Theor Appl Genet (1998) 96:475–483

R. Ortiz · E. N. Ruiz-Tapia · A. Mujica-Sanchez

Sampling strategy for a core collection of Peruvian quinoa germplasm

Received: 1 May 1997 / Accepted: 20 October 1997

Abstract Quinoa (*Chenopodium quinoa*) is an Andean crop with a high potential for cultivation under temperate agricultural conditions. A new quinoa cultivar for such an environment requires plant characteristics that may be available primarily largely in genetic resources held in gene banks. A core collection may simplify management and enhance the utilisation of quinoa genetic resources. This paper describes the development of a core subset of the whole quinoa gene bank (1029 accessions) of the Universidad Nacional del Altiplano (UNAP). All accessions available in this gene bank have location and altitude descriptors and a partial description for qualitative and quantitative descriptors. The core collection (103 accessions) contains chosen ecotypes or landraces that capture most of the genetic variability available in this Peruvian germplasm. The accessions were selected for the core collection based on a geographically stratified non-overlapping sampling procedure. The number of accessions that were allocated to the core subset was determined using a proportional method adjusted by the relative importance of the quinoa crop in each geographical cluster as determined by its acreage. The sampling method also considered the morphological diversity within four geographical clusters of at least 100 accessions. The multivariate pattern of morphological variation was defined within each of these clusters by independent principal component analyses. A comparison of phenotypic diversity between the entire collection and its core subset confirmed that the proper sampling strategy for this core collection of Peruvian quinoa germplasm had been applied. The most important phenotypic correlations between quantitative descriptors observed in the entire collection, which may be under the control of co-adapted gene complexes, were also preserved by the core collection. The most comprenhensive quinoa core collection should consider accessions from other gene banks in Bolivia and Ecuador, a few accessions from coastal Chile and wild sympatric cross-compatible *Chenopodium* species. This core collection will be a point of entry to the proper exploitation of the genetic resources available in respective quinoa gene banks.

Key words *Chenopodium quinoa* · Co-adapted gene complexes · Genetic resources · Phenotypic diversity

Communicated by P. M. A. Tigerstedt

R. Ortiz (✉)
The Royal Veterinary and Agricultural University (KVL), Department of Agricultural Sciences, 40 Thorvaldsensvej, DK-1871 Frederiksberg C, Copenhagen, Denmark
Fax: +45 3528 3468
E-mail: ro@kvl.dk

E. N. Ruiz-Tapia · A. Mujica-Sanchez
Universidad Nacional del Altiplano-Puno (UNAP), Project Quinoa Danida-CIP-KVL-UNAP, Av. del Ejercito 329, Puno, Peru

Introduction

Quinoa (*Chenopodium quinoa* Willd.) is an important food crop for the Andean population because of its high nutritional value (Weber 1978). Compared with the most important cereals, quinoa has more protein, lysine, fat and fibre (Coulter and Lorenz 1990; Dini et al. 1992). A quinoa kernel (11.2% humidity) consists of 11% protein, 5.3% fat, 68% carbohydrates, 5% fibre and 3% ashes (Carmen 1984). This Andean crop is an important source of lysine (6.8% of total protein), methionine (2.1%), threonine (4.5%) and tryptophan (1.3%). Peru grows almost one-third of the world's total quinoa crop with the highest grain yield per unit area (735 kg ha^{-1}). Bolivia, the world's largest producer, and Ecuador are the other two Andean countries in South America growing quinoa on a relatively large commercial scale, although grain yields in both

countries are lower than those recorded in Peru (511 kg ha^{-1} in Bolivia, and 408 kg ha^{-1} in Ecuador).

Quinoa ($2n = 4x = 36$) is a relatively unimproved pseudocereal native to the Andean region of South America (Gandarillas-Santa Cruz 1976). Farmers grow quinoa for its grain in two distinct ecological zones: Quechua (2000–3000 m above sea level) and Puna (3000–4000 m above sea level). Despite its thermo- and photo-sensitivity, quinoa has a potential as grain crop in Argentina (Gallardo and Gonzalez 1992) and Brazil (Spehar and Souza 1993). Furthermore, its tolerance to cold, drought (Espindola-Canedo and Rodriguez-Ontiveros 1988) and salinity (Gonzalez and Prado 1992), adaptation to marginal or poor environments and excellent nutritional value make quinoa a crop of potential value to areas of the world that lack a diversity of crops with high nutritional value (Risi and Galwey 1984).

Quinoa has been suggested as a potential new crop for animal and human nutrition in Europe (Jacobsen and Stølen 1993) and the USA (Johnson and Ward 1993) because many landraces of this originally short-day-adapted plant flower and set seed in long-day environments (Darwinkel and Stølen 1997). Promising results for acclimatizing quinoa to northern Europe have been obtained in Denmark (Jacobsen et al. 1994, 1996), England (Risi and Galwey 1991a) and Finland (Carmen 1984). New cultivars have been commercially released in the UK, the Netherlands and Denmark, and guidelines for growing quinoa in the temperate regions of northwestern Europe has been published recently (Darwinkel and Stølen 1997). However, the wide cultivation of quinoa in Europe requires the development of short, non-branched, early-maturing cultivars with large, white, low-saponin seeds (Galwey 1992). Therefore, quinoa genetic resources should be explored for further genetic betterment of this crop. For example, Risi and Galwey (1989a) concluded that plant characteristics required for temperate agriculture of quinoa are available to a large extent in accessions collected near sea level in southern-central Chile, whereas seed characteristics are scattered throughout the germplasm. Conversely, Carmen (1984) recommended the utilisation of earlier genetic material available from the quinoa collection of the Banco de Germoplasma de Cultivos Andinos of Universidad Nacional del Altiplano-Puno (UNAP).

