RESEARCH ARTICLE

Simple sequence repeat marker development and genetic mapping in quinoa (*Chenopodium quinoa* Willd.)

D. E. JARVIS[1], O. R. KOPP[2], E. N. JELLEN[1], M. A. MALLORY[1], J. PATTEE[2], A. BONIFACIO[3], C. E. COLEMAN[1], M. R. STEVENS[1], D. J. FAIRBANKS[1] and P. J. MAUGHAN[1,*]

[1]*Brigham Young University, Department of Plant and Animal Sciences, Provo, UT 84602, USA*
[2]*Utah Valley State University, Biology Department, Orem, UT 84058, USA*
[3]*Fundacion PROINPA Casilla Postal 4285, Cochabamba, Bolivia*

Abstract

Quinoa is a regionally important grain crop in the Andean region of South America. Recently quinoa has gained international attention for its high nutritional value and tolerances of extreme abiotic stresses. DNA markers and linkage maps are important tools for germplasm conservation and crop improvement programmes. Here we report the development of 216 new polymorphic SSR (simple sequence repeats) markers from libraries enriched for GA, CAA and AAT repeats, as well as 6 SSR markers developed from bacterial artificial chromosome-end sequences (BES-SSRs). Heterozygosity (H) values of the SSR markers ranges from 0.12 to 0.90, with an average value of 0.57. A linkage map was constructed for a newly developed recombinant inbred lines (RIL) population using these SSR markers. Additional markers, including amplified fragment length polymorphisms (AFLPs), two 11S seed storage protein loci, and the nucleolar organizing region (NOR), were also placed on the linkage map. The linkage map presented here is the first SSR-based map in quinoa and contains 275 markers, including 200 SSR. The map consists of 38 linkage groups (LGs) covering 913 cM. Segregation distortion was observed in the mapping population for several marker loci, indicating possible chromosomal regions associated with selection or gametophytic lethality. As this map is based primarily on simple and easily-transferable SSR markers, it will be particularly valuable for research in laboratories in Andean regions of South America.

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

Introduction

Quinoa (*Chenopodium quinoa* Willd.) is an allotetraploid ($2n = 4x = 36$) that shows disomic inheritance for most-qualitative traits (Simmonds 1971; Risi and Galwey 1984; Ward 2000; Maughan *et al*. 2004). It is an important South American pseudo-cereal crop that has recently gained international attention for the high nutritional value of its grain (Jacobsen 2003). Quinoa is grown primarily in the altiplano regions of Bolivia, Ecuador, Chile and Peru. It has served as an important staple crop for subsistence farmers for thousands of years (Pearsall 1992; Wilson 1988). It is well suited as a staple crop in the altiplano due to its high protein content (7.5–22.1%; Tapia *et al*. 1979) as well as its ability to grow in the harsh environments that characterize much of the altiplano, including high altitudes (up to 4000 m), frequent frosts and saline soils (Risi and Galwey 1984; Vacher 1998; Prado *et al*. 2000; Jacobsen 2003).

Despite of its many desirable nutritional and abiotic stress characteristics, quinoa is plagued by a number of biotic stressors. Significant quinoa diseases includes bacterial stem rot and downy mildew (Danielsen *et al*. 2003). Quinoa is also affected by avian, arthropod and nematode pests (Franco 2003; Rasmussen *et al*. 2003), all of which reduce the grain yield. Thus, a major breeding objective for quinoa includes the development of disease resistant, high-yielding varieties. Unlike the other major cereal crops, which have benefited greatly from the modern plant breeding techniques and genetic research, genetic improvement of quinoa has, until recently, received limited attention. Indeed, since from past 10 years only breeding programs for

*For correspondence. E-mail: Jeff_Maughan@byu.edu.

Keywords. microsatellites; heterozygosity; genetic map; NOR; 11S.

quinoa have been established in many Andean countries. Development of appropriate molecular tools to aid selection and germplasm conservation programs for there improvement programs. Genetic markers are particularly important for germplasm conservation and core-collection development (Diwan *et al*. 1995; Tanksley and McCouch 1997), as well as in enhanced breeding applications, such as marker-assisted selection (Staub *et al*. 1996). Crucial to the successful implementation of marker-aided breeding strategies is the development of highly informative, easily transferable and reliable-genetic markers.

The first step towards the development of genetic markers for quinoa was the development of a genetic linkage map by Maughan *et al*. (2004). This map, which covered an estimated 60% of the genome, was based primarily on amplified fragment length polymorphism (AFLP). Unfortunately, the difficulties associated with AFLP marker technologies and the associated transfer of this technology to developing world countries where quinoa is cultivated have limited the use of these markers to enhance quinoa improvement programmes. The next step in quinoa marker development was the characterization of 208 SSR markers reported by Mason *et al*. (2005). These SSR markers have already been utilized to assess the genetic diversity among quinoa accessions within the USDA collection (Christensen *et al*. 2007) and efforts to genetically characterize Andean and Chilean germplasm are currently under way (F. Fuentes, personal communication). Unfortunately, less than 10% of the 208 SSR markers identified by Mason *et al*. (2005) have been mapped genetically, and only 67 were considered highly polymorphic (H > 0.7)—highlighting the need for additional SSR marker development and genetic mapping.

The objectives of the present study is: (i) the development of a new set of polymorphic SSR markers to augment the number of SSR markers already available in quinoa and (ii) the construction of a new genetic linkage map of quinoa based primarily on the SSRs described here and by Mason *et al*. (2005) using an immortalized recombinant-inbred line (RIL) population.

Materials and methods

Plant material

For SSR development and characterization, seeds from 22 quinoa accessions representing the geographical distribution of cultivated quinoa were kindly provided by Angel Mujica, National University of the Altiplano, Puno, Peru and Alejandro Bonifacio, PROINPA, La Paz, Bolivia (table 1). Seeds from the out-group *Chenopodium berlandieri* Moq. (pitseed goosefoot, PI 595315) were kindly provided by David Brenner (USDA, *Chenopodium* curator, Ames, Iowa).

For genetic map construction, a RIL population was developed (designated Ku-2/0654). The Ku-2/0654 population consists of 82 F_6 plants from a cross of 'Ku-2' (Chilean coastal ecotype) and '0654' (Peruvian altiplano ecotype). The RIL population was produced by self-fertilizing a single F_1 plant and allowing plants of subsequent generations to self-fertilize.

All plants were grown in greenhouse in Provo (Utah, USA) in 15 cm pots using Sunshine Mix II (Sun Grow, Bellevue, WA), supplemented with 1 g per pot of Osmocote Plus® fertilizer (Scotts, Marysville, OH). Plants were maintained at 25°C under broad-spectrum halogen lamps with a 12-h photoperiod.

DNA extraction

Genomic DNA from all the plants were extracted from 30 mg freeze-dried leaf tissue following procedures described by Sambrook *et al*. (1989), with modifications described by Todd and Vodkin (1996).

SSR discovery and analysis

Enriched SSR libraries for GA, CAA and AAT repeat motifs were produced by Genomic Identification Services, (Chatsworth, CA) using genomic DNA from the Bolivian altiplano ecotype 'Surimi' according to protocol described by Mason *et al*. (2005). Libraries were plated in S-gal media (Sigma-Aldrich, S. Louis, MO) supplemented with 50 mg l-1 ampicillin, for blue-white detection of recombinant clones. Recombinant clones were sequenced bidirectionally using M13 forward (5′ GTA AAA CGA CGG CCA GT) and M13 reverse (5′ CAG GAA ACA GCT ATG AC) primers at the Arizona Genomics Institute (Tucson, AZ) using standard ABI Prism *Taq* dye terminator cycle-sequencing methodologies. The computer program, Contig Express (InforMax, Frederick, MD) was used to determine consensus sequences, eliminate redundant clones and identify SSRs. Primers flanking each unique SSRs were designed using the web-based computer program Primer3 v2.0 (Rozen and Skaletsky 2000) according to the program's default parameters with the following exceptions: product size = 150–250, max T_m difference = 1°C and max poly-X = 3. Oligonucleotide primers were synthesized by Integrated DNA Technologies, (Iowa City, IA). All primers were screened on a panel of DNA from seven quinoa accessions (table 1) and one *C. berlandieri* accession (PI 595315). This panel was used to eliminate monomorphic primer pairs or primer pairs that failed to amplify. Primers that successfully amplified on this panel and showed simple-amplification patterns were subsequently run on a full panel consisting of 22 quinoa accession and one *C. berlandieri* (PI595315) accession. All data analysis, including calculation of heterozygosity (H) values were performed using data obtained from this full panel.

BAC (bacterial artificial chromosome)-end sequence SSRs (BES-SSRs) were identified using the web-based computer program Tandem Repeats Finder (Benson 1999) and 2580 quinoa BAC-end sequences. The sequences were obtained from clones of a quinoa BAC library previously reported by Stevens *et al*. (2006). Only sequences with total repeat lengths greater than 20 bp (n = 10 for dinucleotides;

Table 1. *Chenopodium* accessions used in the SSR diversity analysis. The SSR preliminary screening panel consisted of the samples 1–8. Heterozygosity values were determined with all samples (samples 1–23).

