Theor Appl Genet (2004) 109: 1188–1195
DOI 10.1007/s00122-004-1730-9

ORIGINAL PAPER

P. J. Maughan · A. Bonifacio · E. N. Jellen · M. R. Stevens · C. E. Coleman · M. Ricks · S. L. Mason · D. E. Jarvis · B. W. Gardunia · D. J. Fairbanks

A genetic linkage map of quinoa (*Chenopodium quinoa*) based on AFLP, RAPD, and SSR markers

Received: 25 November 2003 / Accepted: 13 May 2004 / Published online: 12 August 2004

Abstract Quinoa (*Chenopodium quinoa* Willd.) is an important seed crop for human consumption in the Andean region of South America. It is the primary staple in areas too arid or saline for the major cereal crops. The objective of this project was to build the first genetic linkage map of quinoa. Selection of the mapping population was based on a preliminary genetic similarity analysis of four potential mapping parents. Breeding lines 'Ku-2' and '0654', a Chilean lowland type and a Peruvian Altiplano type, respectively, showed a low similarity coefficient of 0.31 and were selected to form an F_2 mapping population. The genetic map is based on 80 F_2 individuals from this population and consists of 230 amplified length polymorphism (AFLP), 19 simple-sequence repeat (SSR), and six randomly amplified polymorphic DNA markers. The map spans 1,020 cM and contains 35 linkage groups with an average marker density of 4.0 cM per marker. Clustering of AFLP markers was not observed. Additionally, we report the primer sequences and map locations for 19 SSR markers that will be valuable tools for future quinoa genome analysis. This map provides a key starting point for genetic dissection of agronomically important characteristics of quinoa, including seed saponin content, grain yield, maturity, and resistance to disease, frost, and drought. Current efforts are geared towards the generation of more than 200 mapped SSR markers and the development of several recombinant-inbred mapping populations.

Communicated by D. Hoisington

P. J. Maughan (✉) · E. N. Jellen · M. R. Stevens · C. E. Coleman · M. Ricks · S. L. Mason · D. E. Jarvis · B. W. Gardunia · D. J. Fairbanks
Department of Plant and Animal Sciences, Brigham Young University,
Provo, UT, 84602, USA
e-mail: Jeff_Maughan@byu.edu
Tel.: +1-801-4228698
Fax: +1-801-4220008

A. Bonifacio
The Foundation for the Promotion and Investigation of Andean Products (PROINPA),
La Paz, Bolivia

Introduction

Quinoa (*Chenopodium quinoa* Willd.) is one of the most important food crops in the Andean region of South America, including regions of Bolivia, Peru, Ecuador, Colombia, Argentina, and Chile. It is a member of the family Amaranthaceae (formerly Chenopodiaceae), which also contains the economically important species spinach (*Spinacea oleracea* L.) and sugar beet (*Beta vulgaris* L.). Quinoa is an allotetraploid ($2n=4x=36$) and shows disomic inheritance for most qualitative traits (Simmonds 1971; Risi and Galwey 1984; Ward 2000). Quinoa is especially important as a major food stable in the Altiplano (high plain) regions of Bolivia and Peru (Wilson 1988). The Altiplano covers 255,000 km^2 at an altitude of 3,500–3,850 m above sea level. It has a relatively cold and arid climate with frequent frosts and annual rainfall ranging from 80–600 mm/year. In spite of such harsh conditions, more than two million people reside there, half in rural areas. Quinoa is remarkably adapted to the high altitudes, arid and saline soils, and frequent frosts characteristic of the Altiplano (Risi and Galwey 1984; Vacher 1998; Prado et al. 2000).

For people who live in the Altiplano region, quinoa is a traditional, widely consumed, food crop that is a part of their daily diet. Quinoa grain has an excellent balance of carbohydrates, lipids, and protein and provides an ideal balance of essential amino acids for human nutrition (Risi and Galwey 1984; Chauhan et al. 1999). Quinoa constitutes the principal source of protein in the daily diet of many people who inhabit the Andean region, especially the rural poor. The protein content of quinoa varies between 7.5–22.1% (Tapia et al. 1979), substantially higher than that of the cereal grains. Oil content in quinoa ranges from 4.5–8.75%, with the majority of these oils being oleic (24%), linoleic (52%) and linolenic acids (4%) (Ruales and Nair 1993; Wood et al. 1993; Fleming and Galwey 1995). Several anti-nutritional triterpenoid compounds, called saponins, are found on the pericarp of mature quinoa seeds; these compounds adversely affect digestibility and nutrient uptake, although they may play a

positive role in deterring avian predation (Chauhan et al. 1992). The development of saponin-free varieties is consequently a major breeding goal and one where marker-assisted selection (MAS) could be highly effective (Ward 2001). Because of quinoa's nutritional qualities, its use as a subsistence staple across the Altiplano, and its potential as an export crop, increasing quinoa productivity is viewed as critical for addressing malnourishment problems in Andean region.

While historical research in quinoa genetics and breeding is limited (see reviews by Risi and Galwey 1984; Fleming and Galwey 1995), modern breeding programs, supported by private and government agencies, are quickly becoming established in Bolivia and Peru. Principal objectives for these programs include traditional breeding goals for grain yield, disease resistance (Ochoa et al. 1999), drought tolerance, and saponin content, as well as the development of molecular markers to manage germplasm and facilitate traditional plant breeding.

To facilitate the application of molecular tools and to enhance the knowledge base in quinoa, we report the development of several molecular markers and a preliminary genetic linkage map for quinoa. Molecular markers and genetic maps are powerful tools for plant breeding research programs (Patterson et al. 1991; Staub et al. 1996). They are particularly important for germplasm conservation, including core-collection development and cultivar identification (Diwan et al. 1995; Tanksley and McCouch 1997). Furthermore, they allow for development of enhanced breeding methodologies, including MAS and marker-aided backcross breeding. The molecular markers and genetic map reported here will be of particular value in ongoing efforts to genetically characterize germplasm collections of quinoa and to develop core collections that can be utilized by traditional breeding programs throughout the Andean region.

