2007-08-01

Assessment of genetic diversity in the USDA and CIP-FAO international nursery collections of quinoa (Chenopodium quinoa Willd.) using microsatellite markers

S. A. Christensen

D. B. Pratt

C. Pratt

P. T. Nelson

M. R. Stevens

See next page for additional authors

Original Publication Citation

Plant Genetic Resources 5.2(Aug 27): 82-95.

ISSN 1479-2621
Plant Genetic Resources: Characterization and Utilization 5(2); 82–95
DOI: 10.1017/S1479262107672293

Assessment of genetic diversity in the USDA and CIP-FAO international nursery collections of quinoa (*Chenopodium quinoa* Willd.) using microsatellite markers

S. A. Christensen[1], D. B. Pratt[2], C. Pratt[1], P. T. Nelson[1], M. R. Stevens[1], E. N. Jellen[1], C. E. Coleman[1], D. J. Fairbanks[1], A. Bonifacio[3] and P. J. Maughan[1]*

[1]*Brigham Young University, Department of Plant & Animal Sciences, 275 WIDB, Provo, UT 84602, USA,*[2]*Stephen F. Austin State University, Department of Biology, Nacogdoches, TX 75961, USA and* [3]*Fundacion PROINPA, Casilla Postal 4285, Cochabamba, Bolivia*

Received 19 January 2006; Accepted 18 September 2006

Abstract

Quinoa (*Chenopodium quinoa* Willd.) is a staple food crop for millions of impoverished rural inhabitants of Andean South America where it has been cultivated for millennia. Interest in quinoa, due largely to its superior nutritional characteristics, is fuelling a growing export market and has led to an increased focus on genetic research and the development of quinoa breeding programmes throughout South America. The success of these breeding programmes will rely heavily on the development of core germplasm collections and germplasm conservation. We report the development of a set of fluorescence-tagged microsatellite molecular markers that can be used to characterize genetic diversity within quinoa germplasm and we use this set of 36 microsatellites markers to genetically characterize the diversity of 121 accessions of *C. quinoa* held in the USDA germplasm bank, 22 accessions from the CIP-FAO international nursery collection and eight accessions representing parents from genetic mapping populations. A total of 420 alleles were detected among the quinoa accessions with an average of 11 alleles detected per microsatellite locus. Genetic heterogeneity was observed in 32% of the quinoa accessions at a given locus and suggests that many of these accessions represent heterogeneous seed lots or landraces. Both unweighted pair-group method with arithmetic averages (UPGMA) and principle components analysis (PCA) analyses partitioned the quinoa accessions into two main clusters. The first major cluster consisted of accessions from the Andean highlands of Peru, Bolivia, Ecuador, Argentina and extreme northeastern Chile. The other main cluster contained accessions from both the lowlands of Chile and a set of USDA accessions with no known passport data, collected by Emigdio Ballón. Using the patterns of genetic diversity detected within the *C. quinoa* accessions we discuss hypotheses regarding quinoa's centre of diversity, including highland and lowland ecotype clustering patterns, origin of lowland varieties, origin of domestication, and diversity levels in the USDA and CIP-FAO collections.

Keywords: *Chenopodium quinoa*; genetic diversity; microsatellites; SSRs

Introduction

Quinoa (*Chenopodium quinoa* Willd.) is one of the most important food crops of the Andean region of South America (Cusack, 1984). It is an allotetraploid

* Corresponding author. E-mail: jeff_maughan@byu.edu

($2n = 4x = 36$) that shows disomic inheritance for most qualitative traits (Simmonds, 1971; Risi and Galwey, 1984; Ward, 2000). The seeds of quinoa are high in protein (~16%) and contain an excellent balance of the amino acids essential to the human diet (Ruales and Nair, 1992; Ruales *et al.*, 2002). Quinoa is uniquely adapted to the harsh conditions that characterize much of the Andean Altiplano (high plains), including high altitudes (3500–3850 m asl), arid (80–600 mm/year) and saline soils as well as frequent frost.

Heightened awareness of quinoa's role in food security issues in Andean South America, and its unique potential as a crop for marginal soils worldwide, and a growing health food export market for quinoa, has led to increased interest in quinoa and the establishment of several new breeding programmes throughout South America. Essential to these breeding programmes is the development of quinoa germplasm collections. We estimate that there are approximately 5000 diverse landraces of quinoa maintained in collections worldwide, including 2700 in Bolivia, 1029 in Peru (Ortiz *et al.*, 1998) and 232 in the USDA National Plant Germplasm System (NPGS). An undocumented number of accessions exist in Chile, Argentina, Ecuador and Colombia (Jacobsen, 2000; Rojas *et al.*, 2000; F. Fuentes, personal communication).

There are relatively few published efforts to characterize genetic diversity within quinoa germplasm collections. Wilson (1988b) used morphological data combined with data from six isozyme loci to characterize 98 diverse quinoa accessions. His research suggested that (i) the centre of quinoa diversity is likely located in the southern Andean highlands; (ii) accessions of quinoa cluster into two main groups: Andean highland and coastal lowland (hereafter denoted as highland and lowland); (iii) the lowland varieties originated from the southern highland populations; and (iv) the origin of domestication was likely located in the Peruvian/Bolivian region with subsequent dispersals to the north and south.

Ortiz *et al.* (1988) characterized the 1029 accessions of the Peruvian quinoa germplasm bank at the Universidad National del Altiplano-Puno (UNAP) using eight phenotypic descriptors. Their goal was to derive a representative core collection of 103 accessions using a sampling strategy based primarily on altitude of origin in a geographically stratified, non-redundant procedure. They reported that quantitative trait variation and altitude were highly correlated and demostrated the ability to represent a majority of the range of variation with a 10% core subset.

In another study, Rojas *et al.* (2000) analysed phenotypic variation within 1512 accessions of the Bolivian national quinoa collection using three multivariate procedures for 17 traits (15 quantitative, two qualitative). Their study included genotypes collected in Chile, Argentina and Peru, as well as native Bolivian materials. Using principal component, cluster and discriminate function analyses, they classified quinoa germplasm into seven distinct groups, including five within the Altiplano and two from lower altitude valleys of the eastern Andean mountain range. Their characters did not discriminate for the Chilean lowland germplasm, which grouped with one of their Altiplano clusters.

The ability to identify genetic variation is indispensable to effective germplasm management and conservation (Roa, 2004). Molecular markers provide unique and effective tools for evaluating and characterizing plant genetic diversity in a way that is unaffected by the environment (Gupta and Varshney, 2000). Indeed, molecular marker information has proven valuable in several aspects of germplasm management, including (i) identification of genetic identities within collections, (ii) development of genetically diverse core collections (a genetically representative subset of the reserve collection), (iii) monitoring of natural and artificial changes in the genetic collections (e.g. contaminants, hybridizations or mixtures) and (iv) identification of phylogenetic relationships within the germplasm collection and with related weedy species (Hokanson *et al.*, 1998; Dean *et al.*, 1999; McGregor *et al.*, 2002). Among the various molecular markers, microsatellites (also known as simple sequence repeats or SSRs), have emerged as the genetic marker system of choice for plant genetic resource analysis, owing to their ease of use, highly polymorphic nature and abundance throughout plant genomes (Maughan *et al.*, 1996; Mace and Godwin, 2002).

The conclusions derived from Wilson (1988a, b) and the development of the first set of quinoa microsatellite markers (Mason *et al.*, 2005) provide a foundation for an in-depth molecular survey of genetic diversity in this species. We report on the use of 36 fluorescence-tagged SSRs to genetically characterize 153 quinoa accessions from the USDA-NPGS and CIP-FAO collections. Using the patterns of genetic diversity detected within the *C. quinoa* accessions we discuss Wilson's (1988a, b) hypotheses regarding quinoa's centre of diversity, including highland and lowland clustering patterns, origin of lowland varieties and the origin of domestication.