Samples	Name	Species type	Country origin	Source
1	PI 595315	*C. berlandieri*	lowa, USA	USDA
2	0654	valley	Peru	PROINPA
3	Sayana	Andes	Boliva	UNA
4	KU-2	coastal	United Kingdom[c]	PROINPA
5	Real	Salares	Bolivia	UNA
6	Ecu-420	valley	Ecuador	UNA
7	NL-6	coastal	Holland[c]	PROINPA
8	Jujuy	Andes	Argentina	UNA
9	Ratuqui	Andes	Bolivia	UNA
10	L-P	Andes	Bolivia	UNA
11	Ollague	salares	Chile	UNA
12	Surumi	Andes	Bolivia	UNA
13	Chucapaca	Andes	Bolivia	PROPINPA
14	Mocko	salares	Bolivia	UNA
15	Cica-127	valley	Peru	UNA
16	Illpa-INIA	Andes	Peru	UNA
17	E-DK-4	coastal	Denmark[c]	UNA
18	Kamiri	Andes	Bolivia	UNA
19	Kancolla	Andes	Peru	UNA
20	Ingapirca	valley	Ecuador	UNA
21	Nariño	valley	Colombia	UNA
22	G-205-95	coastal	Denmark[c]	UNA
23	Salcedo-INIA	Andes	Peru	UNA

[a]Altiplano types are subdivided in valley, salares and high Andes.
[b]PROINPA, The foundation for the promotion and investigation of Andean Products, La Paz, Bolivia; UNA, national university of the altiplano, Puno, Peru; USDA, Ames, IA.
[c]Coastal types selected from Chilean material.

$n = 7$ for trinucleotides etc.) were selected for primer design using the program Primer3 v2.0 (Rozen and Skaletsky 2000) as described previously.

PCR amplifications of the SSRs were performed in 10-μl PCR reactions containing 30 ng genomic DNA, 0.2 mM of each dNTP, 2.5 mM $MgCl_2$, 1× PCR buffer, 0.1 mM cresol red and 2% (w/v) sucrose, 0.5 U JumpStart *Taq* polymerase (Sigma-Aldrich, Saint Louis, MO), 1.0 μM forward primer, and 1.0 μM reverse primer. Thermal cycling profiles was as follows: 94°C for 60 s, followed by 19 cycles of 94°C for 60 s, 64°C for 30 s (decreasing 0.5°C every cycle), 72°C for 60 s; 30 cycles of 94°C for 60 s, 55°C for 60 s, 72°C for 60 s, followed by a final extension at 72°C for 10 min. PCR products were separated on 3% Metaphor agarose gels (Cambrex Bio Science, East Rutherford, NJ) at 120 V for 4–5 h. All gels were run in 0.5 × TBE and were visualized using ethidium bromide staining with UV transillumination.

Heterozygosity analysis

The information content for each new SSR was described using the H value. In a multiallele system, H values estimate the probability that any two individuals taken at random from a population will be polymorphic and is determined using the following equation: $H = 1 - \sum_{i=1}^{k} P_i^2$, where P_i is the frequency of the ith allele and k is the number of alleles (Nei 1978).

Additional markers

Amplified fragment length polymorphisms: To increase the number of markers on the map, AFLP analysis was performed on the Ku-2/0654 RIL population following the procedures described by Vos *et al*. (1995), with minor modifications for quinoa as described by Maughan *et al*. (2004).

Nucleolar organizing region (NOR) mapping: Maughan *et al*. (2006) reported the cloning and sequencing of the intergenic spacer (IGS) region of the 45S NOR in quinoa. Sequence analysis of the parents of the Ku-2/0654 population revealed a 43-bp indel polymorphism - present in 'Ku-2' (GenBank # DQ187958) and absent in '0654' (DQ187960). Segregation analysis of the NOR was performed using standard PCR (as described above) on the Ku-2/0654 population with primers

flanking the indel (5′-TTT GAA ACC ATA ACA CAC CTA TAA AG and 5′-TGG TCC AAA GAA TGG GTA TTT). PCR products were resolved on 1.4% agarose.

IIS seed protein mapping: Stevens *et al*. (2006) reported the isolation of two BAC clones containing homeologues of the 11S seed storage protein gene in quinoa. Two 11S loci, 11SA and 11SB, presumably one from each of quinoa's subgenomes, have been cloned and sequenced by M. R. B. Balzotti 2007 (Submitted to *Int. J. Plant Sci.*). Sequence analysis of 11SA revealed a polymorphism between parents of the Ku-2/0654 population in a *DraI* restriction site, allowing for mapping of the polymorphism using a standard cleaved amplified polymorphic sequence (CAPS) assay (Konieczny and Ausubel 1993). Briefly, DNA from the parents and the RIL population was amplified in a 10μl reaction containing 30 ng genomic DNA, 0.2 mM of each dNTP, 2.5 mM $MgCl_2$, 1× PCR buffer, 0.1 mM cresol red and 2% (w/v) sucrose, 0.125 U JumpStart *Taq* polymerase, 0.5 mM forward primer (5′ ACA ACA CCG GAA ATG AGC CT), and 0.5 mM reverse primer (5′ CCA CTG AAT ACG TTG CCG C). PCR conditions were as follows: 95°C for 5 min; 40 cycles of 95°C for 30 s, 65°C for 30 s, and 72°C for 30 s; followed by a hold at 72°C for 7 min. The PCR product was brought up to the volume of 20μl with water, 1× Tango buffer (Fermentas, Hanover, MD) and 5 U *DraI* restriction endonuclease (Fermentas, Hanover, MD), and was incubated at 37°C for a minimum of 2 h. Restriction fragments were size-separated on 1% agarose at 150 V for 2 h, and visualized using ethidium bromide staining with UV transillumination.

A Taqman allelic discrimination assay was used to map the 11SB locus. The allelic discrimination reactions were performed using Applied Biosystems (Foster City, CA) PCR Supermix according to the manufacturer's protocol. The final reaction consisted of 30 ng of quinoa genomic DNA, 0.4 μM forward (5′ GCG CTT TTT CCA ATA TTA GAC TCA A) and reverse (5′ TGT TGA AGT TGG TAC GTA AGC ATC A) primers, 0.2 μM of each discrimination probe (5′ 6FAM-TTG TTT GCT ACA TTC A; 5′VIC-TAT TGT TTG ATA CAT TCA AT; SNP is underlined) and a 1× concentration of the PCR Supermix, which includes an internal ROX standard dye. PCR amplifications were carried out on an ABI 7300 RT-thermocycler using the following thermal cycling conditions: 50°C for 2 min, 95°C for 10 min, 40 cycles of 95°C for 15 s and 60°C for 60 s. The analysis of the allelic discrimination assays was performed using the SDS v2.0 software (Applied Biosystems, Foster City, CA). Genotype calls for each accession were determined by inspecting the plot of the fluorescence signals (standardized with ROX values) from each of the allelic discrimination probes (VIC versus 6FAM) generated from the post-PCR fluorescence reads (end-point analysis). Fluorescence of only the 6FAM probe or only the VIC probe indicated homozygosity for a particular allele while intermediate fluorescence from both reporters indicated heterozygosity at the locus. DNA samples with allelic genotypes, verified via sequencing, were utilized as internal standards to validate each TaqMan SNP assay.

Betalain locus: The betalain colour locus (scored as stem colour, red versus green) was phenotypically scored on each RIL from the Ku-2/0654 population.

Map construction

For map construction, markers were scored as codominant (as was the case with a majority of the SSR markers, NOR and 11S loci) or dominant (majority of the AFLP markers and betalain locus). Marker segregation was analysed for conformation to Mendelian ratios expected in RILs using a chi-squared test, with two and one degrees of freedom for codominant and dominant markers, respectively. Linkage groups were constructed with a minimum LOD score of 3.0 using the default mapping parameters (LOD>1.0, recombination threshold, 0.4; ripple value, 1; jump threshold, 5; Kosambi mapping function) of the computer program JoinMap, v3.0 (Ooijen and Voorrips 2001). Linkage groups from the two different populations that shared at least one common marker were combined using the 'Combine groups for map integration' function of JoinMap (Stam 1993).

Results and discussion

SSR discovery and analysis

Libraries were enriched for motifs GA, CAA and AAT based on the results reported by Mason *et al*. (2005) that suggested the quinoa genome contains high frequencies of these repeats. A total of 1172 clones were sequenced, including 490 clones from each of the GA and CAA libraries and 192 clones from the AAT library. A total of 436 (37%) clones were identified that contained unique SSR sequences, of which 402 were suitable for primer design (178, 85 and 139 from the GA, AAT and CAA libraries, respectively). As expected from enriched libraries, the most common repeats observed in the study were GA (49%), CAA (35.6%), and AAT (12.9%). Other repeat motifs, including CA, CGA, GAA and GGT, were also observed, albeit infrequently. Of the 402 SSRs tested, 216 (54%) were polymorphic when tested on the screening panel of seven quinoa accessions (table 1). An additional 19 (4.7%) were polymorphic when the *C. berlandieri* accession was included in the analysis (interspecies polymorphism). The remaining primers (165) were monomorphic or amplified poorly. In only nine (2.2%) cases did a primer successfully amplify in quinoa but not in *C. berlandieri*, suggesting that these two *Chenopodium* species share a high degree of DNA sequence homology. Indeed, gene flow between quinoa and *C. berlandieri* has been reported previously (Wilson and Manhart 1993). Most polymorphic markers had repeat lengths greater than 20 bp, confirming the conclusions of Mason *et al*. (2005) who suggested that the future development of SSR markers in quinoa should focus on the identification of markers with repeat

lengths of >20 bp in order to maximize polymorphism (H values). All 216 polymorphic SSRs, including 111 dinucleotide, 104 trinucleotide and one hexanucleotide repeats were screened on the larger panel of 22 quinoa accessions and one *C. berlandieri* accession to determine their heterozygosity values (table 1). A total of 888 alleles were observed across all 22 quinoa samples included in the full panel. The observed number of alleles per SSR ranged from 2–13, with an average of four alleles per SSR. H values ranged from 0.12–0.90, with an average value of 0.57 (table 2). These values are similar to those observed previously in quinoa (Mason *et al*. 2005) as well as in related species such as sugar beet (*Beta vulgaris* L.) (Rae *et al*. 2000; Cureton *et al*. 2002). According to Ott (1992), a marker is considered polymorphic if H≥0.10 and highly polymorphic if H≥0.70. Based on these criteria, all 216 markers identified are considered polymorphic, and 53 (25%) are considered highly polymorphic (H≥0.70).