Core collections improve the management and effective use of plant genetic resources (Brown 1989). A core collection is a subset of a large germplasm collection that contains chosen accessions capturing most of the genetic variability of the whole gene bank. Such a core subset provides a proper working collection for the extensive searching of desired alleles and a point of entry to the entire germplasm collection (Holbrook and Anderson 1995; Dussert et al. 1997). For example, the search for a desired characteristic or gene(s) would start with the screening of the core collection. After this desired characteristic or allele(s) is found, plant breeders may go back to the reserve collection to screen accessions of similar clusters or geographical area to incorporate more diversity into their genetic improvement programme.

Core collections have been developed for perennial *Glycine* (Brown et al. 1987), annual and perennial *Medicago* species (Diwan et al. 1994; Basigalup et al. 1995), barley (van Hintum et al. 1990), beans (Tohme et al. 1996), cassava, coffee (Dussert et al. 1997), lentil (Erskine and Muelbauer 1991), maize, okra (Mahajan et al. 1996), peanuts (Holbrook et al. 1993), perennial ryegrass (Charmet et al. 1993), emmer wild wheat (Jaradat 1997) and durum wheat (Spagnoletti Zeulli and Qualset 1993). These collections were defined based on passport descriptors, eco-geographical data and geo-statistics, biochemical and molecular data, morphological variability using qualitative and quantitative characterisation data, genetic structure and pedigree analysis (Brown et al. 1990; Charmet and Balfourier 1995; Diwan et al. 1995; Noirot et al. 1996; Peeters et al. 1990; van Hintum 1994a; van Hintum and Haalman 1994; van Hintum et al. 1995). This paper describes the assemblage of a core collection of the Peruvian quinoa germplasm based on information obtained from the database of UNAP.

Materials and methods

A total of 1029 quinoa ecotypes or landraces collected from farms in the Peruvian Andes are held in the gene bank of UNAP (Table 1). All accessions have location and altitude descriptors and a partial description for morphological descriptors (IBPGR 1981). Morphological data were recorded on a categorical scale (stem and inflorescence colour and inflorescence type) or in SI units (days to flowering, inflorescence length, plant height, plant biomass and grain yield). A catalogue of this gene bank containing this information has been published recently (Mujica-Sanchez et al. 1991). These investigators recorded the data for morphological descriptors after growing the collection at the experimental station of the UNAP in Camacani (3825 masl), near the shores of Lake Titicaca (Puno, Peru). The areas surrounding this lake are regarded as the primary centre of diversity for cultivated quinoa.

Statistical analyses of geographical and morphological variation

One-way analysis of variance was carried out to confirm the geographical partitioning of quinoa diversity in the Peruvian Andes. The source of variation was the political department, which almost follows a southwards latitudinal trend in Peru. Similarly, correlation and regression analyses were performed to verify how much of the phenotypic variation was affected by the altitude at which the quinoa landraces were originally collected. Principal component analysis (PCA) for all accessions was not considered since the pattern of variation changed according to the department of collection. Hence, PCA was performed to determine the pattern of multivariate variation in Puno, Cusco, Ancash and Ayacucho, where at least 100 accessions were collected. Specific clusters, based on PCA, were defined within each of these four departments for further random sampling of accessions for the core collection. Altitude intervals of 200 m were defined for subsampling within Huancavelica, Cusco and Ayacucho.

Table 1 Collection sites for Peruvian quinoa accessions held at the gene bank of the Universidad Nacional del Altiplano Puno. Altitude classes (lower limit) are separated by 200-m intervals

Ecozone	Altitude (m)	Cajamarca	Ancash	Junin	Ayacucho	Huancavelica	Arequipa	Apurimac	Cusco	Puno	Total
Quechua	2200						9	21			30
	2400				3						3
	2600	5			85				1		91
	2800								2		2
	3000		109		15	3			75		202
Puna	3200			8		6	12		120		146
	3400					1			28		29
	3600					10			34		44
	3800								10	472	482
	Total	5	109	8	103	20	21	21	270	472	1029

Assembly of the core collection

The strategy for choosing accessions for the core collection consisted of a geographically stratified non-overlapping sampling procedure. The number of accessions that were allocated to the core collection was determined using the proportional method (Brown 1989) adjusted by the relative importance of the crop in each department as determined by the quinoa acreage. This information was obtained from the web site of the Ministerio de Agricultura of Peru (http://www.minag.gob.pe/MINAG/estadistica/cosecha/245289d.htm#qui). For example, Puno grows 85% of the quinoa in Peru, and 46% of the accessions from the gene bank at UNAP were collected in this department. Hence, almost two-thirds of the accessions of the core collection were selected from Puno.

The sampling method for the core collection also considered the diversity within groups. Accessions were chosen within each of the clusters defined earlier by PCA for Puno and Ancash. Sampling for accessions collected in Cusco and Ayacucho considered a balance between the geographical origin and the pattern of multivariate variation of morphological descriptors in both departments. Accessions were randomly selected for Apurimac, Arequipa, Cajamarca, and Junin, and within each altitude for Huancavelica. Selection of accessions for the core collection avoided duplication in the same location in order to maintain the diversity of the gene bank in the core collection. Distinct accesions belonging to the same location were only chosen for the core collection for a few districts in the department of Puno, where many accessions were originally collected. Furthermore, cultivars known in Puno as 'Chehuecca', 'Kancolla' and 'Tahuaco' were added to the core collection.

