	-13.4241	-71.8576	3225
36	Apu321B	CKC321-4B	Apurimac	Abancay	Peru	-13.5474	-72.7047	2688
37	Aya377B	CKC377-2B	Ayacucho	Huamanga	Peru	-13.65	-74.65	2746
38	Cus145B	CKC145-4B	Cusco	Canchis	Peru	-14.3730	-71.1778	3975
39	Aya85B	CKC085-4B	Ayacucho	Huamanga (U San Cristobal)	Peru	-14.6939	-74.1241	3022
40	Apu322B	CKC322-4B	Apurimac	Abancay	Peru	-19.3167	-66.0333	3425
41	Caj397B	CKC397-4B	Cajamarca	Cajamarca	Peru	-7.1544	-78.4921	2627
42	Caj398B	CKC398-2B	Cajamarca	Cajamarca	Peru	-7.1584	-78.5015	2627
43	Anc316B	CKC316-1B	Ancash	Huaraz	Peru	-9.5141	-77.5662	3052
44	Cus69B	CKC069-4B	Cusco	Paruro	Peru	-13.7553	-71.8519	3057
45	Cus24B	CKC024-1B	Cusco	La Convencion	Peru	-13.0072	-72.5386	2017

Table 1 (Continued). Passport data of amaranth (*A. caudatus*) accessions used in this study.

No.	BYU No.	CICA No.	Department	Province	Country	Longit	Latit	Alt
46	Cus4B	CKC004-3B	Cusco	Paucartambo	Peru	-13.3168	-71.5958	3294
47	Cus14B	CKC014-2B	Cusco	Paucartambo	Peru	-13.3255	-71.5540	3294
48	Cus777B	CKC077-1B	Cusco	Anta	Peru	-13.4880	-72.4494	3105
49	Aya380B	CKC380-1B	Ayacucho	Cangallo	Peru	-13.5885	-74.3602	3326
50	Cus17B	CKC017-3B	Cusco	Quispicanchi	Peru	-13.6047	-71.5594	3715
51	Caj397BR	CKC397-2BR	Cajamarca	Cajamarca	Peru	-7.1544	-78.4921	2627
52	Caj395BR	CKC395-4BR	Cajamarca	Cajamarca	Peru	-7.1621	-78.5106	2627
53	Anc44BR	CKC044-3BR	Ancash	Huaraz	Peru	-9.5333	-77.5333	3052
54	Jun19BR	CKC019-3BR	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
55	Cus76BR	CKC076-3BR	Cusco	Urubamba	Peru	-13.3042	-72.1167	3105
56	Cus134BR	CKC134-3BR	Cusco	Calca	Peru	-13.3235	-71.9618	3105
57	Cus14BR	CKC014-1BR	Cusco	Paucartambo	Peru	-13.3255	-71.5540	3294
58	Cus154BR	CKC154-2BR	Cusco	Calca	Peru	-13.4241	-71.8576	3225
59	Cus5BR	CKC005-3BR	Cusco	Quispicanchi	Peru	-13.5928	-71.7719	3276
60	Aya377BR	CKC377-3BR	Ayacucho	Huamanga	Peru	-13.65	-74.65	2746
61	Cus145BR	CKC145-2BR	Cusco	Canchis	Peru	-14.3730	-71.1778	3975
62	Are431BR	CKC431-2BR	Arequipa	Arequipa	Peru	-16.7236	-71.8585	1822
63	Chi97BR	CKC097-1BR	Chimborazo	Guano	Ecuador	-1.6064	-78.6413	2919
64	Caj401BR	CKC401-1BR	Cajamarca	Cajamarca	Peru	-7.1419	-78.5153	2627
65	Anc118BR	CKC118-1BR	Ancash	Carhuaz	Peru	-9.2814	-77.6464	2688
66	Anc316BR	CKC316-2BR	Ancash	Huaraz	Peru	-9.5141	-77.5662	3052
67	Jun20BR	CKC020-4BR	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
68	Jun22BR	CKC022-2BR	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
69	Jun379BR	CKC379-2BR	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
70	Lim68BR	CKC068-3BR	Lima	Yauyos	Peru	-12.2273	-75.7672	1179
71	Cus24BR	CKC024-2BR	Cusco	La Convencion	Peru	-13.0072	-72.5386	2017
72	Aya26BR	CKC026-3BR	Ayacucho	Cangallo	Peru	-13.2186	-74.3606	3598
73	Aya27BR	CKC027-3BR	Ayacucho	Lucanas	Peru	-13.2186	-74.2017	3876

Table 1 (Continued). Passport data of amaranth (*A. caudatus*) accessions used in this study.