BLASTN and BLASTX sequence homology searches identified 14 SSR-containing sequences with significant DNA or protein homologies (E-value<$1E^{-10}$) to sequences in the GenBank databases (table 2). Three clones showed homology to known sequences at the nucleotide level only, while 11 showed significant homology to known sequences at the amino acid level only. Significant hits to annotated protein sequences on GenBank included proteins involved in developmental processes including a circadian-clock associated protein (KGA079), and proteins involved in defense responses and protection including Nim1 (non-inducible immunity-like protein; KCAA111). Metabolic proteins including isocitrate dehydrogenase (KGA026), succinyl CoA ligase (KGA171), sucrose transporter (KGA101), and oligosaccharyl transferase (KGA068) were also identified.

Parental screens and linkage analysis

SSR screening: A total of 424 SSR markers were screened on the parents of the Ku-2/0654 population, including the SSR markers developed here and those reported by Mason *et al*. (2005). Of these, 203 SSRs were polymorphic and easily scored, while the rest did not segregate in the population. The 203 polymorphic SSRs amplified 213 segregating loci, with 193 markers amplifying one locus each and 10 primers each amplifying two loci. Quinoa is an allotetraploid and it is likely that the second segregating bands amplified by these 10 SSRs represent amplification products from homoeologous loci from the two subgenomes of quinoa. Of the SSR marker loci scored, 190 (89%) loci were scored in a codominant fashion, while 23 (11%) were scored as dominant loci. Sixteen markers (7.5%) deviated significantly ($P < 0.05$) from the expected 1:1 segregation ratio, eight (3.8%) of which were highly significant ($P < 0.01$; figure 1).

AFLPs and other markers: Twenty-four primer combinations were chosen based on their previously demonstrated ability to amplify polymorphic loci (Maughan *et al*. 2004). A total of 81 polymorphic, easily-scored loci were amplified from the AFLP primer combinations. The number of polymorphic loci per primer combination varied from one to nine, with an average of 3.4 per primer combination. Of the 81 scored polymorphic loci, 79 were dominant and two were codominant. Thirty-one (39%) of the dominant loci were specific to 'Ku-2', while 48 (61%) were specific to '0654'. Fifteen AFLP markers showed distorted segregation at $P < 0.05$ (figure 1). The betalain colour locus, two homeologous 11S seed storage protein loci, and the NOR locus were also scored and none showed distorted segregation.

Segregation distortion: For the SSR markers, the identification of 16 markers that showed significant segregation distortion is slightly higher than those would be expected by chance alone (11 markers are expected at a $P < 0.05$). For the AFLP markers, segregation distortion was much more pronounced with 15 markers showing distorted inheritance patterns (5 markers are expected at $P < 0.05$). Skewed markers mapped to a total of six different LGs, although clustering of some skewed markers is evident. For example, of the 22 skewed markers (SSR and AFLP) that were placed on the map, 11 markers mapped to LG 1. Seven of these markers are skewed towards 'Ku-2', six of which are localized to the first 34 cM on the LG. Six of the remaining 11 skewed markers are localized to a 23 cM interval on LG 13 and all six markers are skewed towards 'Ku-2'. The presence of clusters of markers skewed to one parent or the other is suggestive of chromosomal regions containing possible gametophytic or zygotic factors (Zamir and Tadmor 1986; Lu *et al*. 2002). These skewed chromosomal regions may be associated with QTL conferring a selective advantage under the particular greenhouse growing conditions utilized to produce the RIL populations—we note that some (~10%) of the lines were lost during the population development process.

While segregation distortion is generally believed to be greater in interspecific crosses, reaching levels as high as 68.5% (Paterson *et al*. 1988), levels can also be high in intraspecific crosses (Hall and Willis 2005). Indeed, the extent of segregation distortion appears to be only indirectly related to the type of cross, and more directly related to the extent of genome divergence between the lines being crossed. Our RIL population is the result of crossing highly divergent altiplano and lowland coastal quinoa ecotypes. The parents of the Ku-2/0654 population have very low similarity coefficients (0.304), suggesting a high degree of genome divergence between the parents of this cross (Maughan *et al*. 2004). In addition, extensive introgression of genes via outcrossing with *C. album* and *C. hircinum* has been observed in lowland quinoa ecotypes, which may enhance genomic divergence between these two groups (I. Von Baer, personal communication).

Table 2. Quinoa microsatellite marker name, primary motif, complexity, type, primer sequences, significant GenBank homologies, expected PCR product size (PRO), observed number of alleles (ONA), and heterozygosity value (H).

Marker name	Primary motif	Complexity	Type	Forward primer (5′ – 3′)	Reverse primer (5′ – 3′)	GenBank homology	PRO[b]	ONA	H
KAAT001	(ATT)5GTT(ATT)3GTT(ATT)13	Simple	Imperfect	tggctatatcatatgcgtaatgtg	gggctcagattgtatctcgac		176	6	0.79
KAAT004	(TAA)14	Simple	Perfect	gtgcagctgctcacatcttc	tggcaataatagtttaggttgtgtg		198	5	0.56
KAAT005	(ATT)6(GTT)7(ATT)6(GTT)19(ATT)24	Compound	Perfect	caccactcaagcaatccaaa	gtgggagcccagattgtatc		293	4	0.57
KAAT006	(ATT)10(GTT)7(ATT)3	Compound	Perfect	tctgcaggatcggtaacctt	ttgtatctcggcttcccact		171	6	0.79
KAAT007	(AAT)30	Simple	Perfect	aggtacaggcgcaaggatac	cggtagcatagcacagaacg	BA000009	197	12	0.86
KAAT008	(ATT)27	Simple	Perfect	aggaacaactcgaagccaag	aaaggtgtgatcaagcaataacaa		177	7	0.73
KAAT009	(ATT)10	Simple	Perfect	agttgccaacatgcagagc	cgacgacgcaagacattaga		212	5	0.77
KAAT010	(AAT)16	Simple	Perfect	cggctctccactaacttcttg	atgtctttcgcctacccaaa		182	5	0.68
KAAT011	(AAT)17	Simple	Perfect	tttcagcaggatcgggttc	agccgaccagagcagtgtag		184	8	0.73
KAAT014	(ATT)10	Simple	Perfect	cgctgacgcttaacattcg	cacaaacaataattcaaccgaaga		191	3	0.45
KAAT016	(TAA)13	Simple	Perfect	gagcccgtgctacaactcat	ctgggcagagcagaacagat		186	5	0.57
KAAT018	(ATT)11	Simple	Perfect	gcaccaacctgagtcctagc	cgtgtcgctgctcatattgt		193	4	0.60
KAAT019	(ATT)10(ATC)5(ATT)3	Compound	Perfect	ctgcaaagcaaagtccatga	cttcagtaggatcgggttcg		196	5	0.60
KAAT020	(ATT)14	Simple	Perfect	gcctttattattgttcatttatttgtt	aggagtgggacccatattgt		199	4	0.57
KAAT021	(AAT)21	Simple	Perfect	cggctccctaccaatttctt	gcccaatggtctttgacact		199	5	0.64
KAAT022	(ATT)12	Simple	Perfect	cgggcagaaacatttaccaa	gcggctgctcacatcttta		199	5	0.66
KAAT023	(AAT)19A(AAT)13	Simple	Imperfect	agattgtatctcggctttcca	cacttcattgtattgcatttagga		225	7	0.74
KAAT024	(AAT)28	Simple	Perfect	cctaatgccacggtttccta	ccgctgaatagacacccagt		199	6	0.78
KAAT025	(AAT)14	Simple	Perfect	gagtgggagcccagattgta	agcaaagtaaatttcaacaaagca		160	2	0.34
KAAT026	(ATT)16	Simple	Perfect	cggagtcagatggttctggt	tcaagtgcagctcaatcacc		178	3	0.55
KAAT027	(ATT)12	Simple	Perfect	tttaaactttattgacccttggaaa	ggatgctattgcattgctga		180	3	0.55
KAAT030	(ATT)13	Simple	Perfect	tcaaatatgtgtggaccactctaag	ccaatttcttgtaaattgattgactt		199	4	0.69
KAAT031	(ATT)9	Simple	Perfect	agagaccaatgccggataga	gttcgctatagctagaggagtgg		200	6	0.77
KAAT033	(ATT)18	Simple	Perfect	tgccaactgacgagacaaag	gcgggagctcatatcttcac		200	5	0.62
KAAT034	(ATT)4ACT(ATT)12	Simple	Imperfect	aaagcaatcgaagcgttgtt	ttcgggtttgatgccataat		247	9	0.73
KAAT036	(ATT)18	Simple	Perfect	ggcagcgatcgtgaaata	gggacccaaattgtatctcg		175	6	0.70
KAAT037	(TAA)19	Simple	Perfect	tcaacctccgaatcctatcaa	ggatgctgattggtggataaa		284	13	0.90
KAAT038	(AAT)14	Simple	Perfect	ccttctctgctctgctatgctt	agcctagtgtcttgcgtcgt		363	6	0.64
KAAT039	(TTA)17(CTA)4TTGATAATG(TTA)10TCAAT(GTT)5	Compound	Imperfect	agccgagcagagcagattt	tgcggttgtagtcatttgaa		297	10	0.84
KAAT040	(AAT)2AAC(AAT)7	Simple	Imperfect	gcatgagtggtaatggagga	cttgaaggagcagtattattcaca		166	6	0.79
KAAT041	(ATT)13AG(TAT)4TAG(TAT)5TG(TTA)5	Compound	Imperfect	tgggacttccataaggcaac	atattgcatgtcgagcacca		182	10	0.86
KAAT042	(ATT)33	Simple	Perfect	tgaatcaaatagctttcatacattcaa	tatgttggcttcccaccaat		197	8	0.76
KAAT043	(AAT)24	Simple	Perfect	ggctcccactaatttcttgtg	tcatgcggcttgagtagttt		199	5	0.76
KAAT044	(AAT)11GCGATG(ATA)11ATGCG(ATA)4GT(AAT)20	Compound	Imperfect	gggtggaggcccagattat	cagagcagagctggcagag		272	7	0.81
KAAT045	(AAT)15	Simple	Perfect	cacattgtatctcggctttcc	cagatgcattgaccttcgtg		200	9	0.75
KAAT047	(AAT)26	Simple	Perfect	tctcggttccctactaatttcttg	tttatgcagcaagggttgtaaa		192	9	0.81
KAAT048	(ATT)14	Simple	Perfect	aaccgtgtcgtgcctaagac	ccagtgtgcaccaatgtagc		178	3	0.32
KAAT049	(CAA)8TAAGAATA(AAT)21	Compound	Imperfect	cagattgtatcccggcttc	tcgagtttcggatttgaatg		151	4	0.45
KAAT050	(ATT)9ATC(ATT)6	Simple	Imperfect	tcatgcctaggatcttgcttt	tcgtatacggactaaattgtccac		162	8	0.76
KCAA002	(CAA)11	Simple	Perfect	caatttcagctcccttgatcc	tattgttggtgcgcttgtgt		183	3	0.12
KCAA003	(CAA)7	Simple	Perfect	acctttcggctgctcagata	tgctgatgttgttgcagatg		179	3	0.53
KCAA005	(CAA)9CAT(CAA)6	Simple	Imperfect	tcaccgcccaccttactaac	gatttgcatgcccttcattt		169	2	0.09
KCAA006	(CAA)5	Simple	Perfect	ttgagcaggatgatgtggag	ttggagaaacatacctttgttgg		161	6	0.83
KCAA009	(ATT)6TT(ATT)8T(ATT)6	Simple	Imperfect	aatgacgtggaaccctaccc	tgctagggaacaatcaaggtg		187	3	0.65