Materials and methods

Genetic material

Four potential mapping parents representing two distinct ecotypes for quinoa, coastal ('Ku-2', 'NL-6') and Altiplano ('0654', 'Chucapaca') types, were screened with amplified fragment length polymorphism (AFLP) markers to assess their genetic diversity and to determine the most informative crosses for use in developing the genetic linkage map. The parents selected to develop the mapping population were 'Ku-2' (as the female parent), a Chilean coastal quinoa accession with green plant color, and '0654' (as the male parent), a Peruvian Altiplano quinoa accession with red plant color. The mapping population consisted of 80 individual F_2 plants obtained by selfing a single F_1 plant. All plants used in this study were grown in commercial potting-soil supplement with fertilizer in 25-cm-diameter pots at greenhouses in Provo, Utah, USA. Greenhouse temperature was maintained at 25°C, and plants were grown under broad-spectrum halogen lamps with a 12-h photoperiod. Fleming and Galwey (1995) reported that plant color in quinoa is governed by a single gene with three major alleles: Red (R), dominant to purple (r^P), which is dominant to green (r). Plant color-inflorescence phenotypes (R vs r) were determined for the parents, F_1, and F_2 populations.

DNA extraction

Total genomic DNA from all F_2 individuals as well as from the parents was extracted from 30 mg freeze-dried leaf tissue by procedures described by Sambrook et al. (1989), with modifications described by Todd and Vodkin (1996).

AFLP protocol

The AFLP analysis was performed essentially as described by Vos et al. (1995), with minor adaptations for quinoa DNA. An aliquot of genomic DNA (500 ng) from each F_2 plant and parent were digested with the restriction enzymes *Mse*I and *Eco*RI. Digested fragments were ligated with *Eco*RI and *Mse*I adapters. A preamplification reaction was performed with primers complementary for each adapter having one selective nucleotide each, specifically *Eco*RI adapter+A and *Mse*I adapter+C. The preamplification product was diluted 20-fold in 1/10 TE, and 2 µl was used in the final selective amplification step. Prior to the selective amplification step, the *Eco*RI selective primer was end-labeled with γ–[^{33}P] in a T4 polynucleotide kinase reaction as recommended by the manufacturer (New England Biolabs, Beverly, Mass., USA). The selective primers were complementary to the adaptor sequence and included three selective nucleotides. The selective amplifications protocol consisted of one cycle of 94°C for 30 s, 65°C for 30 s, and 72°C for 60 s. Then annealing temperatures were lowered 1°C for each of the nine cycles, followed by 26 cycles of 94°C for 30 s, 56°C for 30 s, and 72°C for 60 s. PCR products (2 µl) were separated on 5% denaturing polyacrylamide gels with 7.5 M urea at 110 W constant power on a Sequi-Gen 3000 DNA sequencing unit (Bio-Rad Laboratories, Hercules, Calif., USA) Gels were transferred to 3 MM paper, covered with plastic wrap, dried, and exposed to phosphor-imaging screens for 12 h. Imaging screens were scanned with a Bio-Rad molecular imager FX (Bio-Rad Laboratories) and saved as digital images for band analysis. Polymorphic bands were named according to the selective primers used to amplify them and the size (in base pairs) of the fragment scored in the polyacrylamide gel (e.g., EAAC/MCAT-240; see also Fig. 1).

Simple-sequence repeat markers

Thirty-nine putative simple-sequence repeat (SSR) loci, previously identified from an SSR-enriched genomic