Comparison of phenotypic diversity in entire collection and core subset

The mean phenotypic diversity between the entire collection and suggested core collection was compared by *t*-tests for all morphological descriptors. Distribution homogeneity for each morphological descriptor between the entire collection and the core collections was analysed by Chi-square tests. The constant, logarithmic and unadjusted proportional methods (Brown 1989) were explored to avoid biased germplasm sampling, especially for two descriptors whose means and frequency distributions differed significantly between the entire and core collections. Simulated means for each sampling method were contrasted with those of the entire collection and our suggested core collection as defined earlier. The conservation of phenotypic associations between descriptors, which may be under genetic control, was investigated by calculating independently phenotypic correlations in the entire and core collections.

Results

The phenotypic variation for all morphological descriptors among the collection sites was highly significant ($P < 0.001$) (Table 2). Similarly, highly significant ($P < 0.001$) correlations between altitude and morphological descriptors with continous quantitative variation were found (Table 3). Altitude of the original collection affected from 2% (plant biomass) to 56% (days to flowering) the continuous variation of the quantitative descriptors recorded for all accessions in Camacani. Landraces collected at high altitudes showed early flowering, short plants, large inflorescences and high plant biomass, thereby exhibiting a high yield potential in this location.

The pattern of multivariate variation differed between departments (Table 4). For example, inflorescence length and type, plant height and days to flowering were the most important descriptors for grouping quinoa germplasm from Ancash, whereas stem and inflorescence colour, plant height and biomass and days to flowering were the descriptors suggested by the PCA for grouping quinoa germplasm collected in Puno. Inflorescence length was the only common descriptor for grouping germplasm in all departments.

The number of accessions of the core for the other geographical clusters, i.e. altitude within department (Table 5), was allocated as follows: 1 accession for 14 clusters with 20 or fewer accessions in the entire collection; 2 accessions for 3 clusters with 21–34 accessions; 3 for 1 cluster with 75 accessions; 4 for 1 cluster with 85 accessions; 5 for 1 cluster with 109 accessions; and 6 for 1 cluster with 120 accessions (Table 6).

There were significant differences between the core and the entire collections for two quantitative morphological descriptors: days to flowering and inflorescence length (Table 7). The core collection was significantly ($P < 0.001$) biased towards early flowering and big inflorescences owing to the percentage of

Table 2 Variation in the qualitative and quantitative morphological descriptors of Peruvian quinoa germplasm according to the department in which the collection was made (Mean ± standard error)

Department	Stem colour (scale 1–10)	Days to flowering (days)	Inflorescence type (scale 1–8)	Inflorescence colour (scale 1–10)	Inflorescence length (cm)	Plant height (cm)	Plant biomass (g)	Grain yield (t/ha)
Ancash	2.4 ± 0.4	121 ± 2	6.4 ± 0.2	5.3 ± 0.5	29 ± 1	125 ± 2	129 ± 31	1.5 ± 1.0
Apurimac	4.6 ± 1.1	138 ± 2	6.5 ± 0.5	4.6 ± 0.5		126 ± 6	77 ± 16	0.5 ± 0.1
Ayacucho	2.6 ± 0.3	118 ± 2	6.4 ± 0.2	3.1 ± 0.2	25 ± 2	114 ± 2	115 ± 14	1.0 ± 0.1
Cajamarca	6.0	110	5.0	9.0	20	85	40	1.0
Cusco	4.5 ± 0.2	119 ± 1	4.2 ± 0.2	4.1 ± 0.1	24 ± 0.4	106 ± 1	94 ± 3	0.8 ± 0.0
Huancavelica	1.2 ± 0.2	117 ± 3	7.0 ± 0.0	2.1 ± 0.5	16 ± 2	110 ± 2	76 ± 11	0.6 ± 0.1
Junin	3.1 ± 0.0	78 ± 4	5.0 ± 0.0	1.0 ± 0.0	32 ± 4	77 ± 3	110 ± 21	3.8 ± 1.4
Puno	3.0 ± 0.1	76 ± 0.4	6.0 ± 0.1	4.3 ± 0.2	33 ± 0.4	91 ± 1	133 ± 6	2.7 ± 0.1

Table 3 Correlation coefficient (*r*) and regression equation between the phenotypic variation (Y) of the morphological descriptor in Peruvian quinoa germplasm and the altitude of the collection (X)

Descriptor (Y)	Correlation (*r*)	Regression equation	Significance level
Stem colour	−0.064	–	0.109
Days to flowering	−0.750	248.040 − 0.043 X	<0.001
Inflorescence type	0.044	–	0.271
Inflorescence colour	0.029	–	1.000
Inflorescence length	0.436	0.009 X − 1.213	<0.001
Plant height	−0.547	184.452 − 0.024 X	<0.001
Plant biomass	0.158	2.715 + 0.032 X	<0.001
Grain yield	0.494	0.002 X − 4.413	<0.001

accessions allocated to Puno. This was confirmed by the analysis of frequency distribution (Table 8). However, most of the phenotypic variation recorded in the entire collection for both characteristics is available in the core collection. For example, five classes were defined for inflorescence length, ranging from very short (10–19 cm) to very large (50–59 cm). While the core collection has accessions in the five classes it contains a significantly ($P < 0.05$) higher percentage of accessions in the intermediate and large classes and a lower percentage in the short class than the entire collection. Similarly, four intervals were defined for days to flowering: early (67–96 days), intermediate (97–126), late (127–142) and very late (143 days onwards). The core consists of 63% early-; 23% intermediate- and 14% late-flowering accessions, while the entire collection has 39% early-, 42% intermediate-, 19% late- and 1 very late flowering accession. Other approaches such as the constant and logarithmic methods may not solve this bias on flowering, as determined by simulated sampling (data not shown). Both methods may result in a biased core collection with a significantly ($P < 0.05$) higher mean than the entire collection. Only the proportional