Table 2. (*Contd.*)

Marker name	Primary motif	Complexity	Type	Forward primer (5′ – 3′)	Reverse primer (5′ – 3′)	GenBank homology	PRO[b]	ONA	H
KCAA010	(GGT)CT(GGT)GCTA(GGT)4	Simple	Imperfect	tgggtcgtagttctgggttc	cttatcaccagcagcagcac		191	3	0.66
KCAA011	(ATT)10(GTT)7GTA(GTT)15	Compound	Imperfect	tgaacccgcttcaacaatg	ccttcttcaaactccgaatcc		225	9	0.84
KCAA013	(AAT)7(AAC)4AAA(AAC)12	Compound	Imperfect	cctgtaaattgattgactttgtaggtt	gcaaagcacgtaaaccgtct		199	3	0.49
KCAA014	(GTT)12	Simple	Perfect	gaatttgcatgcccttcatt	ccgccctcgctactatgat		170	5	0.68
KCAA015	(GTT)7	Simple	Perfect	tggttggaggcaaacatacc	tgagggtgaagaggaggatg		198	4	0.54
KCAA016	(CAA)15	Simple	Perfect	cgcggttatttaagggaagg	ccaccaggagagctaggttg	AAF27057	188	6	0.50
KCAA019	(GTT)8	Simple	Perfect	gtagttgggcggatgtgtct	gcgactgagctagcaggttt		166	4	0.70
KCAA022	(GTT)12	Simple	Perfect	ccaattgcatgctcctcatt	aatgcaaacatgggaggaga	CR377726	157	5	0.74
KCAA023	(GTT)2(TGT)3TTT(GTT)3GCT(GTT)2	Compound	Imperfect	tgctgttgttgttgttgtgc	caaatagcaacacggcaataga		193	3	0.52
KCAA026	(CAA)4CAT(CAA)2CAG(CAA)3	Simple	Imperfect	gacgacgacgataacaacga	agccaattcccatcatcaga	CR377898	191	2	0.29
KCAA027	(ATTGTT)19	Simple	Perfect	agagcagagccgagtagagc	gctcacctaaatcgtatatgcact		172	5	0.68
KCAA028	(GTT)7	Simple	Perfect	ggtcgtcctacacctcttgc	cccgcagggtaaccataata		196	2	0.28
KCAA029	(CAA)2CAT(CAA)5	Simple	Imperfect	cagactgcaggcaccaca	gttgttgtggttgttgttattggt		193	9	0.75
KCAA031	(CAA)19(CGA)2(CAA)11	Compound	Perfect	ttgtatctcggcttcccact	ggcttcagttcattaacagcacta		186	4	0.67
KCAA032	(GTT)8(GAT)2	Compound	Perfect	cttgtcacatgccaagttgc	aacaacaacagcagcaccaa		156	3	0.66
KCAA033	(GTT)6	Simple	Perfect	ttccatttgggctctcattc	aggactcgggtgtcctacct		182	4	0.42
KCAA036	(CAA)10	Simple	Perfect	ctgctgaccaatggctaggt	tcatcatcatcaccatcatcatc		250	8	0.69
KCAA038	(CAA)7TAACAACATCAG(CAA)6	Simple	Imperfect	caatggtgtgctacccacag	gtatggcaagttgcatgctc		181	2	0.21
KCAA041	(GTT)8	Simple	Perfect	tggtcgtagaccacccattt	cggatcactccacccttgta		197	2	0.09
KCAA044	(CAA)3CAG(CAA)4	Simple	Imperfect	gcaatgagatgcaacgaatg	ttgcaaagcctccaaatctt		160	2	0.30
KCAA051	(GTT)14	Simple	Perfect	catgctcatcatttgctgct	gtctttggagcggaatgcta		196	7	0.77
KCAA053	(GTT)8	Simple	Perfect	ggagtatcctttgttaaattggtctc	aggcaaagtccatggaacag		160	2	0.50
KCAA057	(CAA)6(CAT)3(CAA)2CAG(CAA)2	Compound	Imperfect	tgtgctaccaactgctctgg	ttggttctccatcaggctct		187	6	0.79
KCAA058	(ACA)4AA(ACA)6	Simple	Imperfect	ggcgcaaggaatttgatagt	cctgctccttctccatcaag		166	4	0.63
KCAA063	(CAA)9	Simple	Perfect	tccgatgatgaagaggagga	gatttgcaaaccgctcattt		180	3	0.35
KCAA065	(CAA)5AGC(AAC)2CACACCGAC(AAC)3	Simple	Imperfect	gccatcctagttggcgttt	tctgtccattatcaacttcacca		281	8	0.80
KCAA066	(GTT)9ACGGAATTC(GTT)5	Simple	Imperfect	aaaccgctcatttgctcact	ggcacgttcccaagtcttat		211	2	0.38
KCAA067	(CCA)3(ACA)8	Compound	Perfect	atgagggcacagaggatgag	gagaggtgttgatgggaaaca		187	3	0.54
KCAA068	(ACA)7	Simple	Perfect	cagcaactgaaaccagcaa	gcagctgctgttgctaaatac		186	10	0.79
KCAA069	(CAA)9TGCAT(CAA)2	Simple	Imperfect	tggtggtggagaggaagaac	tcatgtgctccatttgcttt		181	3	0.54
KCAA071	(TGT)7TGGTGTCG(TCT)4TCCA(TGT)10	Compound	Imperfect	tccctgccatatcttgttga	acatagcggtggatttggag		191	4	0.56
KCAA078	(CAA)3AAA(CAA)9	Simple	Imperfect	aggcgaggataacatgatcg	aagaagccatacctccctcac		170	3	0.51
KCAA083	(GTT)7(AGTT)4ATT(GTT)3(AGTT)4	Compound	Imperfect	tgttgttgttaagtttatttgttgttg	ccagaaccctcgatctacataaa		194	2	0.13
KCAA084	(CAA)3AAA(CAA)9	Simple	Imperfect	gcggaatgcttggaatgtat	atgctcaagtctgctcatgc		196	4	0.58
KCAA086	(CAA)4CT(ACA)2C(CAA)3	Compound	Imperfect	ccctgcgtaaattctctcca	gggagctagcatatgggtga		191	3	0.65
KCAA088	(GTT)10CAC(CAT)4	Compound	Imperfect	tgtgcttgcaaagcctctaa	gatgctccgaatgtttggtt		161	5	0.77
KCAA091	(GTT)5(ATT)4	Compound	Perfect	tttgttgttgtcgttgttgttg	atattgcatgtcgagcacca		186	4	0.72
KCAA093	(GTT)2CTTATAGGT(GTT)7	Simple	Imperfect	tgctgatgttgttgcagatg	acctttcggctgctcagata		179	2	0.43
KCAA095[c]	(GTTGGT)2(GTT)2	Compound	Perfect	gctggtggagtcgaattgat	agccacttccccttctcact		150	4	0.71
KCAA101	(CAA)7G(GAA)2GG(CAA)4(CAT)2	Compound	Imperfect	gccatgagagtgtggaggat	gggaagcaaacataccttgc		165	4	0.69
KCAA104	(CAA)12CTACTACAACTA(CAA)5	Simple	Imperfect	caacaacaagtacaacaacatcca	cggaaatttcaggcagatgt		232	4	0.73
KCAA105	(CAA)10	Simple	Perfect	cgaacagcagcaacaataaca	cctttagacgccaccgtact		199	2	0.38
KCAA106	(CAA)20	Simple	Perfect	atatggaagtcggccaacag	gcatgctcatcatttgttgc		153	10	0.83
KCAA107	(CAA)18	Simple	Perfect	caccagaaccctcgatctaca	tggttactgttgttgttgttaatttg		271	4	0.64
KCAA111	(CAA)2TAA(CAA)4CAC(CAA)6	Simple	Imperfect	ctcacattgagcccaacaaa	ctccaacgggtgcataaatc	BAD87216	156	2	0.45
KCAA112	(CAA)8	Simple	Perfect	cgttgtcaagtgattcaagacc	aaagattggaggctttgaagtaaa		190	3	0.61