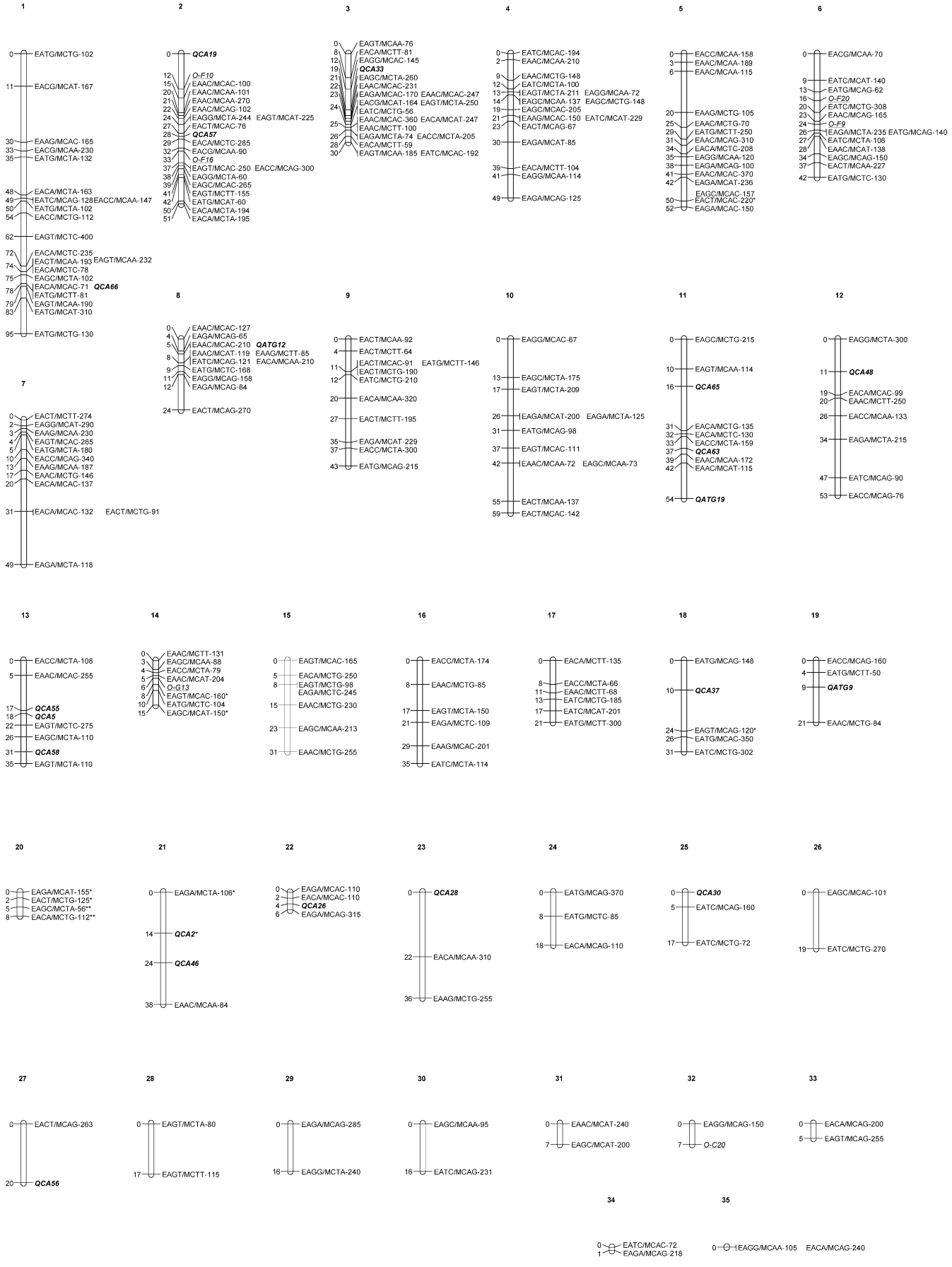
1
0 EATG/MCTG-102
11 EACG/MCAT-167
30 EAAG/MCAC-165
33 EACG/MCAA-230
35 EATG/MCTA-132
48 EACA/MCTA-163
49 EATC/MCAG-128 EACC/MCAA-147
50 EATG/MCTA-102
54 EACC/MCTG-112
62 EAGT/MCTC-400
72 EACA/MCTC-235
74 EACT/MCAA-193 EAGT/MCAA-232
EACA/MCTC-78
75 EAGC/MCTA-102
78 EACA/MCAC-71 QCA66
79 EATG/MCTT-81
83 EAGT/MCAA-190
EATG/MCAT-310
95 EATG/MCTG-130
2
0 QCA19
12 O-F10
15 EAAC/MCAC-100
20 EAAC/MCAA-101
21 EAAC/MCAA-270
22 EAAC/MCAG-102
24 EAGG/MCTA-244 EAGT/MCAT-225
27 EACT/MCAC-76
28 QCA57
29 EACA/MCTC-285
32 EACG/MCAA-90
33 O-F16
37 EAGT/MCAC-250 EACC/MCAG-300
38 EAGG/MCTA-60
39 EAGC/MCAC-265
41 EAGT/MCTT-155
42 EATG/MCAT-60
50 EACA/MCTA-194
51 EACA/MCTA-195
3
0 EAGT/MCAA-76
8 EACA/MCTT-81
12 EAGG/MCAC-145
19 QCA33
21 EAGC/MCTA-260
22 EAAC/MCAC-231
23 EAGA/MCAC-170 EAAC/MCAC-247
24 EACG/MCAT-164 EAGT/MCTA-250
EATC/MCTG-56
25 EAAC/MCAC-360 EACA/MCAT-247
EAAC/MCTT-100
26 EAGA/MCTA-74 EACC/MCTA-205
28 EACA/MCTT-59
30 EAGT/MCAA-185 EATC/MCAC-192
4
0 EATC/MCAC-194
2 EAAC/MCAA-210
9 EAAC/MCTG-148
12 EATC/MCTA-100
13 EAGT/MCTA-211 EAGG/MCAA-72
14 EAGC/MCAA-137 EAGC/MCTG-148
19 EAGC/MCAC-205
21 EAAG/MCAC-150 EATC/MCAT-229
23 EACT/MCAG-67
30 EAGA/MCAT-85
39 EACA/MCTT-104
41 EAGG/MCAA-114
49 EAGA/MCAG-125
5
0 EACC/MCAA-158
3 EAAC/MCAA-189
6 EAAC/MCAA-115
20 EAAG/MCTG-105
25 EAAC/MCTG-70
29 EATG/MCTT-250
31 EAAC/MCAG-310
34 EACA/MCTC-208
35 EAGG/MCAA-120
38 EAGA/MCAG-100
41 EAAC/MCAC-370
42 EAGA/MCAT-236
EAGC/MCAC-157
50 EACT/MCAC-220*
52 EAGA/MCAC-150
6
0 EACG/MCAA-70
9 EATC/MCAT-140
13 EATG/MCAG-62
16 O-F20
20 EATC/MCTG-308
23 EAAC/MCAG-165
24 O-F9
26 EAGA/MCTA-235 EATG/MCAG-140
27 EATC/MCTA-108
28 EAAC/MCAT-138
34 EAGC/MCAG-150
37 EACT/MCAA-227
42 EATG/MCTC-130
7
0 EACT/MCTT-274
2 EAGG/MCAT-290
3 EAAG/MCAA-230
4 EAGT/MCAC-265
5 EATG/MCTA-180
10 EACC/MCAG-340
13 EAAG/MCAA-187
17 EAAC/MCTG-146
20 EACA/MCAC-137
31 EACA/MCAC-132 EACT/MCTG-91
49 EAGA/MCTA-118
8
0 EAAC/MCAC-127
4 EAGA/MCAG-65
5 EAAC/MCAC-210 QATG12
EAAC/MCAT-119 EAAG/MCTT-85
8 EATC/MCAG-121 EACA/MCAA-210
9 EATG/MCTC-168
11 EAGG/MCAG-158
12 EAGA/MCAG-84
24 EACT/MCAG-270
9
0 EACT/MCAA-92
4 EACT/MCTT-64
11 EACT/MCAC-91 EATG/MCTT-146
EACT/MCTG-190
12 EATC/MCTG-210
20 EACA/MCAA-320
27 EACT/MCTT-195
35 EAGA/MCAT-229
37 EACC/MCTA-300
43 EATG/MCAG-215
10
0 EAGG/MCAC-67
13 EAGC/MCTA-175
17 EAGT/MCTA-209
26 EAGA/MCAT-200 EAGA/MCTA-125
31 EATG/MCAG-98
37 EAGT/MCAC-111
42 EAAC/MCAA-72 EAGC/MCAA-73
55 EACT/MCAA-137
59 EACT/MCAC-142
11
0 EAGC/MCTG-215
10 EAGT/MCAA-114
16 QCA65
31 EACA/MCTG-135
32 EACA/MCTC-130
33 EACC/MCTA-159
37 QCA63
39 EAAC/MCAA-172
42 EAAC/MCAT-115
54 QATG19
12
0 EAGG/MCTA-300
11 QCA48
19 EACA/MCAC-99
20 EAAC/MCTT-250
26 EACC/MCAA-133
34 EAGA/MCTA-215
47 EATC/MCAG-90
53 EACC/MCAG-76
13
0 EACC/MCTA-108
5 EAAC/MCAC-255
17 QCA55
18 QCA5
22 EAGT/MCTC-275
26 EAGC/MCTA-110
31 QCA58
35 EAGT/MCTA-110
14
0 EAAC/MCTT-131
3 EAGC/MCAA-88
4 EACC/MCTA-79
5 EAAC/MCAT-204
6 O-G13
8 EAGT/MCAC-160*
10 EATG/MCTC-104
15 EAGC/MCAT-150*
15
0 EAGT/MCAC-165
5 EACA/MCTG-250
8 EAGT/MCTG-98
EAGA/MCTC-245
15 EAAC/MCTG-230
23 EAGC/MCAA-213
31 EAAC/MCTG-255
16
0 EACC/MCTA-174
8 EAAC/MCTG-85
17 EAGT/MCTA-150
21 EAGA/MCTC-109
29 EAAG/MCAC-201
35 EATC/MCTA-114
17
0 EACA/MCTT-135
8 EACC/MCTA-66
11 EAAC/MCTT-68
13 EATC/MCTG-185
17 EATC/MCAT-201
21 EATG/MCTT-300
18
0 EATG/MCAG-148
10 QCA37
24 EAGT/MCAG-120*
26 EATG/MCAC-350
31 EATC/MCTG-302
19
0 EACC/MCAG-160
4 EATG/MCTT-50
9 QATG9
21 EAAC/MCTG-84
20
0 EAGA/MCAT-155*
2 EACT/MCTG-125*
5 EAGC/MCTA-56**
8 EACA/MCTG-112**
21
0 EAGA/MCTA-106*
14 QCA2*
24 QCA46
38 EAAC/MCAA-84
22
0 EAGA/MCAC-110
2 EACA/MCAC-110
4 QCA26
6 EAGA/MCAG-315
23
0 QCA28
22 EACA/MCAA-310
36 EAAG/MCTG-255
24
0 EATG/MCAG-370
8 EATG/MCTC-85
18 EACA/MCAG-110
25
0 QCA30
5 EATC/MCAG-160
17 EATC/MCTG-72
26
0 EAGC/MCAC-101
19 EATC/MCTG-270
27
0 EACT/MCAG-263
20 QCA56
28
0 EAGT/MCTA-80
17 EAGT/MCTT-115
29
0 EAGA/MCAG-285
16 EAGG/MCTA-240
30
0 EAGC/MCAA-95
16 EATC/MCAG-231
31
0 EAAC/MCAT-240
7 EAGC/MCAT-200
32
0 EAGG/MCAG-150
7 O-C20
33
0 EACA/MCAG-200
5 EAGT/MCAG-255
34
0 EATC/MCAC-72
1 EAGA/MCAG-218
35
0 EAGG/MCAA-105 EACA/MCAG-240

