Theor Appl Genet (2006) 112: 1593–1600
DOI 10.1007/s00122-006-0266-6

ORIGINAL PAPER

M. R. Stevens · C. E. Coleman · S. E. Parkinson
P. J. Maughan · H.-B. Zhang · M. R. Balzotti
D. L. Kooyman · K. Arumuganathan · A. Bonifacio
D. J. Fairbanks · E. N. Jellen · J. J. Stevens

Construction of a quinoa (*Chenopodium quinoa* Willd.) BAC library and its use in identifying genes encoding seed storage proteins

Received: 5 December 2005 / Accepted: 7 March 2006 / Published online: 4 April 2006

Abstract Quinoa (*Chenopodium quinoa* Willd.) is adapted to the harsh environments of the Andean Altiplano region. Its seeds have a well-balanced amino acid composition and exceptionally high protein content with respect to human nutrition. Quinoa grain is a staple in the diet of some of the most impoverished people in the world. The plant is an allotetraploid displaying disomic inheritance ($2n=4x=36$) with a di-haploid genome of 967 Mbp (megabase pair), or $2C=2.01$ pg. We constructed two quinoa BAC libraries using *Bam*HI (26,880 clones) and *Eco*RI (48,000 clones) restriction endonucleases. Cloned inserts in the *Bam*HI library average 113 kb (kilobase) with approximately 2% of the clones lacking inserts, whereas cloned inserts in the *Eco*RI library average 130 kb and approximately 1% lack inserts. Three plastid genes used as probes of high-density arrayed blots of 73,728 BACs identified approximately 2.8% of the clones as containing plastid DNA inserts. We estimate that the combined quinoa libraries represent at least 9.0 di-haploid nuclear genome equivalents. An average of 12.2 positive clones per probe were identified with 13 quinoa single-copy ESTs as probes of the high-density arrayed blots, suggesting that the estimate of 9.0× coverage of the genome is conservative. Utility of the BAC libraries for gene identification was demonstrated by probing the library with a partial sequence of the 11S globulin seed storage protein gene and identifying multiple positive clones. The presence of the 11S globulin gene in four of the clones was verified by direct comparison with quinoa genomic DNA on a Southern blot. Besides serving as a useful tool for gene identification, the quinoa BAC libraries will be an important resource for physical mapping of the quinoa genome.

Communicated by E. Guiderdoni

M. R. Stevens (✉) · C. E. Coleman · S. E. Parkinson · P. J. Maughan
M. R. Balzotti · D. J. Fairbanks · E. N. Jellen · J. J. Stevens
Department of Plant and Animal Sciences,
Brigham Young University, Provo, UT 84602-5157, USA
E-mail: mikel_stevens@byu.edu
Tel.: +1-801-4224032
Fax: +1-801-4220008
E-mail: craig_coleman@byu.edu
E-mail: sep@uclink.berkeley.edu
E-mail: pjmaughan@byu.edu
E-mail: weee42@yahoo.com
E-mail: daniel_fairbanks@byu.edu
E-mail: eric_jellen@byu.edu
E-mail: jolynn_stevens@byu.net

Present address: S. E. Parkinson
Department of Plant and Microbial Biology,
University of California Berkeley, 111 Koshland Hall,
Berkeley, CA 94720, USA

H.-B. Zhang
Department of Soil and Crop Sciences and Institute for Plant Genomics and Biotechnology, 2123 TAMU, Texas A&M University, College Station, TX 77843-2123, USA
E-mail: hbz7049@tamu.edu

D. L. Kooyman
Department of Physiology and Developmental Biology,
Brigham Young University, Provo, UT 84602, USA
E-mail: david_kooyman@byu.edu

K. Arumuganathan
Flow and Image Cytometry Core Laboratories,
Benaroya Research Institute at Virginia Mason,
1201 Ninth Avenue, Seattle, WA 98101, USA
E-mail: karu@benaroyaresearch.org

A. Bonifacio
The Foundation for the Promotion and Investigation of Andean Products (PROINPA), La Paz Bolivia
E-mail: quinua@mail.megalink.com

Introduction

Quinoa (*Chenopodium quinoa* Willd.) is an allotetraploid plant species displaying disomic inheritance (Simmonds 1971; Risi and Galwey 1984; Ward 2000) and belongs to the Amaranthaceae family in the subfamily Chenopodioideae. Other agronomically important plant species in this subfamily are spinach (*Spinacea oleracea* L.), amaranth (*Amaranthus* spp.), sugar beet (*Beta vulgaris* L.) and a series of Eurasian and American *Chenopodium*

species of local importance as seed and vegetable crops (Partap et al. 1998). Quinoa grain is an important source of protein for the families of subsistence farmers on the high Andean plateau (the Altiplano). The seed is also grown for commercial production throughout the western regions of South America, including Bolivia, Chile, Ecuador and Peru for domestic markets and emerging export markets in Japan, Australia, Europe and North America.