No.	BYU No.	CICA No.	Department	Province	Country	Longit	Latit	Alt
74	Cus137BR	CKC137-4BR	Cusco	Urubamba	Peru	-13.3042	-72.1167	3105
75	Cus4BR	CKC004-1BR	Cusco	Paucartambo	Peru	-13.3168	-71.5958	3294
76	Cus77BR	CKC077-2BR	Cusco	Anta	Peru	-13.4880	-72.4494	3105
77	Apu321BR	CKC321-2BR	Apurimac	Abancay	Peru	-13.54742	-72.70473	2688
78	Cus17BR	CKC017-4BR	Cusco	Quispicanchi	Peru	-13.6047	-71.5594	3715
79	Cus16BR	CKC016-2BR	Cusco	Quispicanchi	Peru	-13.6356	-71.7375	3253
80	Apu71BR	CKC071-3BR	Apurimac	Cotabambas	Peru	-13.7571	-72.3636	1008
81	Cus141BR	CKC141-2BR	Cusco	Acomayo	Peru	-13.9167	-71.6811	3376
82	Aya85BR	CKC085-3BR	Ayacucho	Huamanga (U San Cristobal)	Peru	-14.6939	-74.1241	3022
83	SanL38BR	CKC038-1BR	San Lorenzo	San Lorenzo	Bolivia	-15.3823	-65.7975	3052
84	Apu322BR	CKC322-2BR	Apurimac	Abancay	Peru	-19.3167	-66.0333	3425
85	Tar36BR	CKC036-1BR	Tarija	Tarija	Bolivia	-21.5310	-64.7754	2114
86	Car195BR	CKC195-4BR	Carchi	Montufar	Ecuador	0.6037	-77.8330	2897
87	LaL691BR	CKC691-2BR	la Libertad	Huamachuco	Peru	-8.1090	-78.0641	3085
88	Cus133BR	CKC133-1BR	Cusco	Chincheros	Peru	-13.3958	-72.0547	3839
89	Cus67BR	CKC067-2BR	Cusco	Canchis	Peru	-14.2853	-71.2247	3903
90	Car195V	CKC195-3V	Carchi	Montufar	Ecuador	0.6037	-77.8330	2897
91	Caj398V	CKC398-1V	Cajamarca	Cajamarca	Peru	-7.1584	-78.5015	2627
92	Anc120V	CKC120-1V	Ancash	Huari	Peru	-9.5928	-77.1785	3180
93	Cus133V	CKC133-2V	Cusco	Chincheros	Peru	-13.3958	-72.0547	3839
94	Aya85V	CKC085-2V	Ayacucho	Huamanga (U San Cristobal)	Peru	-14.6939	-74.1241	3022
95	Apu322V	CKC322-3V	Apurimac	Abancay	Peru	-19.3167	-66.0333	3425
96	Caj395V	CKC395-2V	Cajamarca	Cajamarca	Peru	-7.1621	-78.5106	2627
97	LaL691V	CKC691-3V	la Libertad	Huamachuco	Peru	-8.1090	-78.0641	3085
98	Anc118V	CKC118-3V	Ancash	Carhuaz	Peru	-9.2814	-77.6464	2688
99	Anc44V	CKC044-4V	Ancash	Huaraz	Peru	-9.5333	-77.5333	3052
100	Jun19V	CKC019-2V	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
101	Jun20V	CKC020-2V	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
102	Jun22V	CKC022-4V	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
103	JunC379V	CKC379-1V	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
104	Lim68V	CKC068-1V	Lima	Yauyos	Peru	-12.2273	-75.7672	1179
105	Cus24V	CKC024-4V	Cusco	La Convencion	Peru	-13.0072	-72.5386	2017
106	Aya50V	CKC050-1V	Ayacucho	Huamanga	Peru	-13.1901	-74.2205	2746
107	Aya26V	CKC026-2V	Ayacucho	Cangallo	Peru	-13.2186	-74.3606	3598

Table 1 (Continued). Passport data of amaranth (*A. caudatus*) accessions used in this study.

No.	BYU No.	CICA No.	Department	Province	Country	Longit	Latit	Alt
108	Aya28V	CKC028-2V	Ayacucho	Lucanas	Peru	-13.2186	-74.2017	1802
109	Cus76V	CKC076-2V	Cusco	Urubamba	Peru	-13.3042	-72.1167	3105
110	Cus137V	CKC137-3V	Cusco	Urubamba	Peru	-13.3042	-72.1167	3105
111	Cus4V	CKC004-2V	Cusco	Paucartambo	Peru	-13.3168	-71.5958	3294
112	Cus134V	CKC134-4V	Cusco	Calca	Peru	-13.3235	-71.9618	3105
113	Aya380V	CKC380-2V	Ayacucho	Cangallo	Peru	-13.5885	-74.3602	3326
114	Cus5V	CKC005-4V	Cusco	Quispicanchi	Peru	-13.5928	-71.7719	3276
115	Cus17V	CKC017-1V	Cusco	Quispicanchi	Peru	-13.6047	-71.5594	3715
116	Cus16V	CKC016-4V	Cusco	Quispicanchi	Peru	-13.6356	-71.7375	3253
117	Apu71V	CKC071-2V	Apurimac	Cotabambas	Peru	-13.7571	-72.3636	1008
118	Cus141V	CKC141-1V	Cusco	Acomayo	Peru	-13.9167	-71.6811	3376
119	Cus67V	CKC067-3V	Cusco	Canchis	Peru	-14.2853	-71.2247	3903
120	SanL38V	CKC038-2V	San Lorenzo	San Lorenzo	Bolivia	-15.3823	-65.7975	3052
121	Are431V	CKC431-3V	Arequipa	Arequipa	Peru	-16.7236	-71.8585	1822
122	Tar36V	CKC036-2V	Tarija	Tarija	Bolivia	-21.5310	-64.7754	2114
123	Chi97V	CKC097-2V	Chimborazo	Guano	Ecuador	-1.6064	-78.6413	2919
124	Caj401V	CKC401-2V	Cajamarca	Cajamarca	Peru	-7.1419	-78.5153	2627
125	Aya30V	CKC030-1V	Ayacucho	Huamanga	Peru	-13.0581	-74.1344	3287
126	Aya25V	CKC025-2V	Ayacucho	Huamanga	Peru	-13.2546	-74.3445	3498
127	Cus154V	CKC154-1V	Cusco	Calca	Peru	-13.4241	-71.8576	3225
128	Aya377V	CKC377-1V	Ayacucho	Huamanga	Peru	-13.65	-74.65	2746
129	Cus145V	CKC145-3V	Cusco	Canchis	Peru	-14.3730	-71.1778	3975
130	Anc316W	CKC316-3W	Ancash	Huaraz	Peru	-9.5141	-77.5662	3052
131	Cus69W	CKC069-2W	Cusco	Paruro	Peru	-13.7553	-71.8519	3057
132	Cus24W	CKC024-3W	Cusco	La Convencion	Peru	-13.0072	-72.5386	2017
133	Aya30W	CKC030-3W	Ayacucho	Huamanga	Peru	-13.0581	-74.1344	3287
134	Aya50W	CKC050-2W	Ayacucho	Huamanga	Peru	-13.1901	-74.2205	2746
135	Aya25W	CKC025-1W	Ayacucho	Huamanga	Peru	-13.2546	-74.3445	3498
136	Cus4W	CKC004-4W	Cusco	Paucartambo	Peru	-13.3168	-71.5958	3294
137	Cus77W	CKC077-3W	Cusco	Anta	Peru	-13.4880	-72.4494	3105
138	Aya380W	CKC380-3W	Ayacucho	Cangallo	Peru	-13.5885	-74.3602	3326
139	Cus17W	CKC017-2W	Cusco	Quispicanchi	Peru	-13.6047	-71.5594	3715
140	Cus67W	CKC067-4W	Cusco	Canchis	Peru	-14.2853	-71.2247	3903
141	Car195W	CKC195-2W	Carchi	Montufar	Ecuador	0.6037	-77.8330	2897
142	Caj397W	CKC397-3W	Cajamarca	Cajamarca	Peru	-7.1544	-78.4921	2627
143	Caj398W	CKC398-4W	Cajamarca	Cajamarca	Peru	-7.1584	-78.5015	2627
144	Aya28W	CKC028-1W	Ayacucho	Lucanas	Peru	-13.2186	-74.2017	1802
145	Cus14W	CKC014-3W	Cusco	Paucartambo	Peru	-13.3255	-71.5540	3294
146	Cus154W	CKC154-4W	Cusco	Calca	Peru	-13.4241	-71.8576	3225
147	Aya377W	CKC377-4W	Ayacucho	Huamanga	Peru	-13.65	-74.65	2746
148	Chi97W	CKC097-3W	Chimborazo	Guano	Ecuador	-1.6064	-78.6413	2919
149	Caj401W	CKC401-3W	Cajamarca	Cajamarca	Peru	-7.1419	-78.5153	2627
150	Cus133W	CKC133-3W	Cusco	Chincheros	Peru	-13.3959	-72.0548	3839
151	Caj395W	CKC395-1W	Cajamarca	Cajamarca	Peru	-7.1621	-78.5106	2627
152	Anc118W	CKC118-4W	Ancash	Carhuaz	Peru	-9.2814	-77.6464	2688
153	Anc44W	CKC044-1W	Ancash	Huaraz	Peru	-9.5333	-77.5333	3052
154	Anc120W	CKC120-2W	Ancash	Huari	Peru	-9.5928	-77.1785	3180
155	Jun19W	CKC019-1W	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
156	Jun20W	CKC020-1W	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
157	Jun22W	CKC022-1W	Junin	Huancayo	Peru	-12.2112	-75.1904	3259
158	Jun379W	CKC379-3W	Junin	Huancayo	Peru	-12.2112	-75.1904	3259