Materials and methods

Plant material and DNA isolation

The *C. quinoa* plant material utilized in this investigation were derived from several sources, namely 121 accessions from the USDA North Central Regional Plant Introduction

Station of the US National Plant Germplasm System (Ames, IA; David Brenner), 22 accessions that make up the publicly available international nursery collection developed by the Centro Internacional de la Papa (CIP; Mujica *et al.*, 1998; Angel Mujica, National University of the Altiplano, Puno, Peru) and eight accessions used as parents in several genetic mapping populations provided by Alejandro Bonifacio (The Foundation for the Promotion and Investigation of Andean Products; PROINPA) (Table 1). We note that the USDA *Chenopodium* collection consists of 41 accessions included in this analysis that were collected by Emigdio Ballón (Traditional Native American Farmers Association, NM, USA) for which passport (origin) data are unavailable. Hereafter we refer to this specific portion of the USDA collection as the Ballón collection. Personal conversations with Emigdio Ballón suggest that this material was increased in New Mexico prior to its donation to the USDA collection on 15 May 1990.

Additionally, one accession of *C. giganteum* D. Don (Ames 19046; a related chenopod crop species) and one accession of *C. album* L. (PI 433378; lambsquarters or fat hen) were also provided by David Brenner (USDA-NPGS, Ames, IA) and included in our analysis for comparative purposes.

All plant material was grown at 25°C with a 12 h photoperiod in greenhouses at Brigham Young University, Provo, UT, USA. Equal amounts of leaf material were harvested from three individuals and bulked for DNA isolation using a freeze-dried leaf tissue protocol described by Sambrook *et al.* (1989) with modifications described by Todd and Vodkin (1996).

Fluorescent SSR analysis

In total, 36 highly reproducible microsatellite primer pairs that produced a simple banding pattern (i.e. all markers amplified a single polymorphic band with the exception of one marker (AAT71) that amplified a second monomorphic band) in prescreen tests and exhibited high heterozygosity values (range 0.33–0.85; mean = 0.59; Mason *et al.*, 2005) were selected for genotyping. Thirty-two of the microsatellite loci were previously described by Mason *et al.* (2005), while the remaining four were developed by P. J. Maughan *et al.* (unpublished data). GenBank accession numbers for all loci are reported in Table 2.

Fluorescent labels were added to either the forward or reverse primer with the strongest 3′ end (i.e. the end with the fewest mononucleotide repeats; Table 2). Each fluorescently labelled primer was synthesized by Applied Biosystems (Foster City, CA, USA) with one of the following fluorophores: 6-FAM, TET, HEX, VIC or NED (Table 2). Since all primer pairs for each of the microsatellite loci were originally designed to amplify PCR fragments in a narrow base pair range (150–250 bp) to allow maximal size separation (Mason *et al.*, 2005), allele size multiplexing was not attempted in this experiment. To avoid cross-amplification among primer pairs, each microsatellite was amplified separately in uniplex PCR reactions. PCR reactions consisted of 60 ng of quinoa genomic DNA, 0.2 mM dNTPs, 2.5 mM $MgCl_2$, 1 × PCR buffer, 1.0 μM primer and 1 U of *Taq* polymerase in a reaction volume of 15 μl. The amplification profile was as follows: 18 cycles of 94°C for 60 s, 67°C for 30 s, 72°C for 60 s, followed by 20 cycles using 58°C as annealing temperature, followed by a final 10 min extension step at 72°C. After amplification, PCR products labelled with three different fluorophores were pooled by combining 2 μl aliquots from each and diluting in 64 μl of HPLC-grade H_2O. A 2 μl aliquot of the pooled dilution was then dried down and submitted for size analysis at the BYU DNA sequencing centre (Provo, UT, USA). A size standard (GS-500) containing DNA fragments ranging in size from 35 to 500 bp labelled with the fluorescent dye ROX was added to each pooled sample. Amplified fragments were separated electrophoretically using an ABI 3100 Genetic Analyzer (Applied Biosystems).

Data analysis

Fluorescent peaks were analysed with the GENESCAN prism software version 3.7. Microsatellite alleles sizes were estimated with the GENOTYPER software version 2.1 (Applied Biosystems) and scored using the local southern base calling method for allele sizing (Elder and Southern, 1987). The microsatellite data set was converted to binary digits for presence (1) or absence (0) of specific alleles. Loci and accessions for which 25% or greater of the data were missing were removed from the data set. Various diversity measures were calculated from the resulting data, including mean number of alleles per locus, total number of alleles and number of unique alleles. Heterozygosity values for each marker locus were determined following published procedures (Mason *et al.*, 2005).

The data were subsequently analysed using the SIMQUAL subprogram of NTSYS-PC (Rohlf, 2000) to generate similarity matrices. These similarity matrices were subsequently analysed phenetically using both clustering and ordination algorithms in NTSYS-pc (Rohlf, 2000). An unweighted pair-group method with arithmetic averages (UPGMA) dendrogram was constructed using a double-centred Jaccard's similarity matrix and rooted using *C. album* and *C. giganteum* as outgroups. Principle components analysis (PCA) was performed using a variance–covariance matrix of the data.

Table 1. Passport data on *Chenopodium* accessions used for microsatellite analysis

Sample ID/name	Accession no.	Passport/origin[a]	Source[b]
Jujuy	N/A	Argentina, Jujuy	CIP-FAO
Kamira	N/A	Bolivia	CIP-FAO
Ratuqui	N/A	Bolivia	CIP-FAO
Real	N/A	Bolivia, Oruro	CIP-FAO
02-Embrapa	N/A	Brazil	CIP-FAO
Ollague	N/A	Chile	CIP-FAO
Narino	N/A	Columbia	CIP-FAO
E-DK-4	N/A	Denmark—Peruvian origin	CIP-FAO
G-205-95	N/A	Denmark—Peruvian origin	CIP-FAO
ECU-420	N/A	Ecuador	CIP-FAO
Ingapirca	N/A	Ecuador	CIP-FAO
RU-2	N/A	England—Chilean origin	CIP-FAO
RU-5	N/A	England—Chilean origin	CIP-FAO
NL-6	N/A	Holland—Chilean origin	CIP-FAO
CICA-127	N/A	Peru, Cusco	CIP-FAO
CICA-17	N/A	Peru, Cusco	CIP-FAO
03-21-072RM	N/A	Peru, Puno	CIP-FAO
03-21-079BB	N/A	Peru, Puno	CIP-FAO
Huariponcho	N/A	Peru, Puno	CIP-FAO
Illpa	N/A	Peru, Puno	CIP-FAO
Kancolla	N/A	Peru, Puno	CIP-FAO
Salcedo	N/A	Peru, Puno	CIP-FAO
Maniquena	N/A	Bolivia	PROINPA
Mocko	N/A	Bolivia	PROINPA
Chucapaca	N/A	Bolivia	PROINPA
L-P	N/A	Bolivia	PROINPA
Surumi	N/A	Bolivia	PROINPA
KU-2	N/A	Chile	PROINPA
G-205-95DK	N/A	Denmark—Chilean origin	PROINPA
0654	N/A	Peru, Puno	PROINPA
A1	PI 614881	Argentina, Jujuy	USDA-NPGS
A2	PI 614883	Argentina, Jujuy	USDA-NPGS
A3	PI 614884	Argentina, Jujuy	USDA-NPGS
A4	PI 587173	Argentina, Jujuy	USDA-NPGS
B2	PI 614902	Bolivia, Oruro	USDA-NPGS
B3	PI 614904	Bolivia, Oruro	USDA-NPGS
B4	PI 614905	Bolivia, Oruro	USDA-NPGS
B6	PI 614907	Bolivia, Oruro	USDA-NPGS
B8	PI 614909	Bolivia, Oruro	USDA-NPGS
B9	PI 614910	Bolivia, Oruro	USDA-NPGS
B10	PI 614911	Bolivia, Oruro	USDA-NPGS
B11	PI 614912	Bolivia, Oruro	USDA-NPGS
B12	PI 614913	Bolivia, Oruro	USDA-NPGS
B13	PI 614915	Bolivia, Oruro	USDA-NPGS
B14	PI 614916	Bolivia, Oruro	USDA-NPGS
B15	PI 614919	Bolivia, Oruro	USDA-NPGS
B16	PI 614920	Bolivia, Oruro	USDA-NPGS
B23	PI 614927	Bolivia, La Paz	USDA-NPGS
B24	PI 614928	Bolivia, La Paz	USDA-NPGS
B25	PI 614929	Bolivia, La Paz	USDA-NPGS
B27	PI 614931	Bolivia, Oruro	USDA-NPGS
B28	PI 614932	Bolivia, Oruro	USDA-NPGS
B29	PI 614933	Bolivia, Oruro	USDA-NPGS
B30	PI 614934	Bolivia, Oruro	USDA-NPGS
B31	PI 614935	Bolivia, Oruro	USDA-NPGS
B32	PI 614936	Bolivia, Oruro	USDA-NPGS
B33	PI 614937	Bolivia, Oruro	USDA-NPGS
B34	PI 614938	Bolivia, Oruro	USDA-NPGS
B35	PI 478415	Bolivia, La Paz	USDA-NPGS
B36	PI 478408	Bolivia, La Paz	USDA-NPGS