Table 1 Microsatellite primer sequences, core motif, GenBank accession number, amplified fragment length, scoring method, and average melting temperature

Primer name	Core motif	GenBank accession number	Forward primer sequence (5′→3′)	Reverse primer sequence (5′→3′)	Predicted product size (bp)	Scoring method	Annealing temperature (°C)
QATG12	(ATC)8	ay458232	ggcaaacacatggatctcaa	caaacagcgaagaggtttcc	182	Dominant	59.5
QATG19	(ATC)12	ay458233	ccaaacaaagacaataaggaaacc	cgaggttgaaggagattcca	186	Codominant	60
QATG16	(CAT)10	ay458234	ccacactcatctcaaccatcc	gcgatccggtacatcatttc	170	Codominant	60
QATG9	(GAT)5	ay458235	gactaatgaacggatggaagc	cctcccacataaccttcctct	150	Dominant	59
QCA19	(CA)10	ay458236	tttcatcactcgaccgtatagc	agggtgactgttacacccaaa	181	Dominant	59
QCA2	(CA)20	ay458231	tgaggacacacacacacacac	gggaccgtttaattcagcaa	160	Codominant	59
QCA26	(TG)12	ay458230	ttccaatacagcaccacctc	tgcaagcatacataagacagtca	187	Dominant	58.5
QCA28	(CA)10	ay458237	tgctcaccctagcatttatacact	atgagacggagggagcacta	155	Codominant	59
QCA30	(CA)13	ay458238	tcattggttagatggtggaatg	ccctctagtgcataggagtttctg	177	Codominant	60
QCA33	(CA)10	ay452266	cagggcagtccacctctcta	accttctagtcctatgttcttgtatgg	219	Codominant	59
QCA37	(CA)16	ay458227	ccgttcttccagaccaattc	tcatgagccacttcatacacg	188	Codominant	60
QCA46	(CA)18	ay458239	gcaggtaaatcaacccttgc	tgcatgataaactaagcagacga	165	Codominant	60
QCA48	(CA)13	ay458240	acaatacatacataacccaatattcaa	tggaaatgtcactatgattgga	235	Codominant	58
QCA51	(TG) 10	ay458241	tgtttcgggtagaaacaccaa	tgcaattcaatgcccacata	156	Dominant	60
QCA5	(CA)16	ay458228	gtggttcatggctgatcctt	cttgccatcagggcatatct	185	Codominant	60
QCA56	(CA)12	ay458242	ttggaagagctccacaaggt	cctctgaataggatacccttctgt	172	Codominant	59
QCA57	(CA)22	ay458243	tgcaaggaaaccatctttgg	tgcctcacagtcacacctaca	163	Codominant	60
QCA58	(TG)16	ay458244	ctcgaccagcagggtctg	ctagctaggcgttgcctgac	183	Dominant	60
QCA63	(CA)11	ay458245	gcagcatcacagagcagaaa	ggtgtagattgggagcctga	184	Codominant	60
QCA65	(CA)19	ay458246	ccatgcaagggaacatattg	aagttcgttggcttgctgta	199	Dominant	59
QCA66	(CA)32	ay458247	agagttcttacataagggaagagt	tttcctttggtagtttcttgtt	176	Dominant	55