Table 1 (Continued). Passport data of amaranth (*A. caudatus*) accessions used in this study.

No.	BYU No.	CICA No.	Department	Province	Country	Longit	Latit	Alt
159	Lim68W	CKC068-4W	Lima	Yauyos	Peru	-12.2273	-75.7672	1179
160	Aya50W	CKC50-4W	Ayacucho	Ayacucho	Peru	-13.1901	-74.2205	2746
161	Aya26W	CKC026-4W	Ayacucho	Cangallo	Peru	-13.2186	-74.3606	3598
162	Aya28W	CKC028-3W	Ayacucho	Lucanas	Peru	-13.2186	-74.2017	1802
163	Aya27W	CKC027-1W	Ayacucho	Lucanas	Peru	-13.2186	-74.2017	3876
164	Cus76W	CKC076-1W	Cusco	Urubamba	Peru	-13.3042	-72.1167	3105
165	Cus137W	CKC137-1W	Cusco	Urubamba	Peru	-13.3042	-72.1167	3105
166	Cus134W	CKC134-1W	Cusco	Calca	Peru	-13.3235	-71.9618	3105
167	Apu321W	CKC321-3W	Apurimac	Abancay	Peru	-13.5474	-72.7047	2688
168	Cus5W	CKC005-1W	Cusco	Quispicanchi	Peru	-13.5928	-71.7719	3276
169	Cus16W	CKC016-1W	Cusco	Quispicanchi	Peru	-13.6356	-71.7375	3253
170	Apu70W	CKC070-2W	Apurimac	Andahuylas	Peru	-13.6545	-73.3825	2926
171	Apu71W	CKC071-1W	Apurimac	Cotabambas	Peru	-13.7571	-72.3636	1008
172	Cus141W	CKC141-4W	Cusco	Acomayo	Peru	-13.9167	-71.6811	3376
173	Cus145W	CKC145-1W	Cusco	Canchis	Peru	-14.3730	-71.1778	3975
174	Aya85W	CKC085-1W	Ayacucho	Huamanga (U San Cristobal)	Peru	-14.6939	-74.1241	3022
175	SanL38W	CKC038-4W	San Lorenzo	San Lorenzo	Bolivia	-15.3823	-65.7975	3052
176	Are431W	CKC431-1W	Arequipa	Arequipa	Peru	-16.7236	-71.8585	1822
177	Apu322W	CKC322-1W	Apurimac	Abancay	Peru	-19.3167	-66.0333	3425
178	Tar36W	CKC036-4W	Tarija	Tarija	Bolivia	-21.5310	-64.7754	2114

Table 2. Map positions of amaranth SNP markers used in this study. The 96 loci were distributed among all 16 linkage groups (LG). Molecular markers displayed on the linkage groups were previously mapped by Maughan et al. (2010).