Table 1. *Continued*

Sample ID/name	Accession no.	Passport/origin[a]	Source[b]
B36	PI 478418	Bolivia, Potosi	USDA-NPGS
B37	PI 478410	Bolivia, La Paz	USDA-NPGS
B38	PI 478414	Bolivia, La Paz	USDA-NPGS
B39	PI 614002	Bolivia, Cochabamba	USDA-NPGS
B40	Ames 13215	Bolivia, La Paz	USDA-NPGS
B42	Ames 13217	Bolivia, La Paz	USDA-NPGS
B43	Ames 13218	Bolivia, La Paz	USDA-NPGS
B44	Ames 13219	Bolivia, La Paz	USDA-NPGS
C1	Ames 22153	Chile, Pichilemu	USDA-NPGS
C2	Ames 22154	Chile, Cajon	USDA-NPGS
C3	Ames 22155	Chile, Pichaman	USDA-NPGS
C4	Ames 22156	Chile, Cajon	USDA-NPGS
C5	Ames 22157	Chile, Lo Valdivia	USDA-NPGS
C6	Ames 22158	Chile, Llico	USDA-NPGS
C7	Ames 22159	Chile, Bucalemu	USDA-NPGS
C8	Ames 22160	Chile, Iloca	USDA-NPGS
C9	Ames 22161	Chile, Llico	USDA-NPGS
C10	PI 614880	Chile, Los Lagos	USDA-NPGS
C11	PI 614882	Chile, La Araucania	USDA-NPGS
C12	PI 614885	Chile, Bio-Bio	USDA-NPGS
C13	PI 614886	Chile, Maule	USDA-NPGS
C14	PI 614887	Chile, Bio-Bio	USDA-NPGS
C15	PI 614888	Chile, Bio-Bio	USDA-NPGS
C16	PI 614889	Chile, Bio-Bio	USDA-NPGS
C17	PI 433232	Chile, Groben	USDA-NPGS
C18	PI 584524	Chile, Chillan	USDA-NPGS
E1	Ames 13228	Equador, Otavalo	USDA-NPGS
P1	PI 510532	Peru, Puno	USDA-NPGS
P2	PI 510533	Peru, Puno	USDA-NPGS
P3	PI 510536	Peru, Puno	USDA-NPGS
P4	PI 510537	Peru, Puno	USDA-NPGS
P5	PI 510543	Peru, Puno	USDA-NPGS
P6	PI 510547	Peru, Puno	USDA-NPGS
P8	PI 510551	Peru, Puno	USDA-NPGS
P9	PI 596498	Peru, Cusco	USDA-NPGS
P10	PI 510542	Peru, Puno	USDA-NPGS
P11	PI 510540	Peru, Puno	USDA-NPGS
P12	PI 510550	Peru, Puno	USDA-NPGS
P13	PI 510545	Peru, Puno	USDA-NPGS
P14	PI 510548	Peru, Puno	USDA-NPGS
P15	Ames 26191	Peru, Puno	USDA-NPGS
P16	PI 510546	Peru, Puno	USDA-NPGS
CO1	PI 596293*	USA, Colorado	USDA-NPGS
MD1	NSL 86628*	USA, Maryland	USDA-NPGS
N1	NSL 91567*	USA, New Mexico	USDA-NPGS
N2	Ames 13719*	USA, New Mexico	USDA-NPGS
N3	Ames 13721*	USA, New Mexico	USDA-NPGS
N4	Ames 13722*	USA, New Mexico	USDA-NPGS
N5	Ames 13723*	USA, New Mexico	USDA-NPGS
N6	Ames 13724*	USA, New Mexico	USDA-NPGS
N7	Ames 13726*	USA, New Mexico	USDA-NPGS
N8	Ames 13727*	USA, New Mexico	USDA-NPGS
N9	Ames 13728*	USA, New Mexico	USDA-NPGS
N10	Ames 13729*	USA, New Mexico	USDA-NPGS
N11	Ames 13730*	USA, New Mexico	USDA-NPGS
N12	Ames 13731*	USA, New Mexico	USDA-NPGS
N14	Ames 13733*	USA, New Mexico	USDA-NPGS
N15	Ames 13734*	USA, New Mexico	USDA-NPGS
N16	Ames 13735*	USA, New Mexico	USDA-NPGS
N17	Ames 13736*	USA, New Mexico	USDA-NPGS
N18	Ames 13737*	USA, New Mexico	USDA-NPGS

Table 1. *Continued*

Sample ID/name	Accession no.	Passport/origin[a]	Source[b]
N19	Ames 13738*	USA, New Mexico	USDA-NPGS
N20	Ames 13739*	USA, New Mexico	USDA-NPGS
N21	Ames 13740*	USA, New Mexico	USDA-NPGS
N22	Ames 13741*	USA, New Mexico	USDA-NPGS
N23	Ames 13742*	USA, New Mexico	USDA-NPGS
N24	Ames 13743*	USA, New Mexico	USDA-NPGS
N25	Ames 13744*	USA, New Mexico	USDA-NPGS
N26	Ames 13745*	USA, New Mexico	USDA-NPGS
N27	Ames 13746*	USA, New Mexico	USDA-NPGS
N28	Ames 13747*	USA, New Mexico	USDA-NPGS
N29	Ames 13748*	USA, New Mexico	USDA-NPGS
N30	Ames 13749*	USA, New Mexico	USDA-NPGS
N31	Ames 13750*	USA, New Mexico	USDA-NPGS
N32	Ames 13751*	USA, New Mexico	USDA-NPGS
N34	Ames 13753*	USA, New Mexico	USDA-NPGS
N35	Ames 13754*	USA, New Mexico	USDA-NPGS
N36	Ames 13755*	USA, New Mexico	USDA-NPGS
N37	Ames 13756*	USA, New Mexico	USDA-NPGS
N38	Ames 13757*	USA, New Mexico	USDA-NPGS
N39	Ames 13758*	USA, New Mexico	USDA-NPGS
N40	Ames 13759*	USA, New Mexico	USDA-NPGS
N41	Ames 13760*	USA, New Mexico	USDA-NPGS
N42	Ames 13761*	USA, New Mexico	USDA-NPGS
N43	Ames 13762*	USA, New Mexico	USDA-NPGS
N44	Ames 13725*	USA, New Mexico	USDA-NPGS
N45	Ames 13720*	USA, New Mexico	USDA-NPGS
SC1	NSL 86649*	USA, South Carolina	USDA-NPGS
TX1	Ames 19047*	USA, Texas	USDA-NPGS
WA1	NSL 92331*	USA, Washington	USDA-NPGS
Cal	PI 433378	USA, Iowa, *C. album*	USDA-NPGS
M1	PI 476820	Mexico, *C. berlandieri*	USDA-NPGS
Cgig	Ames 19046	India, *C. giganteum*	USDA-NPGS

[a] Source of materials was determined from the USDA Germplasm Resources Information Network (GRIN). Accessions lacking specific passport data are denoted with an asterisk and were assigned the US state from which they were donated as their source ID. Accessions donated to the USDA by Emigdio Ballón (Ballón collection) were assigned the source ID 'USA, New Mexico'.
[b] USDA-NPGS, USDA North Central Regional Plant Introduction Station of the US National Plant Germplasm System (Ames, IA); CIP-FAO, Centro Internacional de la Papa; PROINPA, The Foundation for the Promotion and Investigation of Andean Products.