Renewed interest in quinoa as a cultivated crop is centered on two important factors. First, the plant is remarkably well adapted to harsh conditions such as the low water content and high alkalinity of soils (pH > 8.0) in the region surrounding Lake Uyuni of the Potosi Department of Bolivia. Additionally, some varieties of quinoa can withstand dramatic fluctuations in temperatures (including frequent frost) common at elevations of 3,500–3,850 m above sea level (Risi and Galwey 1984; Vacher 1998; Prado et al. 2000). Second, quinoa seed is reported to be one of the most nutritionally well-balanced plant foods under cultivation. It contains a higher percentage of protein (7.5–22.1%) (Tapia et al. 1979) than any of the cereal grains and it has an excellent balance of carbohydrates, lipids, and protein (Risi and Galwey 1984; Coulter and Lorenz 1990; Chauhan et al. 1999). The FAO recommendations for human nutrition are exceeded for all essential amino acids in quinoa grain (Ruas et al. 1999).

Recent studies are leading to a better understanding of the quinoa genome and the development of an important set of genetic tools which can be used to enhance future breeding efforts. Over 400 simple sequence repeats (SSR) markers (Mason et al. 2005) have been characterized and a subset of these along with AFLP and RAPD markers have been placed on a preliminary genetic linkage map (Maughan et al. 2004). The creation of cDNA libraries from seed and floral tissue has led to the identification of 424 ESTs and 51 SNPs (Coles et al. 2005).

An important genetic tool that is useful for correlating physical and genetic maps and for the identification of specific genes is a bacterial artificial chromosome (BAC) library. Libraries have been developed for a number of plant species of major and minor economic/agricultural impact; examples include durum wheat (Cenci et al. 2003), coffee (Noir et al. 2004), pepper (Yoo et al. 2003), and ginseng (Hong et al. 2004). These BAC libraries can facilitate physical mapping, synteny evaluation, gene cloning and many other types of studies (Monaco and Larin 1994).

Here, we report the construction of a BAC library containing 74,496 clones representing approximately 9.0 di-haploid equivalents of the quinoa genome. In characterizing our BAC library we have determined the average insert size, calculated the frequency of a set of ESTs and genomic fragments within the library and screened the library for the genes encoding the 11S seed storage protein. The knowledge gained from in-depth studies of this BAC library and the resulting in-depth studies on the genetics of quinoa will have positive agricultural implications for the socioeconomic base of the Andean people. Furthermore, an understanding of a crop with such unique environmental adaptive characteristics and excellent human nutritive value will enhance our ability to broaden the base of the world food supply.

Materials and methods

Plant material, DNA extraction and genome sizing

Tissue for the high-molecular-weight (HMW) DNA needed for the construction of the BAC libraries was isolated from the first two true leaves and apex of 2–3-week-old plants of commercially available quinoa 'Real' from Bolivia. Because of its commercial nature, the specific accession of the landrace series known as 'Real' is unknown. 'Real' is an ecotype consisting of multiple genotypes with similar seed characteristics from the Potosi Department of Bolivia. Tissue harvested from these plants was grown under greenhouse conditions at Brigham Young University (BYU). The harvested tissue was immediately frozen and maintained either on dry ice or in a −80°C freezer until utilized for BAC library construction. Quinoa genome size was determined for the nuclear DNA content with flow cytometry according to Arumuganathan and Earle (1991) at the University of Nebraska Flow Cytometry Core Research Facilities (FACScan flow cytometer, Becton Dickinson Immunocytormetry system, San Jose, CA).