Table 2. (*Contd.*)

Marker name	Primary motif	Complexity	Type	Forward primer (5′ – 3′)	Reverse primer (5′ – 3′)	GenBank homology	PRO[b]	ONA	H
KCAA117	(GTT)9	Simple	Perfect	ccgtggttcctctagagtcg	cctccaacaacctttctctcc		154	11	0.72
KCAA118	(GTT)2ACGATT(GTT)7	Simple	Imperfect	gctgtgtttgacccatgttg	caaccacagcaaaggtgtga	NP_909548	159	4	0.45
KCAA120	(GTT)11TTTGTG(GTT)2	Simple	Imperfect	ccaccaggagagctaggttg	cgacgtaccttcccttaaaca		153	2	0.42
KCAA125	(GTT)5(GTC)2TTA(TTC)2CATGTTTTTAT(TG)2(GTT)9	Compound	Imperfect	ttgcatgctctccatttaagc	ggtgcatgaggaggatgact		184	3	0.09
KCAA126	(GTT)8TTT(GTT)23	Simple	Perfect	catattggtgatgttgctcttga	cgccctccctactatgatga		166	3	0.52
KCAA130	(GTT)7	Simple	Perfect	gggaagcaaacataccttgc	atgagggcacagaggatgag		165	2	0.50
KCAA132	(CAA)5AAG(CAA)2CCA(CAA)3TCAG(CAA)5	Simple	Imperfect	caaactgcaggcaccaca	caacttcaccatacgcattca		220	3	0.64
KCAA133	(CAG)4(CAA)2CAG(CAA)3CAG(CAA)3(CAG)2	Compound	Imperfect	gcctttagctgttgaaggtgt	tgttgttgttgttcctcctga		165	5	0.66
KCAA136	(GTT)5GC(CTT)2TCTT(GTT)6	Compound	Imperfect	catttgggctctcattgctt	ttcgggtgtccttcctaatg		181	6	0.70
KCAA137	(CAA)16AAA(CAA)5	Simple	Imperfect	caatgatgtgctacccaacg	ttgctcaaggctactcatgc		178	4	0.71
KCAA139	(CAA)3CGA(CA A)5	Simple	Imperfect	gaacacccaacctgcaaact	caacttcaccatacgcattca		180	4	0.59
KCAA141	(CAA)19	Simple	Perfect	gacgagtggatgtggtgatg	accgttgtcattgtcttggtt		171	4	0.51
KCAA143	(GAA)2GG(AGA)4	Simple	Imperfect	ccagggtgaatcagggaata	gggcattaattactctctctctctct		150	2	0.30
KCAA152	(CAG)5AG(CAA)2CAT(CAA)5	Compound	Imperfect	aacataaagcgccaacctg	tcaccatacgcattcattcttt		285	3	0.59
KCAA153	(GTT)5(GCT)2	Compound	Imperfect	tgttgggtggcgatcatac	ggaagagttccctccgtttc		177	3	0.58
KGA002	(CT)15	Simple	Perfect	aaagaacgcatccttccaat	aacctagccaacactccctaaa		193	5	0.51
KGA003	(GA)16	Simple	Perfect	attgccgacaatgaacgaat	atgtaaatggcatgtcccaac		150	5	0.65
KGA006	(GA)16	Simple	Perfect	aaacaaattctatcattcggttagg	gccaacgagcctgatgtaa		175	2	0.38
KGA008	(CT)13	Simple	Perfect	ctcaaatttctgcctcctga	aaatctctgcctctgtgcaa		179	3	0.64
KGA009	(CT)18	Simple	Perfect	tccaagaccaaaccctctct	tgtgtgtatatgagagagagagaca		154	4	0.66
KGA010	(TC)11	Simple	Perfect	tgtttcctgcgtccctattc	gctgaaggtgaaatagtgga		198	2	0.50
KGA014	(GA)15	Simple	Perfect	gaccacatgcataaattaatacgact	tcgtaggtcgaggatcttgc		165	3	0.44
KGA015	(CT)11CC(CT)10	Simple	Imperfect	accagcttgcttgtcttcct	ggataaccgctccaatgcta		173	3	0.45
KGA016	(GA)22AA(GA)5	Simple	Imperfect	ccctgcttaatctccgtgaa	ccgaaccaagactacgaaaca		174	3	0.57
KGA019	(GA)33	Simple	Perfect	tcaccacctttgcaaacaac	cacgagaccaagcctctctt		173	3	0.14
KGA020	(CT)22	Simple	Perfect	tcacctacctcggtaaaggaaa	ggagcagatgatgaacatgg		177	2	0.44
KGA021	(CT)24GTGC(GT)10	Compound	Imperfect	gacctattaaaggttccgcaca	ggtccacacacacacagagc		195	4	0.73
KGA023	(CT)31	Simple	Perfect	cacgagaccaagcctctctt	tcaccacctttgcaaacaac		171	5	0.68
KGA024	(AG)31	Simple	Perfect	caagaaggtgttgggatgtgt	tgtggaattgtgagcggata	XP_731877	165	3	0.41
KGA025	(CT)22	Simple	Perfect	gaggtcgtatcatcccgttt	gcgagtacaggaggatttgc		170	8	0.54
KGA026	(CA)15	Simple	Perfect	gtggttcatggctgatcctt	caccacccttctggtgaact	AAY58195	250	2	0.39
KGA027	(GA)34	Simple	Perfect	ttgtacagaggaagtggcaaga	catcttacagctctggctttcc		153	6	0.56
KGA030	(CT)14	Simple	Perfect	tcttgatcccatcttacccaac	tcgtggagttgtggttcatc		175	5	0.53
KGA035	(CT)15	Simple	Perfect	cattgccggacttctgattt	ccctgcattgacaagcatta		151	6	0.61
KGA036	(GA)26TA(GA)7	Simple	Imperfect	cccaaatgtgaggtttcatta	ttgcccagaatatgacaagtt		174	3	0.45
KGA037	(AG)15	Simple	Perfect	tcgaatatggctaggtgtttct	cattcaccaattacaaccaatttt		200	3	0.58
KGA038	(CT)8	Simple	Perfect	atggacctccaataatcacca	gagagagaaagaggagagagaaagtg		150	3	0.56
KGA039	(CT)14	Simple	Perfect	tgtttcaccttcccttagcttt	tttggttcttaagagggatgc		260	4	0.23
KGA040	(TC)14	Simple	Perfect	accctcctcctttccacagt	ggaacgtcgggtcgagtat		189	6	0.74
KGA041	(GA)20	Simple	Perfect	tttggtgcaaatgttgttca	ttccaagaccaaaccctctc		225	4	0.63
KGA042	(AG)27	Simple	Perfect	ttggtagtgggtaagagaacctg	ctcctccagccacataatc	NP_178206	185	4	0.55
KGA044	(AG)3CG(AG)7	Simple	Imperfect	taacctgcacagggtgacaa	gaaccaattacaacggaaagga		189	2	0.48
KGA046	(GA)23	Simple	Perfect	tccaacttcagatggatgaaga	atcgttggcattctccaaat		187	5	0.70
KGA047	(CT)24	Simple	Perfect	gcagtgcatgaatttggaca	gaagctggcaccttatacatgc		184	3	0.54
KGA048	(GA)38	Simple	Perfect	acgtcgaggatggctaggt	ccaacaatcatcatcaataccc		199	5	0.74
KGA049	(AG)17	Simple	Perfect	cgagaaaggagccggatatag	tttcctcccaacctttctctc		173	4	0.52