Results

Microsatellite diversity

A total of 420 alleles were identified by the 36 microsatellite loci among the 152 quinoa accessions analysed (control species were not included in the analysis). The number of alleles per locus ranged from three (QAAT10) to 27 (QAAT50 and QAAT76) with an average of 11.7 alleles per locus (Table 2). The tri-nucleotide repeat primers amplified a significantly greater number of alleles (*t*-test; $P \leq 0.05$) per locus (average 15.2; $n = 15$) than did the di-nucleotide repeats (average 9.1; $n = 21$).

Heterozygosity (H) values for the microsatellite loci ranged from 0.45 to 0.94 with a mean value of 0.75. As expected, the mean H value in this experiment is larger than the average H value (0.57) reported by Mason *et al.* (2005). The increase in the average heterozygosity value in this study is likely due to: (i) the inclusion of only highly informative markers; (ii) better resolution of marker alleles (single base pair resolution) due to the use of capillary electrophoresis; and (iii) the screening of a much larger and likely more diverse panel of genotypes. All of these factors should lead to an increase in number of alleles observed and consequently increased H values. According to the definition of Ott (1992), markers are considered polymorphic if $H \geq 0.1$ and highly polymorphic if $H \geq 0.7$, thus all of the

Table 2. Number of alleles, number of rare alleles, heterozygosity value, primer sequences and GenBank accession number for each microsatellite marker used to assess genetic variation in 150 accessions of quinoa

Primer	Fluorescent label	No. of alleles	No. of rare alleles	*H* value	Allelic size range (bp)	Forward primer (5′–3′)	Reverse primer (5′–3′)	GenBank accession no.
KGA03*	HEX	21	15	0.85	140–182	HEX-attgccgacaatgaacgaat	***gcttct***atgtaaatggcatgtcccaac	DQ462129
KGA16*	6FAM	9	3	0.76	155–196	6FAM-ccctgcttaatctccgtgaa	***gcttct***ccgaaccaagactacgaaaca	DQ462130
KGA20*	NED	17	7	0.87	155–185	***gcttct***tcacctacctcggtaaaggaaa	NED-ggagcagatgatgaacatgg	DQ462131
KGA27*	NED	16	8	0.90	126–158	NED-ttgtacagaggaagtggcaaga	***gcttct***catcttacagctctggctttcc	DQ462132
QAAT06*	6FAM	12	5	0.81	193–226	***gcttct***cacaaacaataattcaaccgaaga	6FAM-cgctgacgcttaacattcg	DQ462133
QAAT10	6FAM	3	0	0.45	148–161	***gcttct***tgaaatcgaaacggaggaag	6FAM-gggataggactgatgccaga	DQ462134
QAAT22*	6FAM	26	15	0.94	153–235	***gcttct***tggtcgatatagatgaaccaaa	6FAM-ggagcccagattgtatctca	DQ462135
QAAT24	6FAM	20	13	0.85	201–257	***gcttct***accataacagcacccacctt	6FAM-agggatcaatcttgttcattca	DQ462136
QAAT50*	VIC	27	19	0.89	158–246	VIC-ggcacgtgctgctactcata	***gcttct***atggcgaatggttaatttgc	DQ462137
QAAT69*	NED	15	7	0.88	193–266	NED-gtttcctttgaggcttggac	***gcttct***ggatttgtacgaatagttgggatt	DQ462138
QAAT70*	NED	17	6	0.91	158–208	NED-tgaacaggatcgtcatagtcaa	***gcttct***cgttcatcatctgacccaat	DQ462139
QAAT71*	TET	25	14	0.92	122–200	TET-catcacccgctgaatagacac	***gcttct***taccctaatgccacgattcc	DQ462140
QAAT74*	6FAM	16	7	0.87	169–224	***gcttct***atggaacacccatccgataa	6FAM-atgcctatcctcatcctcca	DQ462141
QAAT76*	HEX	27	17	0.92	145–227	HEX-gcttcatgtgttataaaatgccaat	***gcttct***tctcggcttcccactaatttt	DQ462142
QAAT78*	HEX	13	4	0.85	183–226	HEX-agcgaaggaaatttggaact	***gcttct***taacgatacgctccaaggaa	DQ462143
QATG19	6FAM	7	2	0.74	175–193	***gcttct***ccaaacaaagacaataaggaaacc	6FAM-cgaggttgaaggagattcca	DQ462144
QATG52	6FAM	11	4	0.73	200–253	6FAM-tgcagtgaagtgaacaagagaga	***gcttct***gcaccaaatcacaccattca	DQ462145
QATG60	6FAM	4	1	0.59	174–189	***gcttct***tggtgaaagcagttatgaagaca	6FAM-tttcagctcaacctagaacacg	DQ462146
QATG86	VIC	5	1	0.61	189–200	VIC-aatcgcagcctaaactgagc	***gcttct***agttccatttcgaccatatgataa	DQ462147
QCA06	6FAM	6	3	0.56	109–127	6FAM-gctctattaaggaaatgaggttaca	***gcttct***gccattcaattcagcaaagg	DQ462148
QCA11*	VIC	4	0	0.60	199–206	VIC-caggagcctcaagaatgaca	***gcttct***tttaaattcaagcctaccaccttt	DQ462149
QCA14	VIC	8	2	0.59	196–211	***gcttct***cctgagctgatttatcaaaggac	VIC-cctcttgcgagatttctgct	DQ462150
QCA19*	VIC	6	2	0.69	183–212	VIC-tttcatcactcgaccgtatagc	***gcttct***agggtgactgttacacccaaa	AY458236
QCA24*	VIC	7	2	0.75	235–254	VIC-agatgagcttgaatcattacatc	***gcttct***tacatactgtaaatcatgccaaa	DQ462151
QCA37*	VIC	9	1	0.77	186–206	***gcttct***ccgttcttccagaccaattc	VIC-tcatgagccacttcatacacg	AY458227
QCA38*	VIC	5	1	0.71	198–209	***gcttct***catttcccaaactgcatgaat	VIC-atgtgtgttgcgtgtgagtg	DQ462152
QCA48*	6FAM	6	2	0.65	246–258	***gcttct***acaatacatacataacccaatattcaa	6FAM-tggaaatgtcactatgattgga	AY458240
QCA57*	NED	8	2	0.78	168–189	***gcttct***tgcaaggaaaccatctttgg	NED-tgcctcacagtcacacctaca	AY458243
QCA58*	NED	8	2	0.80	170–193	NED-ctcgaccagcagggtctg	***gcttct***ctagctaggcgttgcctgac	AY458244
QCA65*	NED	9	3	0.79	168–206	***gcttct***ccatgcaagggaacatattg	NED-aagttcgttggcttgctgta	AY458246
QCA71*	6FAM	16	10	0.84	140–177	6FAM-aacaacgaaattacgagaatgtca	***gcttct***tctcacgagagtcttcccta	DQ462153
QCA88	NED	8	6	0.49	126–246	***gcttct***tctggctgcttccacctaat	NED-cagtcccggaatcgtaactc	DQ462154
QCA107*	NED	5	1	0.70	153–167	***gcttct***acaggctgtgggtccactt	NED-tcaagcaatactcaccttgtgg	DQ462155
QCA120	NED	6	0	0.69	171–198	***gcttct***gacgcacataacgtgtagaattg	NED-tcctcattcccttctccatc	DQ462156
QGA03*	NED	10	3	0.79	188–209	***gcttct***gaacctttaataggtctgtaccaaatc	NED-aagaaatgtcacaagcaagca	DQ462157
QGA17	NED	8	5	0.52	145–164	NED-ttacggttctcccggtctc	***gcttct***tgcaaacaagagaagcatgaag	DQ462158
Mean		12	5	0.75				

A total of 420 alleles were amplified across the *C. quinoa* genotypes ranging from three to 27 alleles per locus (mean = 11.7). Primers that differed for the most common allele between the lowland and highland groups are denoted with an asterisk. Bolded portions of the primer sequences represent the nucleotides (gcttct) added to the primer to facilitate PCR.

microsatellite markers utilized in this study were polymorphic and 25 (69%) were highly polymorphic.