library (P.J. Maughan et al., unpublished data) were screened for polymorphism in the parents of this cross. Twenty-one polymorphic SSR loci (Table 1) were amplified with DNA from individuals of the F_2 population, according to standard PCR conditions as described by Yu et al. (1994) at 2.5 mM $MgCl_2$. All SSR amplifications were performed with a touchdown amplification protocol as follows: 94°C for 1 min, followed by five cycles of 94°C for 30 s, 55°C for 30 s (decreasing 1°C every cycle), and 72°C for 1 min; ten cycles of 94°C for 30 s, 50°C for 30 s, and 72°C for 1 min; five cycles of 94°C for 30 s, 50°C for 30 s (decreasing 1°C every cycle), and 72°C for 1 min; ten cycles of 94°C for 30 s, 45°C for 30 s, and 72°C for 1 min; and hold at 72°C for 5 min. PCR products were separated on 3% Metaphor agarose gels (Cambrex Bio Science, East Rutherford, N.J., USA), run in 0.5× TBE. PCR products were visualized by ethidium-bromide staining and UV transillumination.

◀ **Fig. 1** Linkage map of *Chenopodium quinoa*. Cumulative distances are in centiMorgans and are indicated on the *left side* of the linkage groups. Microsatellite markers are in *boldface*. RAPD markers are *underlined*. All other markers are amplified length polymorphism (AFLP) markers. AFLP markers are named as described in Materials and methods. Loci marked with * or ** showed significant distortion a $P<0.05$ and $P<0.01$, respectively

RAPD procedures

A total of 60 random decamer primers (primer sets C, F and G, Operon, Alameda, Calif., USA) were screened for polymorphism between the parents of this cross, according to procedures described by Williams et al. (1990) with modifications as described by Ruas et al. (1999). RAPD markers were given the prefix "O-," followed by the Operon primer set and number. DNA amplification was performed in a total volume of 15 µl. Amplified products were separated on a 1.4% agarose gel run in 0.5× TE buffer and visualized by ethidium-bromide staining and UV transillumination.

Data analysis and map construction

The majority of AFLP and RAPD markers were scored as dominant markers, whereas the majority of SSR markers were scored as codominant markers. Only unambiguous markers were scored. Genetic similarities, based on simple matching coefficients, were analyzed phenetically with the computer program NTSYS-pc, version 2.1 (Rohlf 2000). Linkage groups and map construction were accomplished with the computer program JoinMap, version 3.0 (van Ooijen and Voorrips 2001). Linkage groups were determined as markers linked with a minimum LOD score of 3.5. The order of the markers of each linkage group was determined by JoinMap's default parameters (LOD>1.0, recombination threshold = 0.4, ripple value =

Table 2 Similarity matrix based on simple matching coefficients of potential parents

	Chucapaca	NL6	0654	Ku-2
Chucapaca	1.000			
NL6	0.245	1.000		
0654	0.576	0.327	1.000	
Ku-2	0.229	0.866	0.304	1.000

1, jump threshold = 5, Kosambi mapping function) for all linkage groups.

Results and discussion

Population selection

Eighty-eight AFLP primer combinations were screened for polymorphism among the four potential mapping parents ('Ku-2', 'NL-6', '0654', and 'Chucapaca'). Analysis of the AFLP gels for all primer combinations identified a total of 597 polymorphic bands across the four potential parents. The average number of bands detected for individual primer pairs ranged from a low of 19 to a high of 52, with an average of 6.8 polymorphic bands per primer combination. The similarity coefficients (Table 2) ranged from 0.23 to 0.87, with the least genetic similarity between the Bolivian accession 'Chucapaca' and the Chilean accession 'Ku-2', and the greatest similarity between the two Chilean coastal accessions 'NL-6' and 'Ku-2'. These findings agreed well with the previous morphological and isozyme studies, which separated the quinoa germplasm into two distinct fundamental elements, Chilean coastal types and Andean Altiplano types (Wilson 1988; Risi and Galwey 1989). In this manuscript we report the mapping results of the cross of 'Ku-2', a Chilean accession, and '0654', an Altiplano type from Peru.

Marker segregation

Of the 88 AFLP primers combinations screened for polymorphism between the mapping parents, 68 were polymorphic and highly reproducible based on duplicated samples. These 68 primer combinations produced a total of 4,280 amplification products, of which 284 (6.6%) were polymorphic, easily scored, and reproducible on the F_2 population. The number of polymorphic markers per primer combination segregating in the F_2 population ranged from a low of one to a high of nine. On average, 4.2 polymorphic markers were scored per primer combination and 3.4 polymorphic markers per primer combination were placed on the map. Twenty-six (8%) of the markers were identified and scored as segregating codominantly. Of the dominant AFLP markers, 160 (53%) were specific to '0654' and 144 (47%) were specific to 'Ku-2'. All AFLP markers were tested for segregation distortion from the expected Mendelian ratio (3:1 or 1:2:1) with the chi-square test. Out of the 284 AFLP markers tested, 12 (4.2%) loci showed distorted segregation at $P<0.05$, and two (0.7%) loci showed distortion at $P<0.01$, fully within expectations for random sampling error. Thus, there was no evidence of statistically significant deviation from a disomic segregation model in the data set. Interestingly, Ward (2000) reported erratic multivalent pairing and tetrasomic inheritance for two monogenic morphological traits in quinoa. Indeed, tetrasomic segregation of "erratic multivalents" (Ward 2000) would lead to distorted segregation ratios and may also contribute to the low levels of segregation distortion seen in our data set. The low level of segregation distortion for AFLP markers observed in this study is comparable to that observed in allotetraploid mapping experiments with AFLP markers in cotton (Lacape et al. 2003), durum wheat (Lotti et al. 2000), and *Brassica juncea* (Pradhan et al. 2003) and suggests that the AFLP markers are heritable and should be reliable for genetic analysis in *C. quinoa*.