Table 4 Eigen vector values, latent roots and percentage of variation explained by principal components (Prin) of the accessions collected in the respective departments. The most important loading descriptors are in brackets

Department	Prin 1	Prin 2	Prin 3
Ancash	0.698 (Inflorescence length) 0.650 (Plant height)	0.708 (Days flowering)	0.631 (Inflorescence type)
Latent root	1.635	1.435	0.967
Variation (%)	*32.700*	*29.097*	*19.343*
Ayacucho	0.472 (Inflorescence length) 0.411 (Inflorescence type) −0.443 (Stem colour)	−0.591 (Altitude) −0.405 (Days flowering)	−0.564 (Grain yield)
Latent root	3.279	2.020	1.450
Variation (%)	*36.430*	*22.442*	*16.116*
Cusco	0.527 (Stem colour) 0.453 (Inflorescence colour)	0.576 (Plant height) 0.419 (Inflorescence length)	0.689 (Grain yield)
Latent root	2.418	1.507	1.205
Variation (%)	*26.868*	*16.750*	*13.885*
Puno	0.480 (Inflorescence length) 0.467 (Plant height) 0.462 (Plant biomass)	0.607 (Inflorescence colour) 0.545 (Stem colour)	−0.527 (Days flowering)
Latent root	2.432	1.890	1.435
Variation (%)	*24.918*	*18.898*	*14.352*

Table 5 Sampling for a core collection of Peruvian quinoa germplasm based on the proportional method adjusted by acreage. Altitude interval class (lower limit) for collection is in brackets

Regions department	Number of accessions in gene bank	Percentage of collection	Acreage by region (ha)	Percentage of acreage (%)	Number of accessions for core subset
MARANON	*5*	*0.5*	*61*	*0.4*	*1*
Cajamarca (2600)	5	0.5	61	0.4	1
CHAVIN	*109*	*10.6*	*124*	*0.8*	*5*
Ancash (3000)					
CACERES	*8*	*0.8*	*374*	*2.3*	*1*
Junin (3200)	8	0.8	159	1.0	1
WARI	*144*	*14.0*	*748*	*5.9*	*13*
Ayacucho	103	10.0	312	1.9	6
(2400)	3	0.3			1
(2600)	85	8.3			4
(3000)	15	1.5			1
Huancavelica	20	1.9	240	1.5	5
(3000)	3	0.3			1
(3200)	5	0.5			1
(3300)	1	0.1			1
(3400)	1	0.1			1
(3600)	10	1.0			1
Apurimac	21	2.0	396	2.5	2
AREQUIPA	*21*	*2.0*	*60*	*0.4*	*2*
(2200)	9	0.9			1
(3200)	12	1.2			1
INCA (*Cusco*)	*270*	*26.2*	*730*	*4.6*	*16*
(2600)	1	0.1			1
(2800)	2	0.2			1
(3000)	75	7.3			3
(3200)	120	11.7			6
(3400)	28	2.7			2
(3600)	34	3.3			2
(3800)	10	1.0			1
MARIATEGUI					
Puno (*3800*)	*472*	*45.8*	*13.615*	*84.9*	*65*

Table 6 Accessions selected for the core collection of Peruvian quinoa germplasm

Department	Accession number
Ancash	495, 502, 506, 519, 1002
Apurimac	991, 1000
Arequipa	1003, 1012
Ayacucho	440, 573, 579, 919, 940, 955
Cajamarca	585
Cusco	620, 627, 632, 694, 701, 703, 744, 791, 800, 804, 822, 827, 834, 841, 878, 887
Huancavelica	966, 972, 974, 975, 980
Junin	876
Puno	7, 12, 18, 19, 25, 35, 36, 43, 49, 50, 55, 61, 69, 77, 80, 89, 97, 106, 112, 118, 125, 140, 143, 146, 157, 165, 173, 175, 177, 186, 197, 211, 230, 246, 249, 256, 280, 291, 301, 322, 333, 346, 354, 368, 371, 377, 382, 446 (cv Chehuecca), 584 (Kancolla Rosada-Camacani), 854, 856, 859, 864, 865, 870, 880, 882, 889, 892, 894, 908, 910, 1029 (cv Tahuaco), cv Kancolla

method, without any adjustment for acreage, may provide a core collection with an intermediate but significant ($P < 0.05$) mean between the core and the entire collections (data not shown).

Stem and inflorescence colour were highly correlated ($r > 0.51$; $P < 0.001$) in the entire and core collections (Table 9). The discrete variation of qualitative descriptors was not significantly associated ($P > 0.05$) in the core collection with the continuous variation observed on morphological quantitative descriptors (Table 9). However, these associations were significant ($P < 0.05$) but with a low correlation coefficient in the entire collection. There were significant ($P < 0.05$) phenotypic associations between days to flowering with inflorescence length, plant biomass, plant height and grain yield in the entire and core collections.