The allelic size range variation (defined as base pairs from smallest to largest allele) at a microsatellite locus was highly correlated ($P \leq 0.01$) with the total number of alleles identified per locus ($r = 0.71$; Table 2). Since different microsatellite loci carrying the same fluorescent label can be simultaneously analysed in the same gel lane/capillary column if their allele size ranges do not overlap, the determination of the full allelic size range for these microsatellite markers represents an important step towards the future development of multiplex marker sets based on allele size ranges. Such multiplexed sets will dramatically reduce the cost associated with genotyping. Several researchers have reported the development of six- and nine-plex marker systems for a wide variety of crop species including soybean (Diwan and Cregan, 1997), sunflower (Tang *et al.*, 2003) and rape seed (Tommasini *et al.*, 2003). In the present study, the average allelic range across the microsatellites was 37 bp (Table 2).

Rare and high-frequency microsatellite alleles

Rare alleles, defined as those that were observed in less than 5% of the accessions were detected at all loci, except QAAT10, QCA11 and QCA120 (Table 2). QAAT50 displayed the highest number of alleles (27), and the highest percentage (70%) of rare alleles (19/27 alleles; Table 2). Not unexpectedly, and as seen by other researchers (Jain *et al.*, 2004), a positive significant correlation existed between the number of rare alleles and the number of alleles per locus ($r = 0.96$). Rare alleles are of particular value for variety identification via DNA fingerprinting.

High-frequency alleles (common alleles) were seen within the majority of the loci analysed. Indeed, on average 76 (50%) of the 152 accessions shared a common allele at any given locus. A negative correlation between the frequency of the most common allele at a locus and the number of alleles per locus was observed ($r = -0.78$). Within the group of 81 highland accessions the average frequency of shared alleles was 53%, whereas within the group of 71 lowland (including the Ballón collection) ecotypes the frequency of shared alleles rose to 61%, suggesting a lower level of genetic diversity within the coastal ecotypes. Moreover, a comparison of the most common allele observed in the highland and lowland ecotype groups indicated that the most common allele differed between the groups in 25 of the 36 microsatellite loci (microsatellite markers that differed for the most common alleles between the two groups are marked with an asterisk in Table 2). Based on the highest allele frequency alone, the two groups could be distinguished from each other at 69% of the microsatellite loci.

Heterogeneity

Genetic heterogeneity, detected as multiple alleles at a genetic locus within an individual accession (allele mixtures), was observed on average in 32% of the quinoa accessions per microsatellite locus. The presence of heterogeneity in the quinoa accessions was not unexpected and may be explained by low levels of residual heterozygosity at individual loci due to outcrossing during seed maintenance (outcrossing estimates range from 8 to 10%; A. Bonifacio, personal communication) or as a result of seed mixtures of contrasting homozygous (or heterozygous) lines inherent to the heterogeneous quinoa seed lots, which would be detected as a result of the leaf-bulking strategy utilized in the DNA preparation of each accession. Indeed, phenotypic inspection of field-grown seed lots revealed low levels of phenotypic diversity in many of the accessions utilized in these experiments, suggesting the presence of genetic heterogeneity within accessions and the need for future characterization of within-line genetic diversity. Loci that produced multiple alleles were scored as having all alleles present.

Phenetic analysis and genetic similarities

Jaccard's similarity coefficients for quinoa accessions ranged from 0.038 to 1, with the least genetic similarity detected between the Bolivian highland accession B15 (PI614002) and the lowland Chilean accession C15 (PI614888), whereas the greatest genetic similarity was observed between the genetically identical lowland accessions WA (NSL92331) and SC (NSL86649). Distinct genetic profiles were identified for all other accessions.

The UPGMA and Jaccard's similiarity coefficients were used to produce a dendrogram of quinoa accessions and the two control *Chenopodium* control species (Fig. 1). *C. album* and *C. giganteum* clustered at the root of the UPGMA dendrogram with the Mexican accession M1 (Fig. 1; see Discussion). This Mexican accession (M1) was extremely distinct from all other quinoa accessions as shown by it low pairwise similarity value (0.073) and was thus removed from subsequent analyses. The remaining 150 quinoa accessions, clustered into two large, but clearly distinct, clusters corresponding to the highland Andean and lowland coastal accessions (Fig. 1). The highland cluster consisted of accessions from Peru, Bolivia, Ecuador and Argentina and are broken down into two subgroups: northern highland and southern highland. All four of the Ecuadorian accessions clustered together with all but four (P11, P15, P6 and P17) of the Peruvian accessions in the northern highland subgroup, whereas all but three Bolivian accessions

(B37, B38 and B39) clustered with the majority of the Argentinian accessions in the southern highland subgroup (Fig. 1). One Argentinian accession (A2) clustered with the coastal lowland accessions. The Chilean accessions all clustered in the coastal lowland group with the exception of 'Ollague', a highland ecotype originating from northern Chilean Altiplano. Most of the USDA-Ballón collection clustered in the lowland coastal

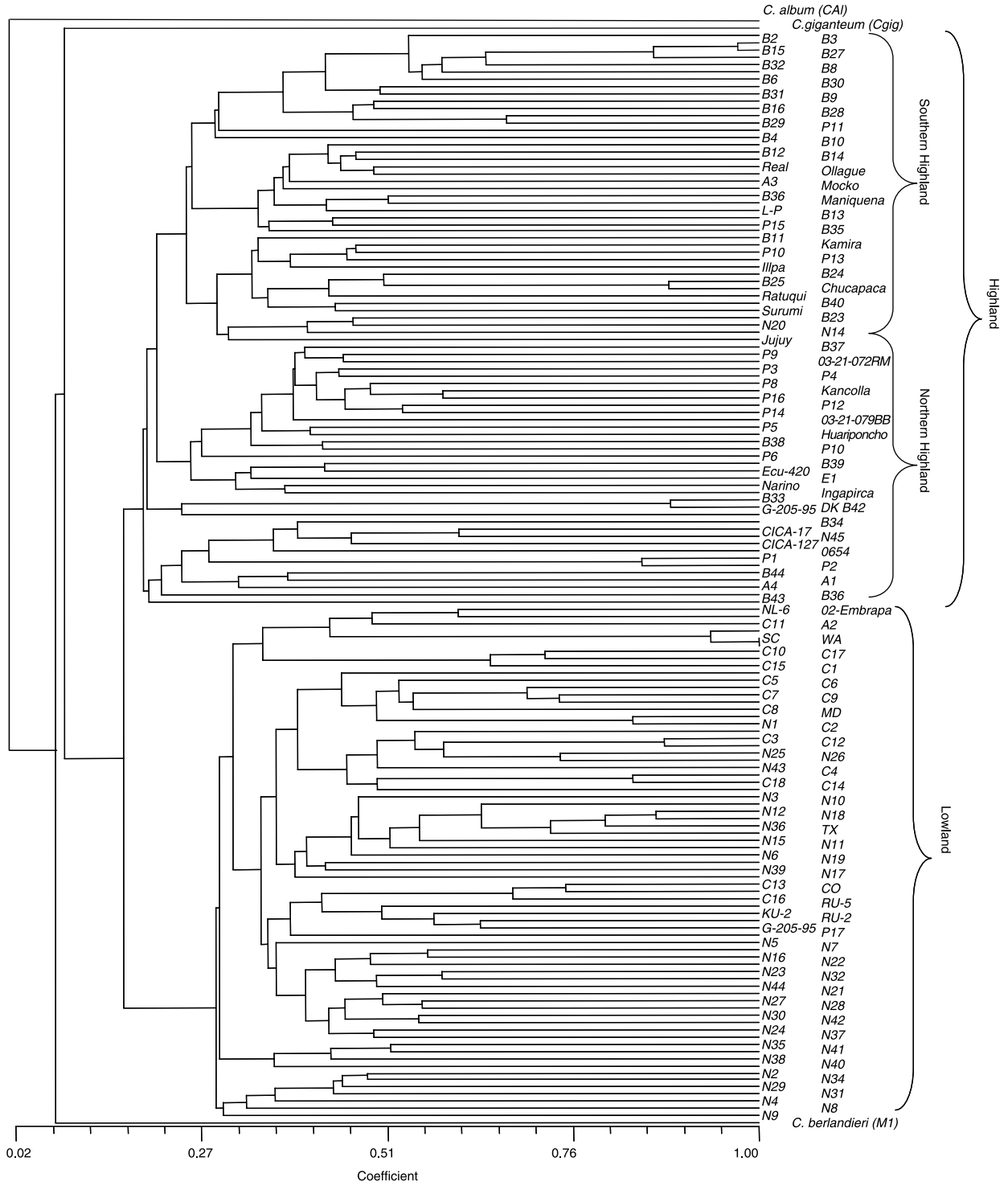

Fig. 1. Dendrogram showing the 150 accessions of *Chenopodium quinoa*, one accession of *C. berlandieri* Moq. ssp. *nuttaliae* (M1) and the two control species (*C. album*—Cal; *C. giganteum*—Cgig), from UPGMA cluster analysis using allelic data from 36 microsatellite markers. *C. quinoa* accessions are clustered into two main groups: highland (top) and lowland (bottom) ecotypes.

group, with the exception of N20 and N14 which grouped with the accessions in the southern highland subgroup (Table 1; Fig. 1).