From the 21 SSR markers identified as polymorphic in this population, 13 produced simple monogenic banding patterns and were easily scored in a codominant fashion (Table 1). Most of the other SSR markers yielded amplification products with complex banding patterns that made scoring the marker in a codominant fashion problematic. In these markers, a single, unambiguous, and clearly segregating band was scored in a dominant fashion. One possible explanation for these complex banding patterns is the presence of duplicate chromosome regions (Rae et al. 2000), a relic of quinoa's probable allotetraploid origin. Indeed, our laboratory has identified several SSR markers (not utilized in this study) that clearly amplify two codominantly segregating loci, presumably orthologous loci from each of the ancestral genomes of quinoa (P.J. Maughan et al., unpublished data). Such markers will be valuable tools in identifying the putative diploid ancestors of *C. quinoa*. Of the SSR markers tested, only QCA2 deviated significantly ($P<0.05$) from its expected Mendelian ratio (1:2:1).

Of the 60 RAPD primers screened, six (10%) produced reproducible polymorphic markers and were included in the linkage analysis. One polymorphic marker (band) was scored from each of the six RAPD primers, with an average of 3.8 prominent bands per RAPD reaction. One RAPD marker (O-F10) was scored in a codominant fashion, while the remaining five RAPD markers were scored as dominant markers. None of the RAPD markers deviated significantly ($P>0.05$) from their expected segregation ratios. Plant color was scored as a dominant phenotypic marker for each F_2 plant [red vs green (*R/r*)] and did not deviate from expected segregated ratios ($P>0.05$).

In this study, AFLP markers were shown to be reliable and efficient in generating a linkage map of quinoa. SSR markers were very informative and reliable, but less efficient than AFLP markers and often produced complex dominant and/or multilocus banding patterns which detracted from their utility for genetic mapping. Jeuken et al. (2001) described a similar situation with SSR

markers in an interspecific cross of lettuce (*Lactuca sativa* × *L. saligna*). In that study only 4 out of 12 SSRs could be scored codominantly. RAPD markers, while attractive due to their technical simplicity, typically are less efficient and prone to reproducibility issues.

Linkage analysis

Genetic segregation in the F_2 population was analyzed for a total of 313 loci, including 284 AFLP loci, 21 SSR loci, six RAPD loci, and one morphological locus. Of the 313 loci, 275 (88%) were grouped at a minimum LOD score of 3.5. Thirty-eight loci remained unassigned, including one SSR marker (QATG16) and the plant-color locus (*R/r*). Ordering of markers within groups identified an additional 19 loci (18 AFLP and one SSR marker) that showed suspect linkages (REC value >0.60) and thus were removed from the final groups. Unassigned markers were not unexpected and have been found in many previous genome-mapping experiments (Atienza et al. 2002; Sharma et al. 2002). Thus, the final map consists of a total of 230 AFLP, 19 SSR, and six RAPD markers; spans 1,020 cM; and contains 35 linkage groups. The linkage groups were temporarily numbered from highest to lowest based on the number of linked markers, pending cytogenetic confirmation. The number of markers per linkage group ranged from a low of two (0.8%) to a high of 22 (8.6%; linkage group 1). Linkage distance spanned by individual linkage groups ranged from a low of 0 cM (linkage group 35) to a high of 95 cM (linkage group 1). Out of the 35 groups, 25 contained at least three linked markers. The average linkage distance between pairs of markers among all groups was 4.5 cM, considering only one marker in regions with cosegregating markers, and 4.0 cM, taking all markers into account. The largest gap (22.5 cM) is located on linkage group 23. Eight additional gaps (>15 cM) are located on linkage groups 1, 7, 11, and 26–30. All of these intervals have LOD score values above the threshold LOD. The vast majority of intervals (86%) were less than 10 cM in length.

The number of linkage groups reported in this study is higher than the haploid chromosome number of *C. quinoa* (n=18). However, many of the smaller linkage groups will undoubtedly coalesce with other linkage groups when additional markers are added to the map. The requirement for large numbers of markers or integrated mapping populations to reduce the linkage group number to haploid chromosome numbers has been seen in many mapping experiments (Jeuken et al. 2001; Sharma et al. 2002). The excess of linkage groups when compared to the haploid chromosome number suggests that several areas of the genome remain undetected with this set of markers. Indeed, employing the method of Hulbert et al. (1988), we predict the total length of the quinoa map to be approximately 1,700 cM. Thus, this map covers approximately 60% of the total quinoa genome. Of the species related to quinoa, sugar beet (*Beta vulgaris* L., $2n$=$2x$=18) has the best-characterized genetic map. Sugar beet is a true diploid member of the Amaranthaceae family and has a haploid chromosome number of nine (half that of quinoa). The total length of the several dense linkage maps reported in sugar beet range from 621 cM (Halldén et al. 1996) to 815 cM (Barzen et al. 1995), roughly half the size predicted for the total length (1,700 cM) of the quinoa genetic map and between 61% and 80% of the genetic map we report for quinoa (1,020 cM). Considering the probable allotetraploid origin of quinoa, a genetic map roughly twice the size of those for diploid members of the family is not unexpected.