Table 7 Means (±standard error) for morphological descriptors for the entire collection and for the core collection of Peruvian quinoa germplasm. Ranges are in brackets

Descriptor	Entire collection	Core collection	Comparison
Stem colour (scale 1–10)	3.4 ± 0.1 (1–10)	3.5 ± 0.3 (1–10)	NS
Days to flowering	103 ± 1 (67–197)	93 ± 2 (67–145)	*
Inflorescence type (scale 1–8)	5.4 ± 0.1 (1–8)	5.6 ± 0.2 (1–8)	NS
Inflorescence colour (scale 1–10)	4.1 ± 0.1 (1–10)	4.5 ± 0.3 (1–10)	NS
Inflorescence length (cm)	28 ± 0.3 (10–58)	31 ± 1 (15–53)	*
Plant height (cm)	103 ± 1 (45–155)	101 ± 2 (60–150)	NS
Plant biomass (g)	112 ± 3 (6–750)	139 ± 12 (6–750)	NS
Grain yield (t/ha)	1.7 ± 0.1 (0.07–8.09)	2.0 ± 0.2 (0.07–8.09)	NS

NS and * indicate non-significant or significant mean difference as determined by respective overlapping or non-overlapping confidence interval at the 95% level

Table 8 Comparison of frequency distribution for morphological descriptors of the entire collection and the core collection of Peruvian quinoa germplasm

Descriptor	Number of classes	χ^2	P
Stem colour	4	1.041	0.904
Days to flowering	4	18.799	0.0003
Inflorescence type	4	4.721	0.193
Inflorescence colour	10	20.133	0.017
Inflorescence length	5	11.647	0.020
Plant height	8	3.163	0.870
Plant biomass	6	7.652	0.177
Grain yield	7	11.045	0.087

Discussion

The sampling strategy used to develop the suggested core collection from the Peruvian quinoa germplasm was based on a hierarchical structure, as defined by the ecological zone of origin of the accessions. Also, the multivariate pattern of morphological variation within the clusters for departments with at least 100 accessions was considered in the sampling procedure. Our core subset, which consists of 103 accessions (Table 6), or 10% of the entire collection, should include with minimum redundancy the genetic diversity of the Peruvian quinoa germplasm.

Geographical variation of morphological descriptors in Peruvian quinoa germplasm

The hierarchical sampling structure adopted to develop the core collection must be reasonable on the basis of prior knowledge. Risi and Galwey (1989b) reported that accessions collected around Lake Titicaca were distinct from those collected in the Andean valleys farther north. Similarly, Wilson (1987) indicated that variation among Andean quinoa populations followed a latitudinal trend, with relatively high levels of variation in Bolivia and southern Peru. Analysis of morphological and electrophoretic data also suggests an ecotypic approach to landrace classification (Wilson 1988a). The one-way analysis of variance that we used to compare the phenotypic diversity of quinoa germplasm among Peruvian departments (Table 2) agrees with these earlier findings.

The core subset should be able to predict sources of useful variation. For example, the accessions collected in Puno showed higher grain yield, earlier flowering, larger inflorescences and shorter plant height than those collected elsewhere (except Junin). Highland farmers from Puno have selected short plants to avoid damage caused from hail storms and to increase tolerance to winds, which may occur during the growing season in the Peruvian Plateau (known as Altiplano in Spanish). These farmers selected landraces with enough leaf area to enhance photosynthesis and with big inflorescences to obtain large seeds. This explains the significant correlations (Table 3) between altitude and the quantitative descriptors. Most quantitative variation has polygenic control and its similarity could have arisen from convergence (van Hintum 1994b), which make quantitative descriptors suitable for studying similar adaptation patterns.

Validation of sampling strategy for choosing accessions of the core collection

A core collection consisting of 10% of the whole gene bank often contains 70% of the alleles of entire collection (Brown 1989). This sampling procedure is based on the theory of neutral marker alleles and considers the incorporation of rare widespread alleles in the core collection. A good core collection avoids redundant accessions, although it is large enough but of manageable size to yield reliable results. For example, the frequency distributions for inflorescence colour between the core and the entire collections were significantly ($P < 0.05$) different (Table 8). The core collection has a relatively lower number of accessions with yellow inflorescences (13%) than the entire collection (21%). Conversely, the core collection has relatively more accessions with grey inflorescences (29%) than the entire

Table 9 Correlation coefficients (r) between morphological descriptors (based on accession means) in the entire collection (above diagonal) and core collection (below diagonal). Significance level for false rejection of null hypothesis ($r = 0$) is indicated in brackets

Descriptor	1	2	3	4	5	6	7	8
1 Stem colour		0.237 (<0.001)	−0.333 (<0.001)	0.510 (<0.001)	−0.246 (<0.001)	0.092 (0.026)	−0.089 (0.034)	−0.144 (<0.001)
2 Days to flowering	−0.023 (0.836)		−0.242 (<0.001)	0.027 (0.502)	−0.578 (<0.001)	0.497 (<0.001)	0.157 0.017	−0.673 (<0.001)
3 Inflorescence type	−0.122 (0.247)	−0.031 (0.774)		−0.104 (0.010)	0.275 (<0.001)	−0.008 (0.048)	0.020 (0.630)	0.195 (<0.001)
4 Inflorescence colour	0.539 (<0.001)	−0.136 (0.209)	−0.031 (0.764)		−0.005 (0.905)	0.100 (0.015)	−0.107 (0.010)	0.010 (0.819)
5 Inflorescence length	0.029 (0.795)	−0.499 (<0.001)	−0.003 (0.976)	0.059 (0.587)		0.022 (0.592)	0.368 (<0.001)	0.335 (<0.001)
6 Plant height	−0.076 (0.479)	0.531 (<0.001)	−0.059 (0.580)	−0.014 (0.898)	0.051 (0.640)		0.093 (0.032)	−0.437 (<0.001)
7 Plant biomass	−0.142 (0.191)	−0.261 (0.017)	−0.002 (0.987)	−0.158 (0.139)	0.462 (<0.001)	0.126 (0.244)		0.124 (0.004)
8 Grain yield	0.001 (0.993)	−0.648 (<0.001)	0.091 (0.398)	0.035 (0.744)	0.194 (0.081)	−0.414 (<0.001)	0.061 (0.573)	