Table 2. (*Contd.*)

Marker name	Primary motif	Complexity	Type	Forward primer (5′ – 3′)	Reverse primer (5′ – 3′)	GenBank homology	PRO[b]	ONA	H
KGA050	(CT)14	Simple	Perfect	tgctcaaattactaaactaccgaca	tcctgagtattgatcgcaagg		157	2	0.39
KGA051	(GA)22	Simple	Perfect	gtgaggaatcagtccggtaca	ccaatctggtcaagcacctc		200	4	0.51
KGA052	(CT)14	Simple	Perfect	tttcttggtgttgattcatttatgtt	catctccttctcaaccacagc		175	3	0.56
KGA053	(GA)13	Simple	Perfect	aaatttctgcctctgtgcaa	ctcaaacttctgcctcctga		182	3	0.48
KGA054	(GA)18	Simple	Perfect	tgttgattgataatatgtaatggtgga	cattcataacagcgagagatgg		193	3	0.33
KGA055	(CT)8	Simple	Perfect	cccaacccaccaaacttaca	gaaaggaaagtgattgcaaagaa		162	2	0.49
KGA056	(GA)12	Simple	Perfect	gactaacggtgtccaaactgc	ccttctgcattacaccgtca		176	2	0.31
KGA059	(AG)19	Simple	Perfect	ataaccactccgatggcaaa	cagccacctggcagttaga		195	3	0.61
KGA060	(AG)12	Simple	Perfect	agtggagagaacgctggaa	tctctcctctcctaggatgctc		177	2	0.36
KGA065	(AG)31	Simple	Perfect	tatatccgacaaggcgacaa	tgtaatgttacgagtacatgttcagtt		167	4	0.73
KGA066	(CT)15	Simple	Perfect	aaagaacgcatccttccaat	aacctagccaacactccctaaa		195	3	0.38
KGA068	(CT)13	Simple	Perfect	tcccgctggaattattgtaag	aaacgagcttgcatcagaca	NP_568380	200	3	0.56
KGA069	(CT)13	Simple	Perfect	ggatggtctcttggcacaaa	cccgaaagcatattaaccagaa		186	4	0.70
KGA070	(AG)11	Simple	Perfect	aggttcttggacaaagggaaa	tgaaataaatggccgagagg		165	5	0.70
KGA071	(CT)13CA(CT)6CC(CT)10	Simple	Imperfect	agcatttattacacacacacacaca	aatccgggtttaaccattcc		170	3	0.59
KGA073[c]	(GA)4ACAAGAG(GA)2ACGACA(GA)2	Simple	Imperfect	tcaatgttggtggtgctgtt	aaaccctaagacacgtacaactcc		280	2	0.50
KGA074	(GA)21	Simple	Perfect	tatatgccaccggaatgtca	tgtatccctttgcattctttga		151	2	0.39
KGA079	(TC)19	Simple	Perfect	gccaacgagcctgatgtaa	aaattctatcattcggttaggtaatca	AAQ73524	177	3	0.43
KGA080	(AG)14	Simple	Perfect	ccagggtgaatcagggaata	ctggcaggtgggtcttctat		200	2	0.50
KGA082	(CT)12	Simple	Perfect	cccaccaaattcattcttga	aagagggagagggaagaagc		158	3	0.53
KGA083	(AG)28	Simple	Perfect	aagatggtttgaggtgtgtttc	ataaaggcacccgtgataaa		177	4	0.73
KGA086	(AG)2A(AG)7AA(AG)12	Simple	Imperfect	ccactccgatggtaaagtcaa	gtggacaccaaccactagca		193	3	0.55
KGA088	(TC)14	Simple	Perfect	tgtaaatggcatgtcccaac	ttattgccatttcagggattt		198	4	0.53
KGA091	(CT)14	Simple	Perfect	tagcaaccagcagaggtcaa	tccaaaccaactcacaaacact		153	4	0.54
KGA092	(AG)18	Simple	Perfect	agagcagggataaggctgtg	gtggtacgtagccatcagca		158	4	0.57
KGA093	(TC)13	Simple	Perfect	cctccaagcccaaatcttta	tccggatgaagataaagaagga		195	3	0.66
KGA094	(CT)12CC(CT)T(TC)10C(CT)6	Simple	Imperfect	gacttggtgcctagggtttg	ggaaggagagagtgccatga		177	6	0.77
KGA097	(GA)8	Simple	Perfect	acgacgctgacatttgtagg	tcgtccctctctctctctcc		198	3	0.56
KGA098	(GA)14	Simple	Perfect	tctgggaataaccgctctga	gcaccagcttggttaccttc		195	2	0.35
KGA100	(AG)12	Simple	Perfect	tgcaatgtcgagaatggcta	ccaacaatcatcatcgtcaca		174	2	0.48
KGA101	(CT)25	Simple	Perfect	tgacaatgtaaagttcatgacaaa	gatacttccttgatttaaagacaacc	CAD58887	152	3	0.48
KGA107	(GA)13	Simple	Perfect	ccagggtgaatcagggaata	ctggcaggtgggtcttctat		200	3	0.54
KGA109	(GA)10	Simple	Perfect	accttgaaccacaccgaaac	tcgctgctcatcaccatatt		150	3	0.58
KGA111	(GA)19	Simple	Perfect	aatggtaaacagaccagactagca	tgggttcatttagtagaatcaagg		161	8	0.72
KGA114	(GA)14	Simple	Perfect	tgttgagtgcgctttaatgg	aataggtgtagccgcgtagg		173	3	0.55
KGA116	(CT)12	Simple	Perfect	ccttccttctctacgctctcc	tgggacccaaatctttcatag		199	4	0.75
KGA117	(CT)29	Simple	Perfect	gctttgtagacacctgtcatgg	ccactccgatgataaagttagaatg		197	6	0.74
KGA118	(CT)20	Simple	Perfect	attcccatccacacccatt	tgtcggttcaaacttggtca		187	4	0.56
KGA119	(GA)6A(GA)10	Simple	Imperfect	gggataaccgctatgatgct	ggtggcaccagcttgattat		166	6	0.73
KGA120	(CT)19	Simple	Perfect	tttgcatgccatgtagcc	tgaccactccgatgacaaag		193	7	0.78
KGA121	(GA)17	Simple	Perfect	ttaggaaggcaagtgtttaggg	tgccacgacaatttctatcg		197	2	0.47
KGA124	(GA)18	Simple	Perfect	gggaccaaaccctcagaaat	gatttccttaatccttcattcacc		166	2	0.43
KGA125	(GA)2G(GA)4	Simple	Imperfect	ggtcgttgatgacagtggtg	tcgatctcctctcctcctctc		188	5	0.72
KGA127	(CT)5C(CT)7C(CT)4	Simple	Imperfect	gtgaatcacgcttcgggtat	gctcgtgatcctcttggttt		163	3	0.54
KGA128	(CT)14	Simple	Perfect	tgctagggctctactgaactcaa	ctggctgcacttcctcttct		174	2	0.44
KGA129	(CT)23	Simple	Perfect	aactctccctacaccgtcacc	ttcctttctcaagtttggcatt		151	2	0.35

Table 2. (*Contd.*)