No	SNP marker	Linkage Group	Position	No	SNP marker	Linkage Group	Position	No	SNP marker	Linkage Group	Position	No	SNP marker	Linkage Group	Position
1	AM18245	_		25	AM27618	3	63.225	49	AM17547	6	60.288	73	AM23934	11	1.516
2	AM19009	1	4.948	26	AM19963	3	72.774	50	AM18657	6	72.694	74	AM26018	11	1.838
3	AM19012	1	5.399	27	AM19961	3	73.079	51	AM19353	6	79.329	75	AM17981	11	17.875
4	AM25261	1	17.716	28	AM19885	3	74.433	52	AM18821	6	79.329	76	AM24013	11	37.333
5	AM18740	1	18.655	29	AM17998	3	83.108	53	AM20065	7	0	77	AM19746	11	61.625
6	AM27655	1	38.959	30	AM23684	3	84.493	54	AM22341	7	15.778	78	AM23989	11	63.71
7	AM23857	1	49.015	31	AM19652	4	1.854	55	AM25606	7	36.262	79	AM17978	12	2.458
8	AM23891	1	51.398	32	AM24743	4	6.516	56	AM21605	7	42.483	80	AM17977	12	2.78
9	AM19452	1	69.562	33	AM23741	4	20.277	57	AM21911	7	44.13	81	AM25474	12	3.329
10	AM19663	1	84.572	34	AM27602	4	40.912	58	AM20181	7	75.5	82	AM17120	12	38.176
11	AM25309	1	118.811	35	AM19427	4	56.915	59	AM20180	7	75.9	83	AM21862	12	50.911
12	AM26133	1	139.095	36	AM19425	4	57.221	60	AM24486	7	76.845	84	AM24766	12	53.112
13	AM22487	1	144.431	37	AM18899	4	61.186	61	AM20405	8	19.154	85	AM18603	13	9.348
14	AM23895	2	14.273	38	AM19442	4	65.023	62	AM19810	8	32.048	86	AM18604	13	9.657
15	AM27607	2	47.863	39	AM24501	5	0	63	AM25708	8	43.492	87	AM24519	13	50.584
16	AM18092	2	58.988	40	AM20896	5	54.8	64	AM21961	8	51.494	88	AM19210	14	0.098
17	AM20886	2	66.596	41	AM24451	5	95.654	65	AM19707	8	64.462	89	AM27633	14	18.041
18	AM18082	2	77.715	42	AM20942	6	0	66	AM21154	8	76.806	90	AM27651	14	29.913
19	AM23808	2	110.164	43	AM20153	6	21.524	67	AM24104	9	31.144	91	AM17729	14	42.357
20	AM20845	3	0	44	AM21305	6	22.72	68	AM27631	9	26.431	92	AM23231	14	43.052
21	AM21439	3	18.835	45	AM25154	6	23.126	69	AM27630	9	26.671	93	AM22082	14	55.374
22	AM21437	3	19.278	46	AM25631	6	23.126	70	AM25418	10	0.975	94	AM19503	15	7.445
23	AM21452	3	22.858	47	AM25494	6	27.936	71	AM20788	10	5.065	95	AM24460	15	29.45
24	AM19559	3	33.733	48	AM20022	6	31.593	72	AM24510	10	6.696	96	AM21276	16	0

Table 3 (following page). Genetic diversity parameters calculated for 96 biallelic SNP loci across 178 *A. caudatus* samples according to geographic origin. "North" contains accessions from latitudes of 0.07 to -9.59 (Ecuador through Ancash); "No Central" from -9.60 to -13.25 (Ancash to Ayacucho); "So Central" from -13.26 to -13.63 (Ayacucho to Cusco); and "South" from -13.64 to -21.53 (Cusco to Bolivia). Explanations: H_T = gene diversity for polymorphic loci in the total collection (same seed color); H_S = diversity among loci within a given sample; G_{ST} = relative degree of gene differentiation among accessions, G'_{ST} = coefficient of absolute gene differentiation; D_m = absolute differentiation.

Seed color	Black				Brown			
Latitude group	North	No Central	So Central	South	North	No Central	So Central	South
No samples	11	12	14	13	9	8	12	10
No polymorphic loci (*n*/96)	37	49	44	53	18	27	15	35
Avg gene diversity (%)	0.2045 (+/_0.10)				0.2362 (+/_0.13)			
Observed heterozygosity	0.180	0.187	0.178	0.154	0.197	0.174	0.181	0.207
Expected heterozygosity	0.307	0.373	0.325	0.297	0.362	0.362	0.358	0.355
Among-accession variation (%)	1.54				1.19			
Within-accession variation (%)	98.46				98.81			
Fixation index	0.015				0.012			
Theta (θ)	0.218	0.337	0.266	0.277	0.447	0.292	0.388	0.376
H_S	0.439				0.462			
H_T	0.859				0.864			
G_{ST}	0.489				0.466			
Dm	0.560				0.537			
G'_{ST}	0.431				0.405			
Seed color	Vitreous				White			
Latitude group	North	No Central	So Central	South	North	No Central	So Central	South
No samples	9	11	10	10	10	13	14	12
No polymorphic loci (*n*/96)	39	15	11	8	19	28	47	53
Avg gene diversity (%)	0.3787 (+/_0.22)				0.1776 (+/_0.09)			
Observed heterozygosity	0.092	0.135	0.103	0.203	0.127	0.194	0.175	0.244
Expected heterozygosity	0.143	0.446	0.416	0.419	0.375	0.283	0.343	0.325
Among-accession variation (%)	0.65				17.43			
Within-accession variation (%)	99.35				82.57			
Fixation index	0.007				0.174			
Theta (θ)	0.505	0.791	0.688	0.721	0.293	0.115	0.261	0.250
H_S	0.455				0.434			
H_T	0.863				0.858			
G_{ST}	0.472				0.494			
Dm	0.543				0.565			
G'_{ST}	0.412				0.437			

Figure 1. Distribution map of all 48 accessions.

Figure 2. Principal component analysis (PCA) of 178 SNP-genotyped Andean amaranths using JMP 9. Numbers correspond to geographically assorted accessions as per Table 1. Explanations: white- (red), black- (green), brown- (yellow), and vitreous-seeded (blue) accessions.