collection (15%). Nonetheless, the core collection has accessions in all the defined classes for inflorescence colour. Most importantly, the frequency of quinoa accessions with white inflorescences was similar ($P > 0.05$) in the core (29%) and the entire (25%) collections. Farmers grow quinoa landraces with white inflorescences, which may have sweet seeds, for further sale in the urban markets. Also, white seeds have relatively low saponine content.

Qualitative descriptors such as inflorescence colour are useful for tracing the domestication route, for investigating the relation between environment and diversity and for determining the complete crop gene pool diversity or a specific part of this gene pool (van Hintum 1994b). In quinoa, the darker the colour of the inflorescence the higher the tolerance to frost and drought, which explains the collection of accessions with darker inflorescences in farmers' fields far away from the Lake Titicaca in Puno. However, quinoa landraces with dark inflorescences have bitter seeds due to their high saponine content. Nevertheless, after successive washing of the seeds with water, these quinoa seeds can be eaten (Ruales and Nair 1993). This explains why farmers are still growing quinoa with dark inflorescences in the Plateau of Puno.

Genetically, the core collection contains major subspecific taxa and geographic regions, and broadly adapted rather than ecologically adapted alleles or phenotypes. Common alleles are surely in the core collection, and common localised alleles should have not been overlooked in the core due to the proper strategy for germplasm sampling. Genetic diversity, as measured by number of alleles per locus, is maximised following the above genetic criteria. For example, the sampling strategy for the Peruvian quinoa core collection did not pick up a single accession showing very late flowering, which was collected at Urcos (3120 m above sea level) in Cusco. This was not unexpected because a core collection may miss very rare localised alleles, which are often not considered because of the impracticability of conserving everything (van Hintum 1994b) or owing to the relatively little value of rare localised alleles anywhere (Allard 1996). Nevetherless, the Peruvian quinoa core collection has other accessions collected at the same altitude in this department (Table 5).

Phenotypic correlations and co-adapted gene complexes

A proper sampling for a core collection must conserve the phenotypic associations arising from co-adapted gene complexes. The core collection preserves most of the significant ($P < 0.001$) phenotypic associations ($r > 0.40$) which have a coefficient of determination (R^2) in excess of 10% in the entire collection (Table 9). This finding suggests that co-adapted gene complexes controlling these significant associations in quinoa were sampled properly by our method to develop the core collection of Peruvian quinoa germplasm. Hence, this core collection may be very useful for quinoa researchers (and breeders) focusing their work on specific gene complexes. Furthermore, the significant association between stem and inflorescence colour confirms the partial common genetic control for pigmentation in quinoa (Gandarillas-Santa Cruz 1976).

The lack of correlation between the variation of qualitative and quantitative descriptors in the core collection was not surprising. Gandarillas-Santa Cruz and Espindola-Canedo (1981) found no relationship between inflorescence type and grain yield, nor with its components such as earliness, plant height, inflorescence length or seed diameter, in quinoa cultivars.

They observed significant differences for most of these characteristics within each inflorescence type. Conversely, the significant associations between the continous variation of the morphological quantitative descriptors were not unexpected. Espindola and Gandarillas (1985) reported significant phenotypic correlations between plant height and inflorescence length with grain yield in 36 accessions collected in Bolivia and Peru. These authors also mentioned that path analysis revealed that inflorescence length was the most important component of grain yield in the cultivated accessions held at the Bolivian gene bank. Similarly, Risi and Galwey (1991b) indicated that earliness and yield are associated among quinoa landraces from distinct South American locations.

Our sampling strategy for the development of the core collection of Peruvian quinoa germplasm was able to explain the known phenotypic correlations in quinoa germplasm adequately. Hence, this quinoa core collection may be an effective tool for further investigations with molecular markers of multi-locus structures within each sampling cluster in this species.

Adding other *Chenopodium* germplasm to a quinoa core collection

A core collection is a dynamic entity subjected to change in size and content. New accessions from new distinct sources may be added into a core collection, or other accessions of questionable authenticity may be replaced. Similarly, accessions may be added or replaced in the core after new groupings are defined, based on new data or analytical tools, or due to changes in users' needs.