PCA formed two well-separated groups (highland and lowland; Fig. 2). The plot of PCA Dim-1 against Dim-2 accounted for 21.8% (Dim-1 = 15.9%; Dim-2 = 5.9%) of the total variation observed in the 152 quinoa accessions analysed. The lowland cluster forms a single continuum along the Dim-1 axis. Within this continuum, the Chilean lowland and Ballón accessions segregate to the opposite extremes, with the Ballón accessions stretching towards the highland cluster (Fig. 2). No discrete breaks occur between the Chilean and Ballón accessions. The highland variation forms a single continuum, easily seen along the Dim-2 axis, with northern (Ecuadorian and Peruvian) and southern (Bolivian and Argentinian) accessions stretching to opposite extremes of the Dim-2 axis. Some Peruvian (northern highland) and Bolivian (southern highland) accessions intermingle within the centre of the highland cluster, preventing the separation of northern and southern highland accessions into discrete entities.

Pairwise comparisons of the similarity coefficient within the highland group (mean = 0.23) are significantly lower (*t*-test; $P \leq 0.01$) than comparisons within the lowland group (mean = 0.34), suggesting that the lowland group contains less genetic diversity. Table 3 presents the average similarity value between accessions based on country of origin (or collector in the case of the Ballón collection). The highest similarities between countries were observed between the Chilean genotypes and the Ballón collection (0.31). In contrast, the greatest dissimilarity between countries was observed between the Chilean and Ecuadorian genotypes, followed closely by the comparison of the Chilean and Peruvian genotypes (0.12 and 0.13, respectively). Interestingly, the genetic similarity comparisons between the Ballón collection and Peruvian and

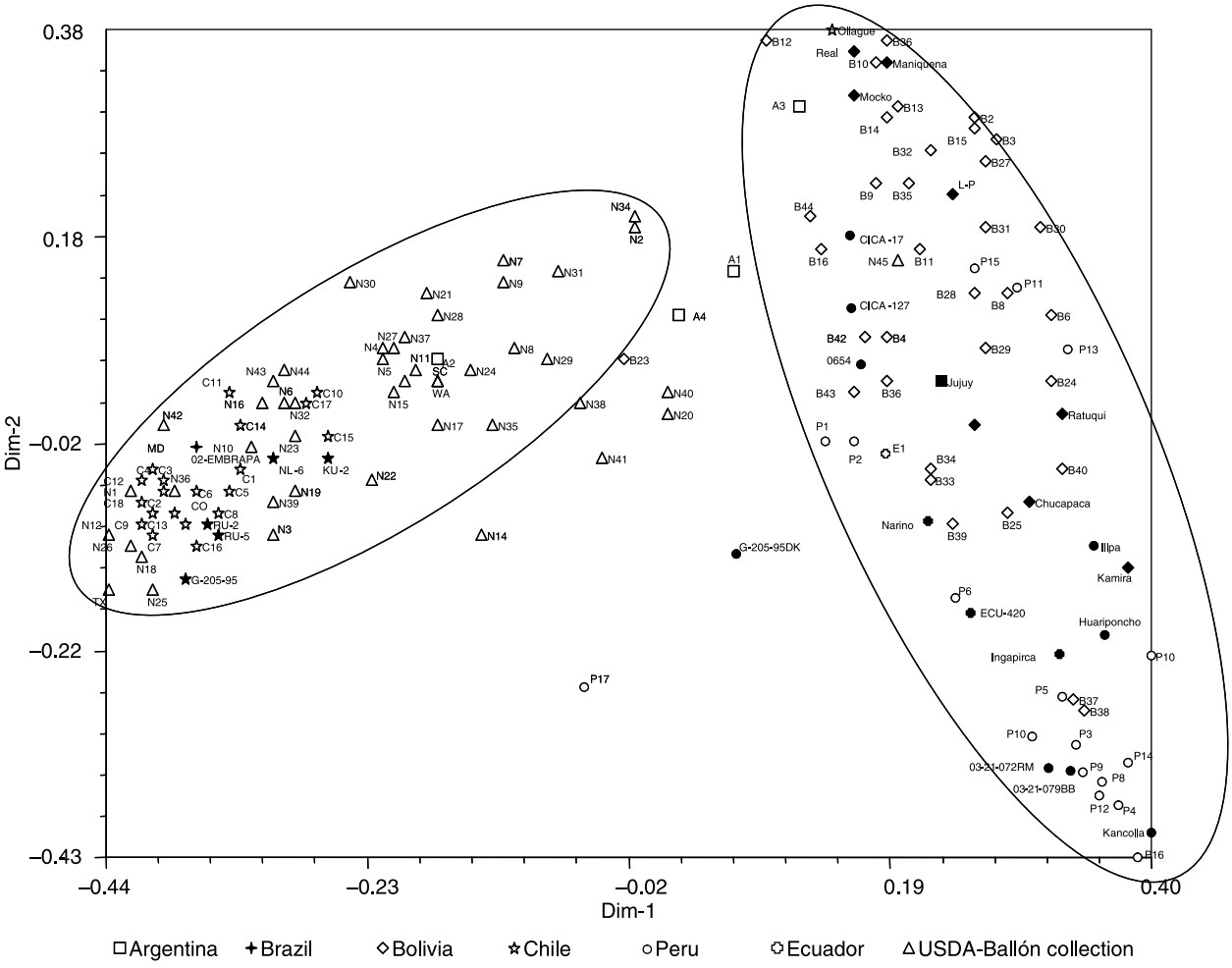

Fig. 2. The two-dimensional scaling resulting from the PCA analysis of 150 *Chenopodium quinoa* accessions using genetic diversity data from 36 microsatellite markers. Circled areas represent the major highland and lowland clusters identified in the dendrogram. Accessions are coded for their country of origin as described in the legend. Blackened symbols indicate accessions from CIP-FAO and PROINPA collections.

Table 3. Total alleles, mean number of alleles per region, unique alleles and pairwise similarity comparisons given for each country represented in the analysis (the single quinoa accession from Brazil (02-Embrapa) is not included in the analysis)

	Ecuador ($n = 4$)	Peru ($n = 28$)	Bolivia ($n = 42$)	Argentina ($n = 5$)	Chile ($n = 24$)	Ballón collection[a] ($n = 48$)
Allelic diversity						
Total alleles	131	271	290	162	198	241
Mean no. of alleles per locus	3.6	7.5	8.1	4.5	5.5	6.7
Unique alleles	5	13	17	10	9	4
Pairwise similarity comparisons among countries						
Ecuador	0.337					
Peru	0.243	0.274				
Bolivia	0.200	0.212	0.262			
Argentina	0.165	0.171	0.209	0.208		
Chile	0.125	0.133	0.144	0.199	0.350	
Ballón collection	0.159	0.167	0.185	0.209	0.309	0.333

[a] The Ballón collection lack specific passport information regarding country of origin. According to the phenetic analysis they are more similar to the lowland accessions originating from Chile.

Bolivian genotypes were intermediate (0.19 and 0.17, respectively) as compared to the distances between the Chilean material and the Peruvian and Bolivian genotypes (0.13 and 0.14, respectively).