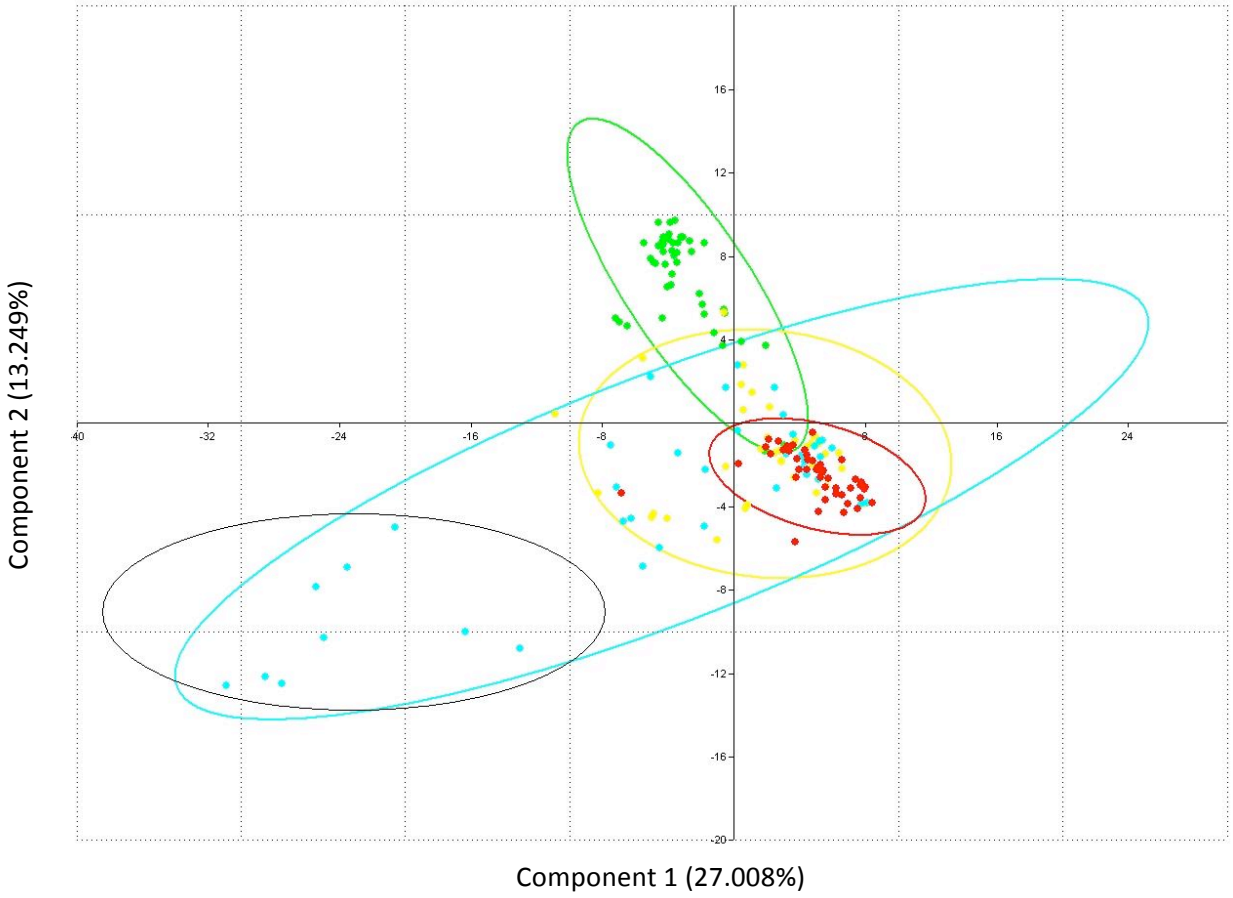


Figure 3 (Following page). Rooted Neighbor-Joining (Nei's genetic distances) dendrogram based on 178 samples and 96 biallelic SNP markers, generated using PAST v2.05 software (Hammer et al., 2001). Accession information is given in Table 1.