Although AFLP markers were used as the backbone of the quinoa linkage map, several sequence-tagged SSR markers were developed and added to the map. These 19 SSR markers are distributed across 14 of the quinoa linkage groups. BLASTN and BLASTX searches of the GenBank database revealed that four of the mapped SSR markers had significant DNA sequence similarity to other plant genes. QCA5 contains the entire sequence of exon 13 of the NADP-specific isocitrate-dehydrogenase gene, and intron sequences on both ends of the exon. QCA26 contains part of exon 11 and part of intron 11 of a gene homologous to the *Arabidopsis thaliana At2 g47210* gene, which encodes a hypothetical protein. QCA37 contains part of exon 5 and part of intron 5 of a gene homologous to the *A. thaliana F6I1.13* gene, which encodes an unknown protein. QCA2 is significantly similar to the *A. thaliana MVI1.11* gene, which encodes an unknown protein. Interestingly, the exon/intron boundaries in the quinoa QCA2 sequence differs from those in the proposed *Arabidopsis* sequence in that an apparent intron in the quinoa sequence spans, but lacks homology to, a region labeled as part of an exon in the *Arabidopsis* sequence. GenBank sequence accession numbers for all of the polymorphic SSR markers are provided in Table 1.

The SSR markers reported in this study constitute an initial set of standard landmarks that can be easily transferred from population to population and can be used as anchor points for future map comparison. Moreover, these SSR markers will be valuable in characterizing quinoa germplasm as well as elucidating the evolutionary relationship and the possibility of genetic exchange between quinoa and the other cultivated *Chenopodium* species, including *C. berlandieri* ssp. *Nutalliae* (cultivated in Mexico, commonly referred to as Hauzontle) and *C. pallidicaule* (cultivated in Bolivia and Peru, commonly referred to as Cañihua).

Nuclear DNA content

The nuclear content estimated by flow cytometry indicates that *C. quinoa* 'Surimi' has a total haploid nucleus DNA content of 967 Mbp/1C (K. Arumuganathan, personal communication). This is ~6.7 times the size of *A. thaliana* and approximately two times the size of rice, but slightly larger than other member of the Amaranthaceae family, including beet (714 Mbp/1C), sugar beet (758 Mbp/1C), and spinach (989 Mbp/1C) (Arumuganathan and Earle

1991). Based on this estimate, we calculate 1 cM in quinoa genetic map is approximately 568 kb. However, caution must be used when relating genetic distance to physical distance as the relationship is usually non-linear, even over short distances within the same chromosome (Lashermes et al. 2001).

AFLP marker clustering

Non-random distribution and clustering of markers at the centromeres and surrounding heterochromatic regions of the genome has been reported by several researchers using AFLP markers based on the restriction endonucleases *Mse*I and *Eco*RI (Keim et al. 1997; Qi et al. 1998; Vuylsteke et al. 1999). Extensive clustering was not apparent in this genetic map of quinoa (Fig. 1). The lack of noticeably clustered markers may simply be an artifact of the relatively low number of AFLP markers (230) on the map. Several of the linkage groups show regions of higher marker density (linkage groups 1, 3, 4, and 8) and may, with additional markers, form more pronounced clusters. Although clustering of AFLP markers is a general characteristic, several researchers have reported the uniform distribution of AFLP markers for several crop species (Jeuken et al. 2001; Vilanova et al. 2003). Jeuken et al. (2001), suggested that if clustering is a result of variants in repetitive elements at the centromeres, the lack of clustering in some species may be the result of a simpler centromeric repeat patterns or that the centromeres of in these species could be of reduced size compared to the other species, leading to smaller regions with suppressed recombination rates.

Conclusions

We report the use of AFLP, SSR, and RAPD markers to construct the first genetic linkage map of *C. quinoa*, an important under-researched food crop of the Andean region of South America. The development of this genetic linkage map is an important step towards the genetic characterization and initiation of MAS for recalcitrant agronomic traits in quinoa. Such qualitative traits as resistance to downy mildew (*Peronospora farinosa* f.sp. *chenopodii*), the most important disease of quinoa, and saponin-free quinoa (both major quinoa breeding objectives) should be readily amenable to genetic mapping studies and MAS protocols.

In conjunction with building a genetic map, we have developed several sequence-tagged SSR markers. In contrast to AFLP markers, the low cost, ease of use, highly polymorphic nature, and abundance of SSR markers make them especially useful as a marker system for the analysis of under-researched crops in developing countries. Once developed, SSRs are easily maintained and shared among laboratories through the publication of primer-pair sequences. These SSRs will undoubtedly be valuable tools for integrating future genetic maps, understanding genome organization and elucidating genetic relationships within the genus *Chenopodium*.

Acknowledgements This research was supported with grants from the McKnight Foundation, the Holmes Family Foundation, and the Erza Taft Benson Agriculture and Food Institute. We thank Robert Jones, Eric Steward, and David Butler of Brigham Young University for their advice and technical assistance with the AFLP analysis.

References

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:208–218

Atienza SG, Satovia Z, Petersen KK, Dolstra O, Martin A (2002) Preliminary genetic linkage map of *Miscanthus sinensis* with RAPD markers. Theor Appl Genet 105:946–952

Barzen E, Mechelke W, Ritter E, Schulte-Kappert E, Salamini F (1995) An extended map of the sugar beet genome containing RFLP and RAPD loci. Theor Appl Genet 89:167–178

Chauhan GS, Eskin NAM, Tkachuk R (1992) Nutrients and anitnutrients in quinoa seed. Cereal Chem 69:85–88

Chauhan GS, Eskin NAM, Tkachuk R (1999) Effect of saponin extraction on quality of quinoa (*Chenopodium quinoa* Willd.). J Food Sci Technol 36:123–126

Diwan N, McIntosh MS, Bauchan GR (1995) Methods of developing a core collection of annual *Medicago* species. Theor Appl Genet 90:775–761

Fleming JE, Galwey NW (1995) Quinoa (*Chenopodium quinoa*). In: Williams JT (ed) Cereals and pseudocereals. Chapman and Hall, London, pp 3–83