Quinoa is a self-pollinated crop (i.e. inbreeding species) bearing hermaphrodite flowers, although 2–9% of outcrossing rates have been reported in this crop (Gandarillas 1967; Gomez-Meza and Lopez-Guadarrama 1990; Simmonds 1965). Gene flow between cultivated quinoa and its sympatric wild form *C. hircinum* have been confirmed by electrophoretic data (Wilson 1981, 1986, 1988b, c). Furthermore, "wild" *Chenopodium* plants growing within and in the periphery of quinoa fields are crop-weed hybrids owing to an assymetric pollen flow from the cultivated to the wild species (Wilson and Manhart 1993). Similarly, Wilson (1990) pointed out that Andean populations of *Chenopodium* sect. *Chenopodium* subsect. *Cellulata* represent a monophyletic crop-weed system developed through cyclic differentiation and introgressive hybridisation. Therefore, a comprenhensive core collection of quinoa germplasm should include this Peruvian core collection, other unique accessions from the gene banks of Bolivia and Ecuador, a few accessions of "quingua" from coastal Chile and some accessions of sexually compatible sympatric *Chenopodium* wild species.

Acknowledgements E. N. Ruiz-Tapia acknowledges the financial support of the Danish International Development Assistance (DANIDA) for his research work in Denmark.

References

Allard RW (1996) Genetic basis of the evolution of adaptedness in plants. Euphytica 92 : 1–11

Basigalup DH, Barnes DK, Stucker RE (1995) Development of a core collection for perennial *Medicago* plant introductions. Crop Sci 35 : 1163–1168

Brown AHD (1989) Core collections: a practical approach to genetic resources management. Genome 31 : 818–824

Brown AHD, Grace JP, Speer SS (1987) Designation of a core collection of perennial *Glycine*. Soybean Genet Newsl 14 : 59–70

Brown AHD, Burdon JJ, Grace PP (1990) Genetic structure of *Glycine canescens*, a perennial relative of soybean. Theor Appl Genet 79 : 729–736

Carmen ML (1984) Acclimatization of Quinoa (*Chenopodium quinoa* Willd.) and Canihua (*Chenopodium pallidicaule* Allen) to Finland. Ann Agric Fenn 23 : 135–144

Charmet G, Balfourier F (1995) The use of geostatistics for sampling a core collection of perennial ryegrass populations. Genet Resource Crop Evol 42 : 303–309

Charmet G, Balfourier F, Ravel C, Denis JB (1993) Genotype X environment interaction in a core collection of French perennial ryegrass populations. Theor Appl Genet 86 : 731–736

Coulter L, Lorenz K (1990) Quinoa-composition, nutritional value, food applications. Lebensm Wiss Technol 23 : 203–207

Darwinkel A, Stlen O (1997) Understanding the quinoa crop: guidelines for growing in temperate regions of N.W. Europe. European Comission (DG VI F.II.3 – Coordination of Agricultural Research), Dynacom

Dini A, Rastrelli L, Saturnino P, Schettino O (1992) A compositional study of *Chenopodium quinoa* seeds. Nahrung 36 : 400–404

Diwan N, Bauchan GR, McIntosh MS (1994) A core collection for the United States annual *Medicago* germplasm collection. Crop Sci 34 : 279–285

Diwan N, McIntosh MS, Bauchan GR (1995) Methods of developing a core collection of annual *Medicago* species. Theor Appl Genet 90 : 755–761

Dussert S, Chabrillange N, Anthony F, Engelmann F, Recalt C, Hamon S (1997) Variability in storage response within a coffee (*Coffea* spp.) core collection under slow growth conditions. Plant Cell Rep 16 : 344–348

Erskine W, Muelbauer (1991) Allozyme and morphological variability, outcrossing rate and core collection formation in lentil germplasm. Theor Appl Genet 83 : 119–125

Espindola G, Gandarillas H (1985) Study of correlated characters and their effects on quinoa yield. Bol Genet 13 : 47–54

Espindola-Canedo G, Rodriguez-Ontiveros JL (1988) Respuestas fisiologicas y del rendimiento de la quinua a deficit hidrico. Agrociencia 74 : 297–312

Gallardo MG, Gonzalez JA (1992) Efecto de algunos factores ambientales sobre la germinacion de *Chenopodium quinoa* Willd. y sus posibilidades de cultivo en algunas zonas de la provincia de Tucuman (Argentina). Lilloa 38 : 55–64

Galwey NW (1992) The potential of quinoa as a multi-purpose crop for agricultural diversification: a review. Ind Crops Products 1 : 101–106

Gandarillas H (1967) Observaciones sobre la biologia reproductiva de la quinua. Rev Sayana 5 : 26–29

Gandarillas-Santa Cruz H (1976) The genetics and origin of quinoa (*Chenopodium quinoa*). Bol Genet 9 : 3–14

Gandarillas-Santa Cruz H, Espindola-Canedo G (1981) Relacion entre el rendimiento y la forma de la panoja en quinua. Turrialba 31 : 385–388