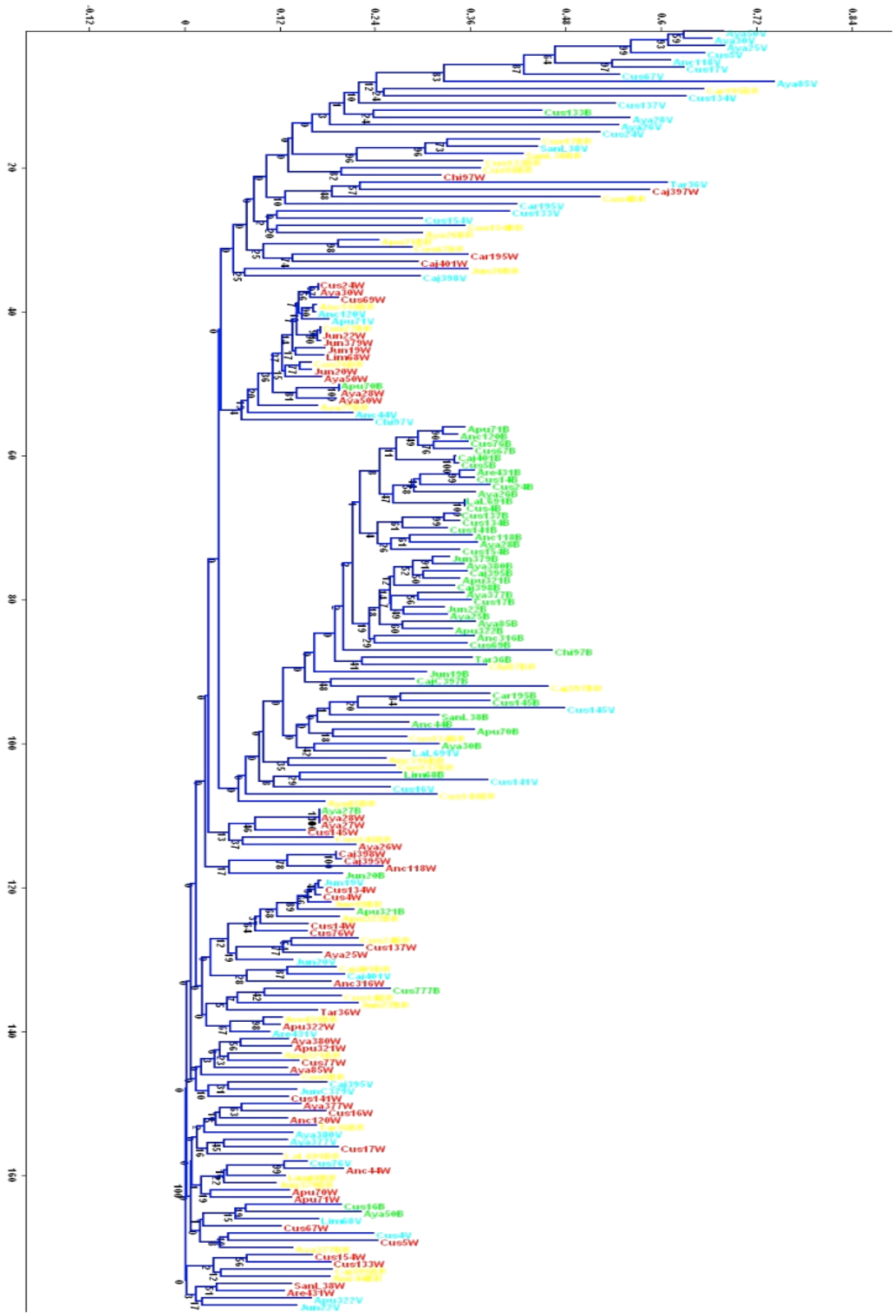

-0.12
0
0.12
0.24
0.36
0.48
0.6
0.72
0.84
20
40
60
80
100
120
140
160
Aya85V
Aya54V
Aya25V
Cus5V
Anc118V
Cus17V
Cus67V
Aya85V
Cus134V
Cus137V
Cus133B
Aya28V
Aya30V
Cus2W
SanL38V
Chi97W
Tar36V
Caj397W
Car195V
Cus133V
Cus154W
Car195W
Caj401W
Caj398V
Cus24W
Aya30W
Cus69W
Jun22W
Jun379W
Jun19W
Lim68W
Jun20W
Aya54W
Apu70B
Aya28W
Aya59W
Anc44W
Chi97V
Apu71B
Anc120B
Cus76B
Cus67B
Caj401B
Cus5B
Anc431B
Cus154B
Cus24B
Aya26B
LaL69B
Cus4B
Cus137B
Cus133B
Cus141B
Anc118B
Aya28B
Cus154B
Jun379B
Aya380B
Caj395B
Apu321B
Caj398B
Aya377B
Cus17B
Jun22B
Aya25B
Aya85B
Apu127B
Anc316B
Cus69B
Chi97B
Tar36B
Jun19B
Caj397B
Car195B
Cus145B
Cus145V
SanL38B
Anc44B
Apu70B
Aya30B
LaL69V
Lim68B
Cus141V
Cus16V
Aya27B
Aya28W
Aya27W
Cus145W
Aya26W
Caj398W
Caj395W
Anc118W
Jun20B
Jun19V
Cus134W
Cus4W
Apu321B
Cus154W
Cus76W
Cus137W
Aya25W
Caj401V
Anc316W
Cus77B
Tar36W
Apu322W
Anc431V
Aya380W
Apu321W
Cus77W
Aya85W
Caj395V
Anc127W
Cus141W
Aya377W
Cus16W
Anc120W
Aya380V
Cus17W
Cus76V
Anc44W
Apu70W
Apu71W
Cus16B
Aya50B
Lim68V
Cus67W
Cus5W
Cus154W
Cus133W
SanL38W
Are431W
Apu322V
Jun22V

Figure 4 (Following page). Cluster analysis dendrogram showing the genetic relationships among 178 amaranth samples based on 96 SNP markers using PAST v2.05 software.

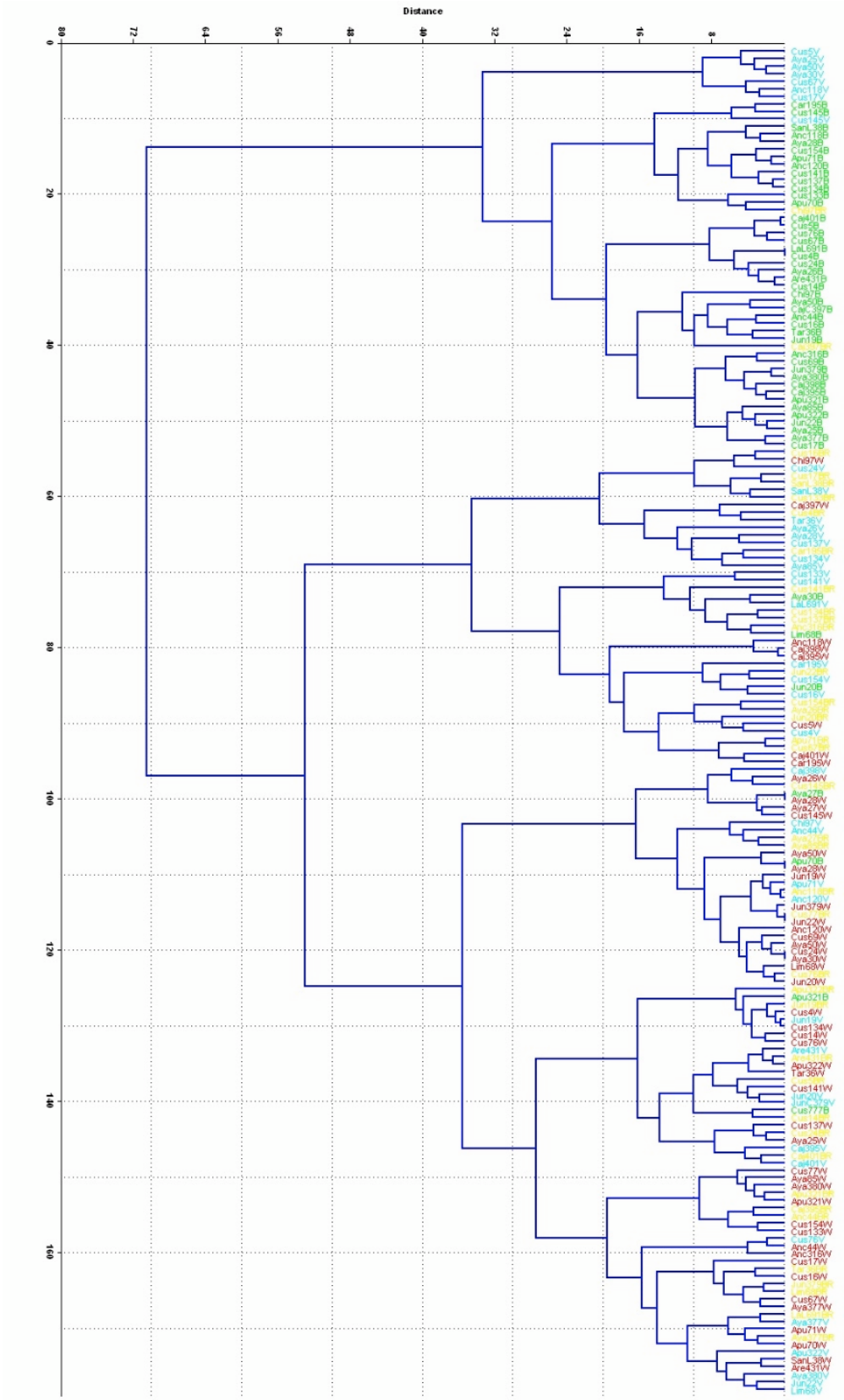

Distance
80
72
64
56
48
40
32
24
16
8
0
20
40
60
80
100
120
140
160

Figure 5. Cluster dendrogram showing the genetic relationships among all white-seeded (red), all black-seeded (green), and nine vitreous-seeded (blue) samples based on 96 SNP markers using PAST v2.05 software. Bootstrap values are given at each node.