Marker name	Primary motif	Complexity	Type	Forward primer (5′ – 3′)	Reverse primer (5′ – 3′)	GenBank homology	PRO[b]	ONA	H
KGA130	(CT)20	Simple	Perfect	ccatgaggttctatggatctgg	acggttgtagcaggatgagc		172	2	0.38
KGA131	(GA)19	Simple	Perfect	acggttgtagcaggatgagc	ccatgaggttctatggatctgg		171	2	0.28
KGA133	(CT)20	Simple	Perfect	cagaaccatcatccctctctct	ctagggtgaaggcaacttcg		163	6	0.79
KGA134	(GA)18	Simple	Perfect	gcggctctgataccaatgat	tgtcagctgtcaagaggtttg		172	2	0.24
KGA135	(CT)17	Simple	Perfect	tctcgccttcattaccctctt	caaataatcggtgggtttgg		196	3	0.62
KGA136	(CT)17	Simple	Perfect	ccgacatttataaaggaagagaca	ccgcacctatcatcaagttaga		183	3	0.54
KGA138	(CT)3CA(CT)20	Simple	Imperfect	cgaaaccacccttctcaaac	taacaaacaaccgaccacca		159	5	0.70
KGA143	(CT)24	Simple	Perfect	gacagtgacaactacctctgtttca	gcgagtcacgagagagagaga		175	2	0.43
KGA145	(AG)12	Simple	Perfect	ccagggtgaatcagggaata	ctggcaggtgggtcttctat		196	2	0.32
KGA148	(CT)12	Simple	Perfect	acttggcgtgggatagtttg	ccactccgatgacaaagtga		159	3	0.66
KGA153	(CT)18	Simple	Perfect	tgttattcctcctcaagacctca	gatgatccgccatttctgtt		193	4	0.56
KGA155	(TG)11(AG)16(TG)9(AG)9(TG)8(AG)9TC(TG)5(AG)12(TG)7	Compound	Imperfect	ctctgttgacaatctaatttcagttct	tgatctgctgcaattctaaacc		156	4	0.66
KGA156	(AG)22	Simple	Perfect	ggcacaccgagagagaagag	agggctcggacaatgagtta		234	4	0.56
KGA157	(TC)9TT(TC)9	Simple	Imperfect	agtttgaccgagggaggatt	gagccctattggaaggacaa		183	4	0.58
KGA158	(TC)14	Simple	Perfect	ccttcaataaccaattatcagcaa	ctttcacgtctagggcgaag		179	3	0.43
KGA159	(TC)14	Simple	Perfect	tcaaatatgccctcttctcca	cctgagtgtggaggttgtca		163	4	0.58
KGA160	(ACT)4AT(TC)17	Compound	Imperfect	ggtccatggagcaaacaaa	aagctgaccaacatcgacaa	AAD21704	157	4	0.68
KGA165	(TG)19(AG)19	Compound	Perfect	caagcatatacctccatgtgc	gaagatatgctgccttgtaatca		274	4	0.71
KGA170	(CT)14	Simple	Perfect	ctggcaggtgggtcttctat	gaagaaggagaagaagaagaagaga		147	2	0.30
KGA171	(CT)12	Simple	Perfect	ttcctctgttcccttaatattgc	tagggatccatgtgatgagg	XP_478842	125	2	0.39
KGA173	(CT)17GT(CT)4	Simple	Imperfect	cccttatcttatgtgaatcaggaaa	cactattgtttggatagaaagatttgg		120	3	0.44
KGA174	(GA)27	Simple	Perfect	aagatgtcccggttgaagg	aaataggcaagttaatccttcattg		150	3	0.66
KGA175	(CT)21	Simple	Perfect	ttcaatcatcaacaactattaagtttc	tctagatcttgtgaatacggaattt		143	3	0.53
KGA177	(CT)32	Simple	Perfect	ctgcttctcatgtgttgatgg	caacaggctgattatgggtgt		125	4	0.67
KGA179	(GA)15	Simple	Perfect	tttcagaaatttgaagtcaagaaaga	gcttactttgactcattcccttct		130	2	0.40
KGA180	(CT)15	Simple	Perfect	atctcagcgagttggaatgg	ggatggtcctcctgaaatagc		149	3	0.52
KGA181	(AG)39	Simple	Perfect	aagttctttaaattgcaacatgg	atcacatggacaagtgtgtga		149	3	0.47
KGA182	(GA)11	Simple	Perfect	tttgttatcgcgcggtgt	ttgtgaagcaagcaaccaaa		173	2	0.34
KGA186	(CT)15	Simple	Perfect	ggaaaccaatggctgatgat	gacggatccatctgttgtca		178	3	0.63
KGA188	(AG)16	Simple	Perfect	ttcagctttcaataccctcaca	gcatttaaagaatcagtgacatgg		177	2	0.50
KGA189	(TC)26	Simple	Perfect	cataaatttgaacccaaacttgc	gggtgcgtgcctgtattact		197	3	0.59
KGA192	(GA)8	Simple	Perfect	cccaatgaccgttattctcg	tgccaatggttaggttgtca		296	2	0.50
KGA193	(TC)24	Simple	Perfect	tggagccagtagttcatgttgt	ggacgcttgtggattctcat		297	2	0.40
KGA199	(GA)2GGAA(GA)4	Simple	Imperfect	gtggtcaggtgggtcagtct	caccgtcaactccaaagaca		195	3	0.27
BAAT051[d]	(ATT)11	Simple	Perfect	aatgatgtcagcggttgct	aaacaagagttaaatgctaaattgct		111	NT[e]	NT[e]
BAAT052[d]	(TTA)11	Simple	Perfect	aaggtggttactttcaaataggg	agttaagtggtggtgaaaccaa		247	3	0.53
BAT002[d]	(AT)25	Simple	Perfect	cgagcatgtgcacctatcaa	ggcggtatttaacaacccaat		290	NT[e]	NT[e]
BAT003[d]	(AT)20	Simple	Perfect	aacttatgcatttgcgacct	tctccgtattatatatgcgtgtgt		329	NT[e]	NT[e]
BAT004[d]	(AT)25	Simple	Perfect	ttacagaaatgtaccaatcctacaaa	ggaacaacgcaacaagctaa		172	NT[e]	NT[e]
BGA200[d]	(GA)20	Simple	Perfect	accagccactttgtcattagg	gccatggttgatgaatgaga		161	4	0.71
						Average:	187	4.1	0.57

[a]GenBank accession numbers of protein and DNA sequences with homologies (E-value <1E-10) to specific SSR sequences.

[b]Reported in base pairs

[c]Markers not considered true SSR on the basis of the definitions of Weber (1990).

[d]BAC-end sequence derived SSRs (BES-SSR).

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

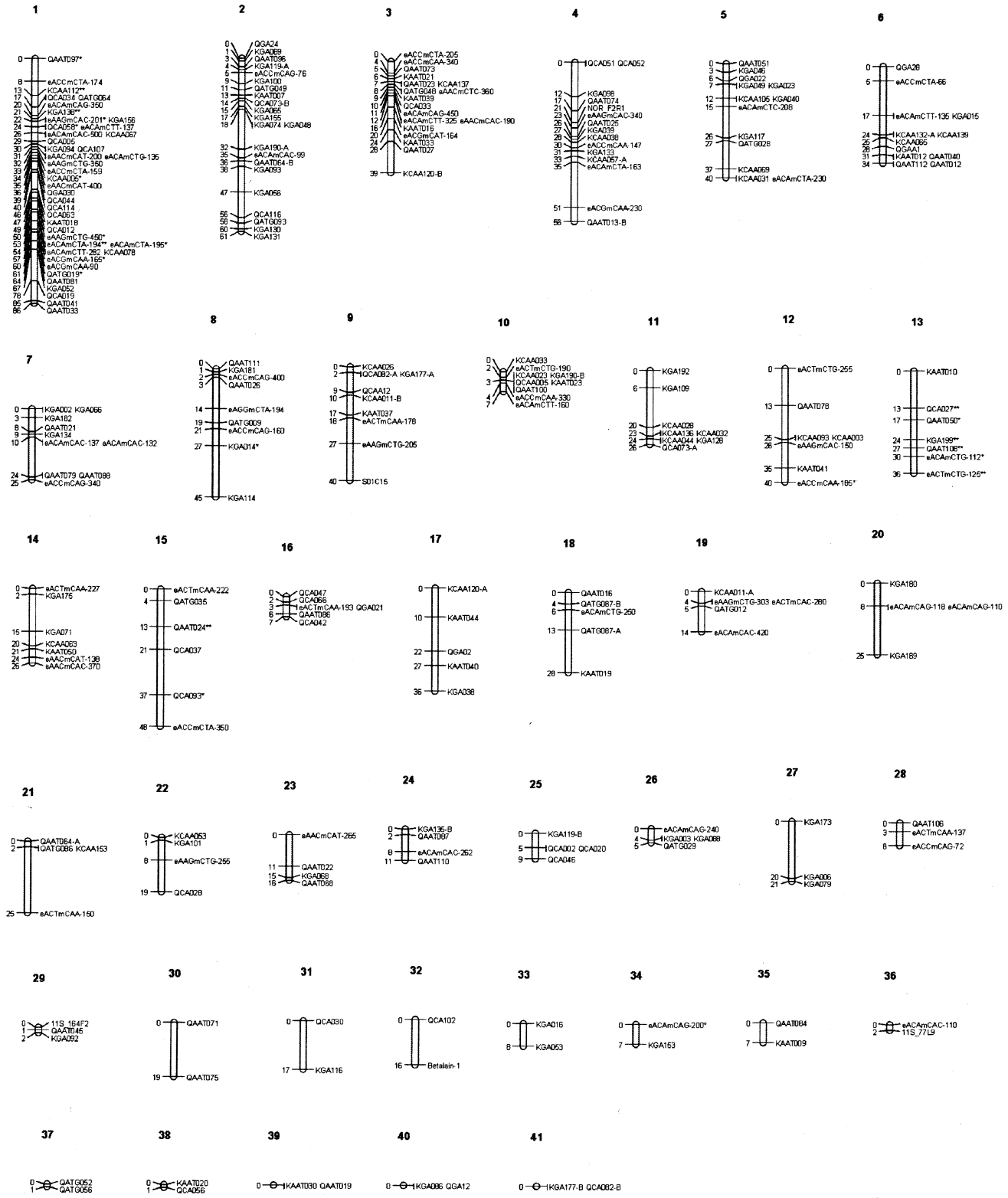

Figure 1. Linkage map of the Ku-2/0654 population. Distances in centiMorgans (cM) are indicated on the left side of each linkage group. All SSR markers begin with Q or K; newly developed SSR markers reported herein begin with K. AFLP markers begin with e. Markers skewed at $P < 0.05$ and $P < 0.01$ are indicated with * and **, respectively.

Map construction

The Ku-2/0654 populatio: A total of 298 loci were included in the linkage analysis; 275 (92%) of these mapped at a minimum LOD of 3.0, including 200 SSR, 70 AFLP, and all four additional markers (11S loci, NOR locus, betalain colour locus). The resulting map consists of 38 linkage groups covering 913 cM, or approximately 54% of the predicted 1700-cM quinoa genome (Maughan *et al*. 2004; figure 1). Linkage groups (LGs) were numbered based on the number of markers, with LG 1 containing the most markers. Linkage group lengths vary from a high of 86 cM (LG 1) to a low of 1 cM (LGs 38). The largest interval between two linked markers is 22 cM (on LG 21), and the average distance between marker loci is 3.3 cM.