Gomez-Meza ME, Lopez-Guadarrama A (1990) Tipo de polinizacion en la quinua (*Chenopodium quinoa* Will.). Rev Chapingo 15 : 71–72
Gonzalez JA, Prado FE (1992) Germination in relation to salinity and temperature in *Chenopodium quinoa* Willd. Agrochimica 36 : 101–107
Holbrook CC, Anderson WF (1995) Evaluation of a core collection to identify resistance to late leafspot in peanut. Crop Sci 35 : 1700–1702
Holbrook CC, Anderson WF, Pittman RN (1993) Selection of a core collection from the US germplasm collection of peanuts. Crop Sci 33 : 859–861
IBPGR (1981) Descriptores de quinua. International Board for Plant Genetic Resources, Rome, Italy
Jacobsen SE, Stlen O (1993) Quinoa-morphology, phenology for its production as a new crop in Europe. Eur J Agron 2 : 19–29
Jacobsen SE, Jørgensen I, Stølen O (1994) Cultivation of quinoa under climatic conditions in Denmark. J Agric Sci 122 : 47–52
Jacobsen SE, Hill J, Stlen O (1996) Stability of quantitative traits in quinoa (*Chenopodium quinoa*). Theor Appl Genet 93 : 110–116
Jaradat AA (1997) Wild emmer wheat in Jordan: a core collection. Isr J Plant Sci 45 : 45–51
Johnson DL, Ward SM (1993) Quinoa. In: Janick J, Simon JE (eds) New crops. John Wiley and Sons, New York, pp 222–227
Mahajan RK, Bisht IS, Agrawal RC, Rana RS (1996) Studies on south Asian okra collection: a methodology for establishing a representative core set using characterization data. Genet Resource Crop Evol 43 : 244–255
Mujica-Sanchez A, Rodriguez-Ponce M, Pino-Valencia S, Apaza-Mamami V, Huapaya-Estrada F (1991) Catalogo de cultivos andinos: quinua (*Chenopodium quinoa* Willd.) Universidad Nacional del Altiplano Puno and IBPGR, Puno, Peru
Noirot M, Hamon S, Anthony F (1996) The principal component scoring: a new method of constituting a core collection using quantitative data. Genet Resource Crop Evol 43 : 1–6
Peeters JP, Wilkes HG, Galwey NW (1990) The use of ecogeographical data in the exploitation of variation from gene banks. Theor Appl Genet 80 : 110–112
Risi J, Galwey N (1984) The *Chenopodium* grains of the Andes: Inca crops for modern agriculture. Adv Appl Biol 10 : 145–216
Risi J, Galwey N (1989a) The pattern of genetic diversity in the Andean grain crop quinoa (*Chenopodium quinoa* Willd.). I. Associations between characteristics. Euphytica 41 : 147–162
Risi J, Galwey N (1989b) The pattern of genetic diversity in the Andean grain crop quinoa (*Chenopodium quinoa* Willd.). II. Multivariate method. Euphytica 41 : 135–145
Risi J, Galwey N (1991a) Effects of sowing rate and development and grain yield of quinoa (*Chenopodium quinoa*) in a temperate environment. J Agric Sci 117 : 325–332
Risi J, Galwey N (1991b) Genotype X environment interaction in the Andean grain crop (*Chenopodium quinoa*) in temperate regions. Plant Breed 107 : 141–147
Ruales J, Nair BM (1993) Saponins, phytic acid, tannins and protease inhibitors in quinoa (*Chenopodium quinoa* Willd.) seeds. Food Chem 48 : 137–143
Simmonds NW (1965) The grain chenopods of the tropical American highlands. Econ Bot 19 : 223–235
Spagnoletti Zeulli P, Qualset CO (1993) Evaluation of five strategies for obtaining a core collection of durum wheat. Theor Appl Genet 78 : 295–304
Spehar CR, Souza PIM (1993) Adaptacao da quinoa (*Chenopodium quinoa* Willd.) ao cultivo nos cerrados do Planalto Central: resultados preliminares. Pesqu Agropecu Bras Ser Agron 28 : 635–639
Tohme J, Gonzalez DO, Beebe S, Duque MC (1996) AFLP analysis of gene pools of a wild bean core collection. Crop Sci 36 : 1375–1384
van Hintum ThL (1994a) Comparison of marker systems and construction of a core collection in a pedigree of European spring barley. Theor Appl Genet 89 : 991–997
van Hintum ThL (1994b) Drowning in the genepool: managing genetic diversity in the genebank collections. The Swedish University of Agricultural Sciences, Department of Plant Breeding Research, Svalov, Sweden
van Hintum ThL, Haalman D (1994) Pedigree analysis for composing a core collection of modern cultivars with examples from barley (*Hordeum vulgare* s. lat.). Theor Appl Genet 88 : 70–74
van Hintum Th, von Bothmer R, Fischbeck G, Knüpffer H (1990) The establishment of the barley core collection. Barley Newsl 34 : 41–42
van Hintum ThJL, von Bothmer R, Visser DL (1995) Sampling strategies for composing a core collection of cultivated barley (*Hordeum vulgare* s.lat.) collected in China. Hereditas 122 : 7–15
Weber EJ (1978) The Inca's ancient answer to food shortage. Nature 272 : 486
Wilson HD (1981) Genetic variation among South American populations of tetraploid *Chenopodium* sect. *Chenopodium* subsect. *Cellulata*. Syst Bot 6 : 380–398
Wilson HD (1986) Allozyme variation and weed/crop differentiation in *Chenopodium quinoa* complex (abstr). Am J Bot 73 : 5
Wilson HD (1987) Biosystematic analysis of *Chenopodium quinoa* Willd. spp. *quinoa* (abstr). Am J Bot 74 : 766
Wilson HD (1988a) Quinua biosystematics I: domesticated populations. Econ Bot 42 : 461–477
Wilson HD (1988b) Quinua biosystematics II: free living species. Econ Bot 42 : 478–494
Wilson HD (1988c) Allozyme variation and morphological relationships of *Chenopodium hircinum* (s.l.). Syst Bot 13 : 215–228
Wilson HD (1990) Quinua and relatives (*Chenopodium* sect. *Chenopodium* subsect. *Cellulata*). Econ Bot 44 [Suppl] : 92–110
Wilson HD, Manhart J (1993) Crop/weed gene flow: *Chenopodium quinoa* Willd. and *C. berlandieri* Moq. Theor Appl Genet 86 : 642–648