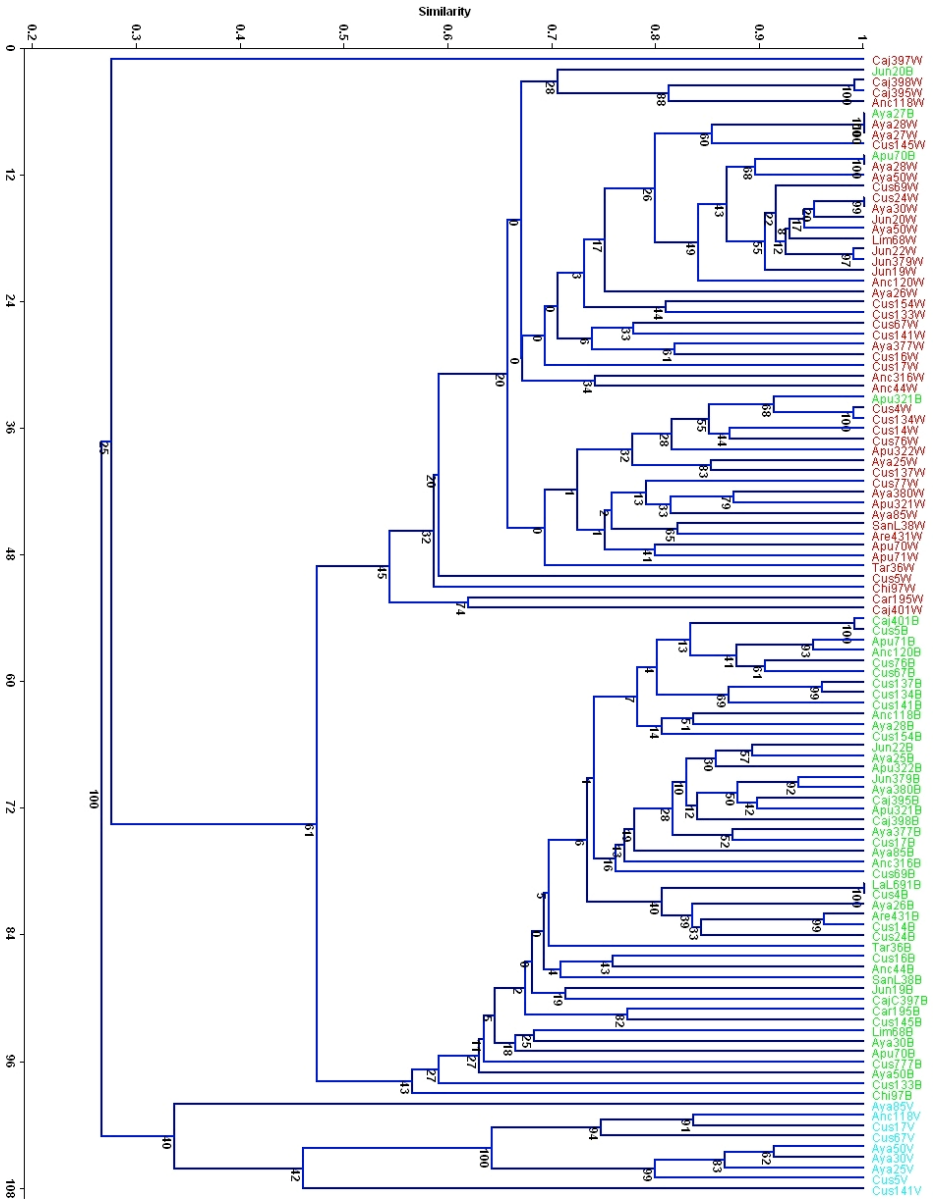
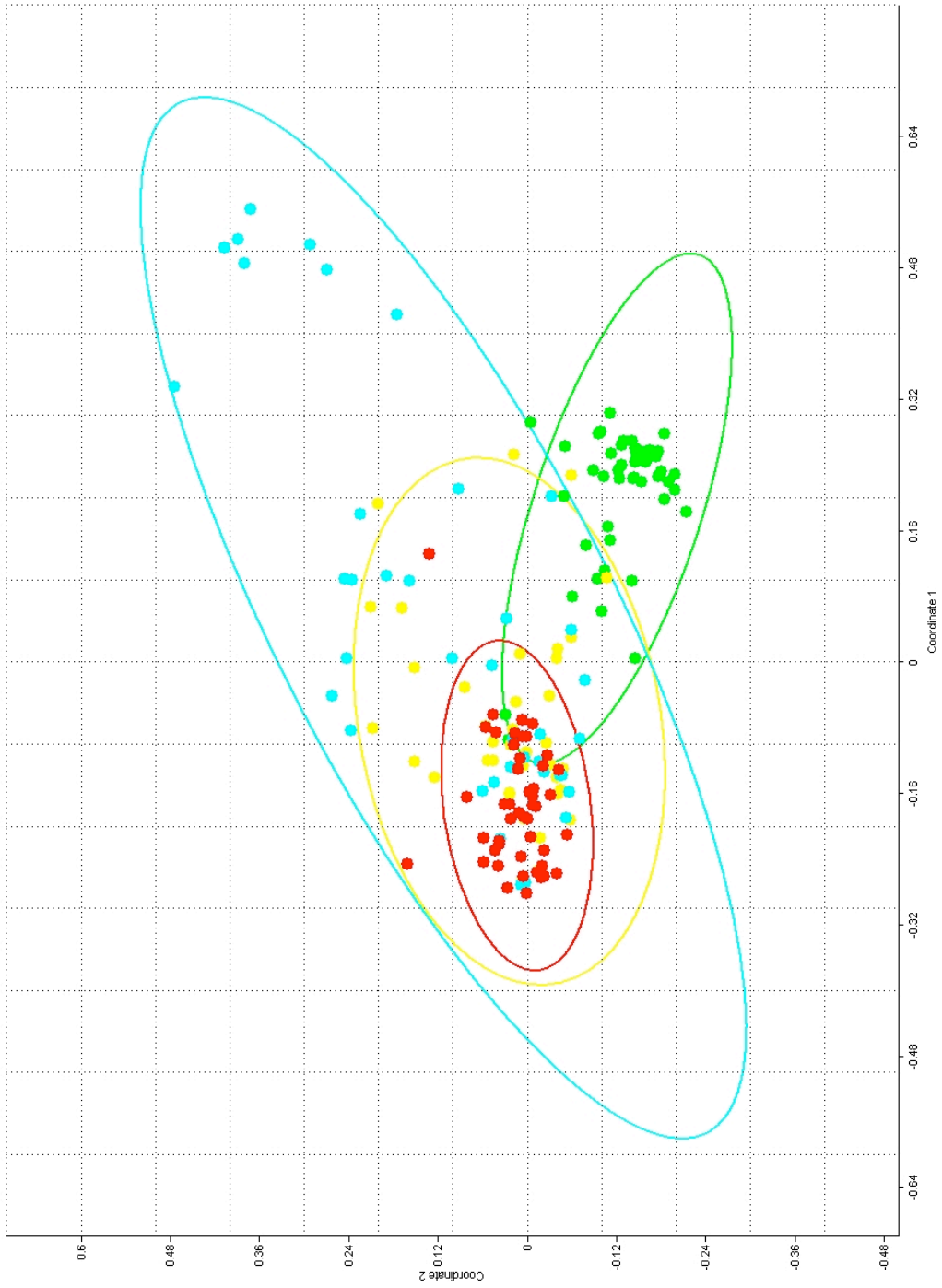
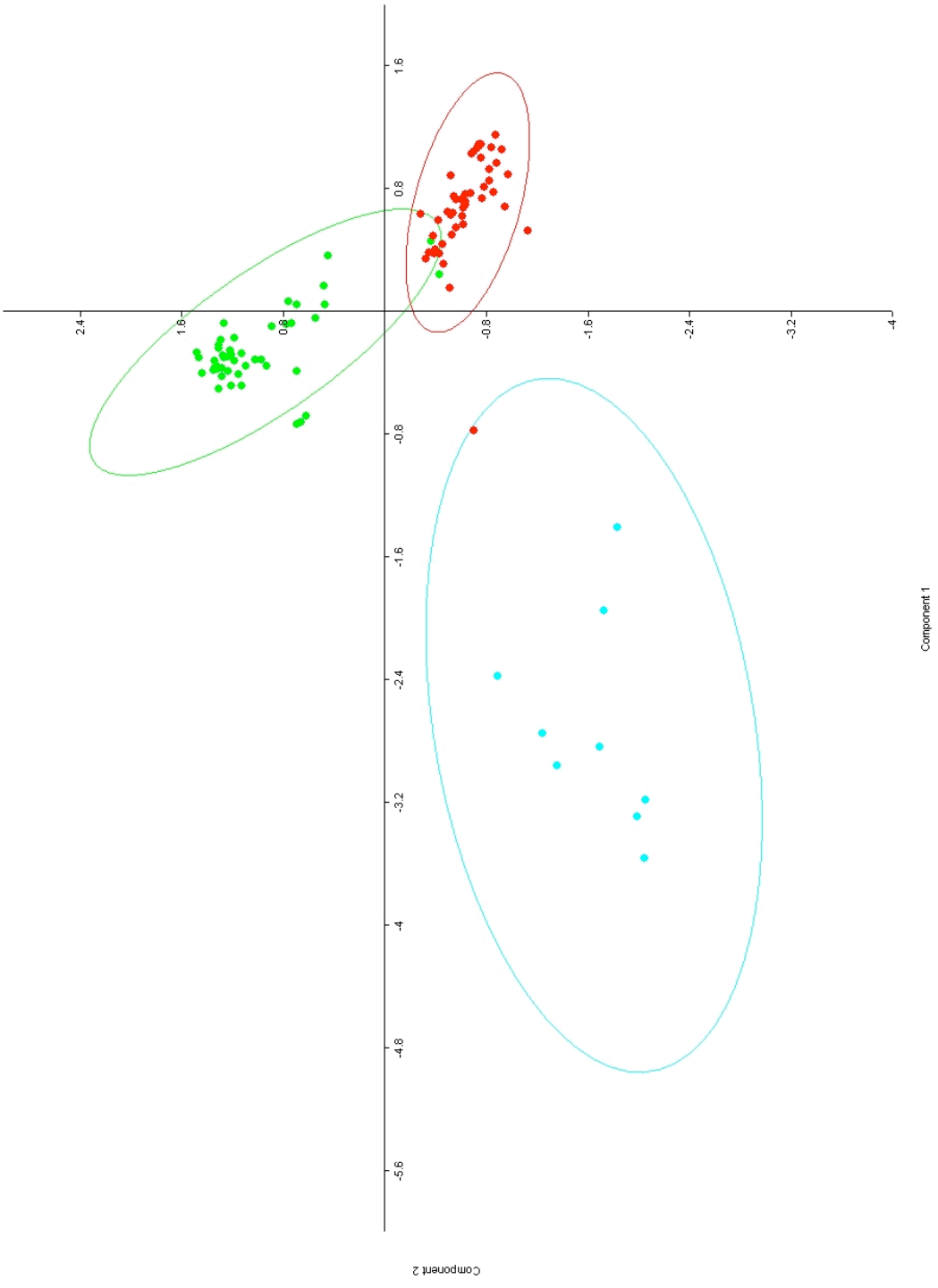


Figure 6 (Following pages). Principal coordinate analysis (PCO), bi-plot of the first two coordinates calculated from SNP-genotype data using PAST v2.05 software. Bootstrap values are given at each node. A) Relationships among all 178 amaranth accessions: white- (red), black- (green), brown- (yellow), and vitreous-seeded (blue). B) Relationships among all black-seeded, white-seeded, and 10 vitreous-seeded samples.

A)


0.64
0.48
0.32
0.16
0
-0.16
-0.32
-0.48
-0.64
Coordinate 1
0.6
0.48
0.36
0.24
0.12
0
-0.12
-0.24
-0.36
-0.48
Coordinate 2

B)


2.4
1.6
0.8
-0.8
-1.6
-2.4
-3.2
-4
1.6
0.8
-0.8
-1.6
-2.4
-3.2
-4
-4.8
-5.6
Component 1
Component 2