Discussion

Phenetic analyses of quinoa accessions show that the lowland and highland accessions segregate into two discrete clusters consistent with Wilson's (1988a) hypothesis. Additionally, the highland cluster segregates into northern (northern Peru/Ecuador) and southern (southern Peru/Bolivia/northern Argentina) highland subgroups (Fig. 2). Gandarillas (1979) and Wilson (1988b) suggested that the southern highlands region near Lake Titicaca represents the centre of diversity for quinoa based on the high levels of morphological and isozyme diversity observed in this region. Diversity measures based on the molecular data reported herein support this view since the regions to the north and south of this region show a decrease in genetic variation (Table 3). The reduction in genetic diversity levels in the northern highland and coastal lowland regions could be a result of founder effects associated with the dispersal of the crop from its centre of origin or alternatively as a response to selective adaptation in the more uniform ecological regions of the northern highlands and coastal lowlands. Given quinoa's ancient association with the Incan civilization any interpretation of quinoa's spread through South America needs to consider the broad implications of human intervention (Wilson, 1988a).

Two hypotheses (Gandarillas, 1979; Tapia, 1979; Wilson, 1988a) have been proposed to explain the origin of quinoa domestication: (i) quinoa originated in Colombia and spread southwards along the Andes; and (ii) quinoa originated in the southern highlands of Peru and Bolivia. As Gandarillas (1979) noted, the evidence in support of a northern origin are tenuous at best, as it appears to be mostly anecdotal and linguistic in inspiration (Tapia, 1979). Gandarillas (1979) noted a Peruvian–Bolivian centre of morphological diversity and subsequently proposed the Peruvian and Bolivian southern highlands as the centre of domestication. Our data support the Peruvian–Bolivian centre of diversity; however the lack of accessions from Ecuador and Colombia in the USDA collection precludes the elimination of a northern hypothesis of the origin of domestication of quinoa and further suggests that concentrated collections of quinoa accessions are needed from these countries.

The phenetic analysis showed an interdigitation of Peruvian and Bolivian accessions in both the UPGMA dendrogram and in the PCA analysis, which is likely due to the geographic proximity of some genotypes along the border of Bolivia and Peru on the southern Andean Altiplano. Ecuadorian accessions form a single group within the northern highland subgroup of the cluster analysis, although they are somewhat scattered within the northern subgroup (Fig. 2), suggesting that the Ecuadorian accessions were likely introduced from the Andean highland group followed by subsequent adaptation. Interestingly, the Argentinian accessions are scattered with no discernible pattern within the dendrogram (Fig. 1) or the PCA (Fig. 2), with one accession (A4) even grouping with coastal Chilean lowland accessions. Although our understanding of Argentinian germplasm is limited by the small sample size available to us in the USDA collection, it does appear that the Argentinian accessions included in this study potentially represent reintroductions of quinoa into Argentina from

both the southern highland and lowland genotypes. Future systematic and phyletic analysis of quinoa would undoubtedly benefit from the inclusion of a wider array of Argentinian accessions collected from indigenous populations.

Tapia *et al.* (1980) classified highland quinoa into four distinct ecotypes, Valle, Yungas, Altiplano and Salares, based on altitude, height of the plants, amount of branching, length of growing period and salt tolerance. Wilson (1988a) noted that these classifications were not well supported by isozyme data. Other than the tendency of the Salares types to cluster together along Dim-2, we could detect no correlation between these ecotypes and genetic diversity as revealed in our phenetic analyses. Furthermore, we observed no significant correlation ($P \leq 0.05$) of genetic diversity using PCA Dim-1 or Dim-2 with altitude, latitude or longitude (data not shown), although we note these correlations were hindered by a lack of reliable passport data for many of the accessions (Pratt and Clark, 2001).

Examination of the lowland cluster shows evidence of two subgroups, the first from lowland Chile and the second made up of populations collected by the Bolivian agronomist Emigdio Ballón. The Chilean subgroup is less diverse as compared to the Peruvian and Bolivian populations (Table 3). The lack of genetic diversity in the Chilean populations supports Wilson's (1988a, b) hypothesis of an ancient founder effect followed by long periods of drift in Chilean populations. *t*-Tests of Jaccard's similarity measures between Chile and Bolivia as well as Chile and Peru showed that the reduction of variation in Chile is statistically significant ($P \leq 0.001$) for both comparisons. Wilson (1988a) additionally hypothesized that Chilean populations originated from the southern Andes. This hypothesis is further supported by pairwise comparisons of the Jaccard's similarity values based on country of origin which indicated that Chilean populations are more similar to the southern highland genotypes from Argentina (0.199) and Bolivia (0.144) than they are to the northern highland genotypes from Peru (0.133) and Ecuador (0.125).

An anomaly in the data is the Mexican accession M1 (PI 476820, 'Santa Elena 7') which grouped by itself in the phenetic analysis and was genetically distinct from all other quinoa accessions, with a substantially reduced average genetic similarity value of 0.073 (7.3%) when compared to all other quinoa accessions. The USDA's Germplasm Resources Information Network (GRIN) lists M1 as *C. quinoa*, but also lists 'huazontle' as an additional name for this accession. Huazontle is the common name used to refer to *C. berlandieri* Moq. ssp. *nuttalliae*, a cultivated chenopod grown in central Mexico. Nelson (1968) refers specifically to 'Santa Elena #7' as *C. nuttalliae* Safford, a former classification for *C. berlandieri* ssp. *nuttalliae*. Huazontle and quinoa are both members of the series Foveosa (subsect Cellulata, sect Chenopodium), and the two cross readily. Aellen (1929) treated the two as conspecific. They are separated geographically and distinguished by slight differences in leaf shape and seed pigmentation (Simmonds, 1965; Nelson, 1968). M1 is from within the range of huazontle, and well outside the area where *C. quinoa* is commonly found. Hence, the data presented here suggest that M1 is misclassified as *C. quinoa* and should be reclassified as *C. berlandieri* Moq. ssp. *nuttaliae*.

USDA-Ballón collection subset

Many of the accessions available in the USDA collection were collected by Emigdio Ballón. Unfortunately, the GRIN database lacks passport information for the vast majority of these genotypes. The UPGMA dendrogram clustered the majority (>90%) of the Ballón genotypes with the Chilean lowland accessions (Fig. 1). The PCA analysis (Fig. 2) also places them within the Chilean lowland group, however the Ballón collection is oriented between the Chilean populations and the southern highland populations, indicating a potential affinity with southern highland accessions that is further supported by pairwise comparisons of Jaccard's similarity values (Table 3).

Shared allelic patterns between the Ballón, Chilean lowland and highland accessions are particularly interesting. Highland, Chilean and the Ballón accessions share a common predominant allele at 31% of the loci. However, at 44% of the loci, the Ballón and Chilean accessions shared a predominant allele while the highland accessions exhibited a different predominant allele. In contrast, the Ballón and highland accessions shared a predominant allele at only 7% of the marker loci. At the remaining loci (18%) the Ballón accessions have a predominant allele that differs from the predominant allele identified in the highland and Chilean lowland groups. The Ballón collections also show higher diversity measures and they harbour a number of alleles not found in the lowland accessions but present in highland accessions. Thus the Ballón collection appears to be coastal Chilean in terms of predominant alleles, but further enriched by highland alleles.

One intriguing hypothesis that may explain the intermediate position of the Ballón collection is that it may represent a distinct geographic ecotype. One potential geographic area to which this collection could belong is the Yungas region (Tapia *et al.*, 1980), which stretches along the eastern slopes of the Bolivian Andes. Evidence in support of this hypothesis includes: (i) the intermediate position of the Ballón populations along Dim-1 of the

PCA; (ii) diversity values that are intermediate between Chilean and highland values; (iii) intermediacy of the similarity values for the Ballón accessions as compared to the other geographical regions; and (iv) patterns of predominant alleles. We note that Wilson (1988a) included a single representative of the Yungas ecotype in his isozyme analysis which he reported was distinctive. However, this ecotype is not present in the USDA collection and was, therefore, unavailable for inclusion in this study. If the Ballón collection represents a distinct geographical ecotype, the lack of similarly clustering accessions from the CIP-FAO international nursery collection suggests that it is potentially underrepresented in even the international quinoa germplasm collections. Tracing the origins of the Ballón collection should be a primary objective for the curation efforts of the USDA quinoa germplasm.