Halldén C, Hjerdin A, Rading IM, Säll T, Fridlundh B, Johannisdottir G, Tuvesson S, Åkesson C, Nilsson N-O (1996) A high-density RFLP linkage map of sugar beet. Genome 39:634–645

Hulbert SH, Ilott TW, Legg EJ, Lincoln SE, Lander ES, Michelmore RW (1988) Genetic analysis of the fungus, *Bremia lactucae*, using restriction fragment length polymorphisms. Genetics 120:947–958

Jeuken M, van Wijk R, Peleman J, Lindhout P (2001) An integrated interspecific AFLP map of lettuce (*Lactuca*) based on two *L. sativa* × *L. saligna* F_2 populations. Theor Appl Genet 103:638–647

Keim P, Schupp JM, Travis SE, Clayton K, Zhu T, Shi L, Ferreira A, Webb DM (1997) A high-density soybean genetic map based on AFLP markers. Crop Sci 37:537–543

Lacape JM, Nguyen TB, Thibivilliers S, Bojinov B, Courtois B, Cantrell RG, Burr B, Hau B (2003) A combined RFLP–SSR–AFLP map of tetraploid cotton based on a *Gossypium hirsutum* × *Gossypium barbadense* backcross population. Genome 46:612–626

Lashermes P, Combes MC, Prakash NS, Trouslot P, Lorieux M, Charrier A (2001) Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meioses. Genome 44:589–596

Lotti C, Salvi S, Pasqualone A, Tuberosa R, Blanco A (2000) Integration of AFLP markers into an RFLP-based map of durum wheat. Plant Breed 119:393–401

Ochoa J, Frinking HD, Jacobs TH (1999) Postulation of virulence groups and resistance factors in the quinoa/downy mildew pathosystem using material from Ecuador. Plant Pathol 48:425–430

Ooijen JW van, Voorrips RE (2001) JoinMap 3.0, software for the calculation of genetic linkage maps. Plant Research International, Wageningen

Patterson AH, Tanksley SD, Sorrells ME (1991) DNA markers in plant improvement. Adv Agron 46:39–90

Pradhan AK, Gupta V, Mukhopadhyay A, Arumugam N, Sodhi YS, Pental D (2003) a high-density linkage map in *Brassica juncea* (Indian mustard) using AFLP and RFLP markers. Theor Appl Genet 106:607–614

Prado RE, Boero C, Gallard M, Gonzalez JA (2000) Effect of NaCl on germination, growth, and soluble sugar content in *Chenopodium quinoa* Willd. seeds. Bot Bull Acad Sci 41:27–34

Qi X, Stam P Lindhout P (1998) Use of locus-specific AFLP markers to construct a high-density molecular map in barley. Theor Appl Genet 96:376–384

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

Risi J, Galwey NW (1984) The *Chenopodium* grains of the Andes: Inca crops for modern agriculture. Adv Appl Biol 10:145–216

Risi J, Galwey NW (1989) The pattern of genetic diversity in the Andean grain crop quinoa (*Chenopodium quinoa* Willd.). I. Associations between characteristics. Euphytica 41:147–162

Rohlf FJ (2000) NTSYSpc 2.1: numerical taxonomy and multivariate analysis system. Exeter Software, New York

Ruales J, Nair BM (1993) Content of fat, vitamins and minerals in quinoa (*Chenopodium quinoa* Willd.) seeds. Food Chem 48:131–136

Ruas PM, Bonifacio A, Ruas CF, Fairbanks DJ, Andersen WR (1999) Genetic relationship among 19 accessions of six species of *Chenopodium* L., by randomly amplified polymorphic DNA fragments (RAPD). Euphytica 105:25–32

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

Sharma R, Aggarwal RAK, Kumar R, Mohapatra T, Sharma RP (2002) Construction of an RAPD linkage map and localization of QTLs for oleic acid level using recombinant inbreds in mustard (*Brassica juncea*). Genome 45:467–472

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

Staub JE, Serquen FC, Gupta M (1996) Genetic markers, map construction, and their application in plant breeding. HortScience 31:729–741

Tanksley SD, McCouch SR (1997) Seed bands and molecular maps: unlocking genetic potential from the wild. Science 277:1063–1066

Tapia M, Gandarillas H, Alandia S, Cardozo A, Mujica R, Ortiz R, Otazu J, Rea J, Salas B, Zanabria E (1979) Quinua y kañiwa: Cultivos andinos. CIID-IICA. Bogota

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

Vacher JJ (1998) Responses of two main Andean crops, quinoa (*Chenopodium quinoa* Willd.) and papa amarga (*Solanum juzepezukii* Buk.) to drought on the Bolivian Altiplano: significance of local adaption. Agric Ecosyst Environ 68:99–108

Vilanova S, Romero C, Abbott AG, Llácer, Badenes ML (2003) An apricot (*Prunus armeniaca* L.) F_2 progeny linkage map based SSR and AFLP markers, mapping plum pox virus resistance and self-incompatibility traits. Theor Appl Genet 107:239–247

Vos P, Hogers R, Bleeker R, Reijans M, Van de Le T, HornesM, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Vuylsteke M, Mank R, Antonise R, Bastiaans E, Senior ML, Stuber CW, Melchinger AE, Lubbersted T, Xia XC, Stam P, Zabeau M, Kuiper M (1999) Two high-density AFLP linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. Theor Appl Genet 99:921–935

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

Ward SM (2001) A recessive allele inhibiting saponin synthesis in two lines of Bolivian quinoa (*Chenopodium quinoa* Willd.). Heredity 92:83–86

Willams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

Wilson HD (1988) Quinoa biosystematics I: domesticated populations. Econ Bot 42:461–477

Wood SG, Lawson LD, Fairbanks DJ, Robison LR, Andersen WR (1993) Seed lipid content and fatty acid composition of three quinoa cultivars. J Food Comp Anal 6:41–44

Yu YG, Saghai Maroof MA, Buss GR, Maughan PJ, Tolin SA (1994) RFLP and microsatellite mapping of a gene for soybean mosaic virus resistance. Phytopathology 84:60–64