Compared to the haploid chromosome number (18) of quinoa, the high number of linkage groups identified in this RIL population indicates that many regions of the genome have not been detected and that additional markers are still needed to coalesce linkage groups and provide complete coverage of the quinoa genome. The genome coverage reported here is similar to that detected in our initial F_2 AFLP map with approximately the same number of AFLP markers and mapping individuals (Maughan *et al*. 2004). However, this map represents a significant step forward in that, (i) it is built on a RIL population that has been selfed to near homozygosity ($F_{2:6}$) and (ii) it is primarily constructed of easily transferable sequence-based SSR markers. RIL populations are essential for the quinoa genetic mapping community, since they alleviate the need to develop new mapping populations each time new genetic markers become available. Indeed we are currently constructing new SSR libraries in the hopes of further populating this genetic map and coalescing linkage groups. We note that while RIL populations facilitate genetic mapping since they are immortal, they also result in an expansion of map distance for linked markers, which may present some difficulties in the initial stages of mapping since more markers will be needed to establish linkage groups. However, this expansion of map distances also gives an advantage over a conventional segregating mapping population (e.g., F_2) since recombination between closely linked markers is more readily detected (Broman 2005). Additionally, since the seed of RIL populations are essentially limitless, these populations allow for replicated field trials necessary for quantitative trait locus (QTL) mapping experiments.

Conclusions

The major objectives of this project were to increase the number of available SSR markers and to build the first SSR-based genetic map of quinoa. We report the development and characterization of 216 new SSRs markers and the development of the first genetic map of quinoa based primarily on sequence-tagged SSR markers. We also report the development of the first RIL mapping population and the development of several SSR markers derived from BAC-end sequences of a newly constructed BAC library (Stevens *et al*. 2006). The development of a physical map via restriction mapping of BAC-end sequences should prove invaluable in the targeted development of genetic markers, helping to coalesce linkage groups and integrate future genetic and physical maps of quinoa (McCouch *et al*. 2002; Mozo *et al*. 1999). The linkage maps reported here can also be used for cytogenetic studies. Several SSR primers amplified more than one segregating locus; of these, eight amplified two loci that both mapped in the segregating population. One of these pairs (QATG087-A, QATG087-B) mapped to the same linkage group, while all the others mapped to different linkage groups. In addition, the two 11S seed storage protein loci each mapped to different linkage groups. Thus, these linkage groups represent putative homoeologous chromosomes in the allotetraploid quinoa genome, and may be useful in cytological analyses and genome evolutionary studies.

The markers and map presented here will be particularly useful for laboratories in the regions of South America where quinoa is cultivated. Compared to other marker techniques, SSRs are relatively inexpensive once they have been developed, highly polymorphic, and easy to use. SSR markers are easily transferred between laboratories and are highly reproducible. These characteristics make them especially applicable in developing countries that may lack the resources required for other marker techniques.

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 (USDA-NPGS, Iowa State University) and A. Mujica (National University of the Altiplano, Puno, Peru), for taxonomic suggestions and seed contribution.

References

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

Broman K. W. 2005 The genomes of recombinant inbred lines. *Genetics* **169**, 1133-1146.

Christensen S. A., Pratt D. B., Pratt C., Stevens M. R., Jellen E. N., Coleman C. E. *et al*. 2007 Assessment of genetic diversity in the USDA and CIP-FAO international nursery collections of quinoa (*Chenopodium quinoa* Willd.) using microsatellite markers. *Plant. Genet. Res.* **5**, 82–95.

Cureton A. N., Burns M. J., Ford-Lloyd B. V. and Newbury H. J. 2002 Development of simple sequence repeat (SSR) markers for the assessment of gene flow between sea beet (*Beta vulgaris* ssp. *maritima*) populations. *Mol. Ecol. Notes* **2**, 402–403.

Danielsen S., Bonifacio A. and Ames T. 2003 Diseases of quinoa (*Chenopodium quinoa*). *Food Rev. Int*. **19**, 43–59.

Diwan N., McIntosh M. S. and Bauchan G. R. 1995 Methods of developing a core collection of annual *Medicago* species. *Theor. Appl. Genet*. **90**, 755–761.

Franco J. 2003 Parasitic nematodes of quinoa in the Andean region of Bolivia. *Food Rev. Int*. **19**, 77–85.

Hall M. C. and Willis J. H. 2005 Transmission ratio distortion in intraspecific hybrids of *Mimulus guttatus*: implications for genomic divergence. *Genetics* **170**, 375–386.

Jacobsen S.-E. 2003 The Worldwide potential for quinoa (*Chenopodium quinoa* Willd.) potential. *Food Rev. Int*. **19**, 167–177.

Jacobsen S.-E., Mujica A. and Jensen C. R. 2003 The resistance of quinoa (*Chenopodium quinoa* Willd.) to adverse abiotic factors. *Food Rev. Int*. **19**, 99–109.

Konieczny A. and Ausubel F. M. 1993 A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *Plant J*. **4**, 403–410.

Lu H., Romero-Severson J. and Bernardo R. 2002 Chromosomal regions associated with segregation distortion in maize. *Theor. Appl. Genet*. **105**, 622–628.

Mason S. L., Stevens M. R., Jellen E. N., Bonifacio A., Fairbanks D. J., Coleman C. E. *et al*. 2005 Development and use of microsatellite markers for germplasm characterization in quinoa (*Chenopodium quinoa* Willd.). *Crop Sci*. **45**, 1618–1630.

Maughan P. J., Bonifacio A., Jellen E. N., Stevens M. R., Coleman C. E., Ricks M. *et al*. 2004 A genetic linkage map of quinoa (*Chenopodium quinoa*) based on AFLP, RAPD, and SSR markers. *Theor. Appl. Genet*. **109**, 1188–1195.

Maughan P. J., Kolano B., Maluszynska J., Coles N. D., Bonifacio A., Rojas Beltran J. *et al*. 2006 Molecular and cytological characterization of ribosomal DNAs in *Chenopodium quinoa* and *Chenopodium berlandieri. Genome* (in press).

McCouch S. R., Teytelman L., Xu Y., Lobos K. B., Clare K., Walton M., Fu B. *et al*. 2002 Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Res*. **9**, 199–207.

Mozo T., Dewar K., Dunn P., Ecker J. R., Fischer S., Kloska S. *et al*. 1999 A complete BAC-based physical map of the Arabidopsis thaliana genome. *Nature Genet*. **22**, 271–275.

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

Ooijen J. W. van and Voorrips R. E. 2001 JoinMap 3.0, software for the calculation of genetic linkage maps. Plant Research International, Wageningen, Netherlands.

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

Paterson A. H., Lander E. S., Hewitt J. D., Peterson S., Lincoln S. E. and Tanksley S. D. 1988 Resolution of quantitative trait into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature* **335**, 721–726.

Pearsall D. 1992 The origins of plant cultivation in South America. In *The Origins of Agriculture* An International Perspective (ed. C. W. Cowan and P. J. Watson), pp. 173–205. Smithsonian Institution Press, Washington.

Prado R. E., Boero C., Gallard M. and Gonzalez J. A. 2000 Effect of NaCl on germination, growth, and soluble sugar content in *Chenopodium quinoa* Willd. seeds. *Bot. Bull. Acad. Sci*. **41**, 27–34.

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

Rasmussen C., Lagnaoui A. and Esbjerg P. 2003 Advances in the knowledge of quinoa pests. *Food Rev. Int*. **19**, 61–75.

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

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

Sambrook J., Fritsch E. E. and Maniatis T. 1989 Molecular cloning. A laboratory manual. 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Simmonds N. W. 1971 The breeding system of *Chenopodium quinoa*. I. Male sterility. *Heredity* **27**, 73–82.

Stam P. 1993 Construction of integrated genetic linkage maps by means of a new computer package: JOINMAP. *Plant J*. **3**, 739–744.

Staub J. E., Serquen F. C. and Gupta M. 1996 Genetic markers, map construction, and their application in plant breeding. *HortScience* **31**, 729–741.

Stevens M. R., Coleman C. E., Parkinson S. E., Maughan P. J., Zhang H. B., Balzotti M. R. *et al*. 2006 Construction of a quinoa (*Chenopodium quinoa* Willd.) BAC library and its use in identifying genes encoding seed storage proteins. *Theor. Appl. Genet*. **112**, 1593–1600.

Tanksley S. D. and McCouch S. R. 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. *et al*. 1979 Quinua y kañiwa: Cultivos andinos. CIID-IICA. Bogotá, Colombia.

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

Vacher J. J. 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 adaptation. *Agric. Ecosyst. Environ*. **68**, 99–108.

Vos P., Hogers R., Bleeker M., Reijans M., Van de Lee T., Hornes M. *et al*. 1995 AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res*. **23**, 4407–4414

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

Weber J. L. 1990 Informativeness of human (dC-dA)n. (dG-dT)n polymorphisms. *Genomics* **7**, 524–530.

Wilson H. D. 1988 Quinoa biosystematics I: Domesticated populations. *Econ. Bot*. **42**, 461–477

Wilson H. and Manhart J. 1993 Crop/weed gene flow: *Chenopodium quinoa* Willd. and C. *berlandieri* Moq. *Theor. Appl. Genet*. **86**, 642–648

Zamir D. and Tadmor Y. 1986 Unequal segregation of nuclear genes in plants. *Bot. Gaz*. **147**, 355–358.

Received 7 May 2007; accepted 31 October 2007

Published on the Web: 5 March 2008