Conclusions

The results of this study hold several implications for quinoa conservation and improvement programmes. Multiple alleles were observed at marker loci for several of the accessions utilized in the study, suggesting that quinoa accessions (both those held by the USDA and those in the CIP-FAO collection) represent potentially heterogeneous lines of mixed genotypes that should be carefully screened and potentially purified prior to breeding. The lack of passport and morphological data associated with many of the publicly available quinoa accessions in the USDA germplasm system may potentially hinder their utility in traditional breeding programmes. We note, however, that extensive collections that are well documented are available in the Peruvian, Bolivian and Chilean germplasm collections. Application of these microsatellite markers to these South American germplasm collections will further elucidate the full range of genetic diversity within *C. quinoa*. Additionally, this study highlights the need for increased collection from two specific geographical areas: the northern Andes (especially Ecuador and Colombia) and the Yungas (eastern slopes of the Bolivian and Peruvian Andes).

Acknowledgements

This research was supported by grants from the McKnight Foundation, Holmes Family Foundation, and the Erza Taft Benson Agriculture and Food Institute. We gratefully acknowledge D. Brenner at USDA-NPGS, Iowa State University, for taxonomic suggestions and seed contribution. We thank R. Holmes and M. Dilg for technical assistance with the microsatellite assays.

References

Aellen P (1929) Beitrag zur systematic der *Chenopodium*—Arten Amerikas, vorweigend auf Grund der Sammlung des United States National Museum in Washington, D.C. *Feddes Repertorium Specierum Novarum Regni Vegetabilis* 26: 31–67, 119–160.

Cusack DF (1984) Quinua: grain of the Incas. *Ecologist* 14: 21–31.

Dean RE, Dahlberg JA, Hopkins MS, Mitchell SE and Kresovich S (1999) Genetic redundancy and diversity among 'orange' accessions in the US National Sorghum Collection as assessed with simple sequence repeat (SSR) markers. *Crop Science* 39: 1215–1221.

Diwan N and Cregan PB (1997) Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. *Theoretical and Applied Genetics* 95: 723–733.

Elder JK and Southern EM (1987) Computer-aided analysis of one dimensional restriction fragment gels. In: Bishop MJ and Rawlings CJ (eds) *Nucleic Acid and Protein Sequence Analysis—A Practical Approach*. Oxford: IRL Press, pp. 165–172.

Gandarillas H (1979) Genética y orígen. In: Tapia ME (ed.) *Quinua y Kañiwa: Cultivos Andinos*. Bogota: Instituto Interamericano de Ciencias Agrícolas, pp. 45–64.

Gupta PK and Varshney RK (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with special emphasis on bread wheat. *Euphytica* 113: 163–185.

Hokanson SC, Szewc-McFadden AK, Lamboy WF and McFerson JR (1998) Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a *Malus* × *domestica* Borkh. core subset collection. *Theoretical and Applied Genetics* 97: 671–683.

Jacobsen SE (2000) Quinoa—research and development at the International Potato Center (CIP). Paper presented at the Annual Meeting of the Consultative Directorate of CONDESAN, November, pp. 1–5 Lima, peru.

Jain S, Jain RK and McCouch Lima, Peru SR (2004) Genetic analysis of Indian aromatic and quality rice (*Oryza sativa* L.) germplasm using panels of fluorescently-labeled microsatellite markers. *Theoretical and Applied Genetics* 109: 965–977.

Mace ES and Godwin ID (2002) Development and characterization of polymorphic microsatellite markers in taro (*Colocasia esculenta*). *Genome* 45: 823–832.

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

Maughan PJ, Saghai Maroof MA, Buss GR and Huestis GM (1996) Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near-isogenic line analysis. *Theoretical and Applied Genetics* 93: 392–401.

McGregor CE, van Treuren R, Hoekstra R and van Hintum TJL (2002) Analysis of the wild potato germplasm of the series *Acaulia* with AFLPs: implications for *ex situ* conservation. *Theoretical and Applied Genetics* 104: 146–156.

Mujica A, Jacobsen SE, Izquierdo J and Marathee JP (1998) *Prueba Americana y Europea de quinoa (*Chenopodium quinoa *Willd.)*. Puno: FAO, Libro de Campo.

Nelson, DC (1968) Taxonomy and origins of *Chenopodium quinoa* and *Chenopodium nuttalliae*. PhD Thesis, Indiana University.

Ortiz R, Ruiz-Tapia EN and Mujica-Sanchez A (1998) Sampling strategy for a core collection of Peruvian quinoa germplasm. *Theoretical and Applied Genetics* 96: 475–483.

Ott J (1992) Strategies for characterizing highly polymorphic markers in human gene mapping. *American Journal of Human Genetics* 51: 283–290.

Pratt DB and Clark LG (2001) *Amaranthus rudis* and *A. tuberculatus*—one species or two?" *Journal of the Torrey Botanical Society* 128: 282–296.

Risi JC and Galwey NW (1984) The *Chenopodium* grains of the Andes: Inca crops for modern agriculture. *Advances in Applied Biology* 10: 145–216.

Roa NK (2004) Plant genetic resources: advancing conservation and use through biotechnology. *African Journal of Biotechnology* 3: 136–145.

Rohlf FJ (2000) *NTSYSpc 2.1: Numerical Taxonomy and Multivariate Analysis System*. New York: Exeter Software.

Rojas W, Barriga P and Figueroa H (2000) Multivariate analysis of the genetic diversity of Bolivian quinoa germplasm. *Plant Genetic Resources Newsletter* 122: 16–23.

Ruales J and Nair BM (1992) Nutritional quality of the protein in quinoa (*Chenopodium quinoa*, Willd) seeds. *Plant Foods and Human Nutrition* 42: 1–11.

Ruales J, de Grijalva Y, Lopez-Jaramillo P and Nair BM (2002) The nutritional quality of an infant food from quinoa and its effect on the plasma level of insulin-like growth factor-1 (IGF-1) in undernourished children. *International Journal of Food Science and Nutrition* 53: 143–154.

Sambrook J, Fritsch EF and Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edn. New York: Cold Spring Harbor Press.

Simmonds NW (1965) The grain chenopods of the tropical American highlands. *Economic Botany* 19: 223–235.

Simmonds NW (1971) The breeding system of *Chenopodium quinoa*. I. Male sterility. *Heredity* 27: 73–82.

Tang S, Kishore VK and Knapp SJ (2003) PCR-multiplexes for a genome wide framework of simple sequence repeat marker loci in cultivated sunflower. *Theoretical and Applied Genetics* 107: 6–19.

Tapia ME (1979) Historia y distribción geográfica. In: Tapia ME (ed.) *Quinua y Kañiwa: Cultivos Andinos*. Bogota: Instituto Interamericano de Ciencias Agrícolas, pp. 11–19.

Tapia ME, Mujica SA and Canahua A (1980) Orígen, distribución geográfica, y sistemas de producción en quinua. *Primera Reunion Sobre Genética y Fitomejoramiento de la Quinua*. Puno: Universidad Nacional Técnica del Altiplano, Instituto Boliviano de Technologia Agropecuaria, Instituto Interamericana de Ciencias Agricolas, Centro de Investicación Internacional para el Desarollo, pp. A1–A8.

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

Tommasini L, Batley J, Arnold GM, Cooke RJ, Donini P, Lee D, Law JR, Lowe C, Moule C, Trick M and Edwards KJ (2003) The development of multiplex simple sequence repeat (SSR) markers to complement distinctness, uniformity and stability testing of rape (*Brassica napus* L.) varieties. *Theoretical and Applied Genetics* 106: 1091–1101.

Ward SM (2000) Allotetraploid segregation for single-gene morphological characters in quinoa (*Chenopodium quinoa* Willd.). *Euphytica* 116: 11–16.

Wilson HD (1988a) Quinoa biosystematics I: domesticated populations. *Economic Botany* 42: 461–477.

Wilson HD (1988b) Quinoa biosystematics II: free living populations. *Economic Botany* 42: 478–494.