BAC library construction

Two separate libraries were constructed by digesting HMW quinoa DNA either with *Bam*HI or *Eco*RI and cloning it into the vector pECBACI [7.5 kb (kilobase)] (for vector map and sequence, see http://www.hbz.tamu.edu) according to Zhang (2000) with minor modifications. Nuclear DNA was obtained from 70 to 100 g of tissue and embeded in 2.5 ml of low-melting-point agarose plugs. The HMW DNA, embedded in LMP agarose plugs, was partially digested with *Bam*HI or *Eco*RI and size-fractionated twice by separation in agarose CHEF gels (Bio-Rad Laboratories, Hercules, CA, USA). The libraries were constructed with approximately 150 ng HMW digested DNA and was used at a rate of 0.6–1.0 ng/μl in ligation reactions of linearized de-phosphorylated pECBAC1 at the recommended 1:4 molecular weight ratio (vector to HMW DNA). ElectroMax DH10B cells (Invitrogen, Carlsbad, CA) were transformed with ligation mix by electroporation with a BioRad (Hercules, CA) Gene Pulser at a voltage of 2.5 kV, a capacitance of 25 μF, and an impedance of 100 Ω, all according to manufacturer's specifications. Cells were plated on LB agar containing 12.5 μg/ml chloramphenicol, 15 μg/ml IPTG, and 60 μg/ml X-gal and incubated at 37°C for 24–36 h. White colonies were manually picked and transferred directly to 384-well

microtiter plates containing 50 µl LB freezing medium with 12.5 µg/ml of chloramphenicol in each well (Zhang et al. 1996). After incubating at 37°C for 24 h microtiter plates were placed in a −80°C freezer for long-term storage.

BAC library insert characterization and screening

BAC DNA was isolated from 200 and 280 clones from the *Bam*HI and *Eco*RI libraries, respectively, by an alkaline lysis method, digested with *Not*I, and separated on 1% agarose CHEF gels. The insert size was determined by comparison with a molecular-weight ladder on the gel and an average insert size was calculated for the set of clones tested from each library.

Colonies containing individual BACs were double-spotted in high-density arrays on nylon membranes using a Genetix Q-Bot at the Arizona Genomics Institute (Tucson, AZ, USA) as described in Tomkins et al. (1999b). Radiolabeled probes were prepared from PCR products by a random priming method using a Prime-A-Gene® kit (Promega, Madison, WI, USA). Nuclear-gene probes were derived from cDNAs (Coles et al. 2005) or from genomic DNA for the 5S rDNA and 45S intergenic spacer (IGS) rDNA probes (unpublished data). The organellar gene probe was derived from a mixture of three barley (*Hordeum vulgare* L.) plastid genes *ndhA*, *rbcL*, and *psbA* (Chen et al. 2004). Hybridizations and the identification of addresses for positive clones were performed using the protocol described at http://www.genome.arizona.edu/information/protocols/addressnew.html. The hybridized blots were visualized by exposure to a phosphorescent screen followed by scanning the screen with a Molecular Imager FX Pro MultImager System and Quantity One software (Bio-Rad). The nylon membranes were stripped of probe between hybridizations by immersion in a boiling 0.1% SDS solution and allowing them to cool to room temperature.

Southern blotting

BAC DNA containing the 11S seed storage protein gene was prepared from positive clones identified by hybridization using a BACMAX™ DNA purification kit (Epicentre, Madison, WI, USA). Genomic DNA was extracted from quinoa plants grown in the greenhouse using a protocol reported elsewhere (Saghai-Maroof et al. 1984). BAC and genomic DNAs were digested using either *Hin*dIII or *Eco*RI restriction enzymes and the fragments were separated on a 0.8% agarose gel by electrophoresis. The gel was blotted to Hybond™-N^+ positively charged nylon membrane (Amersham, Piscataway, NJ USA) using downward capillary transfer and the blot was hybridized and washed according to the manufacturer's instructions. Preparation of the 11S seed storage protein gene probe and visualization of the hybridized blot was the same as described previously.

Results

Estimating quinoa nuclear DNA content

The chromosome number for *C. quinoa* is $2n=4x=36$ as reported by Giusti (1970). These chromosomes are very small, mostly isobrachial and basically without distinguishing features to enable traditional karyotyping (Parkinson 2001). Using the flow cytometry method reported by Arumuganathan and Earle (1991), we determined the di-haploid genome (1C) for the landrace series quinoa 'Real' to be 967 Mbp or 2.01 pg/2C.

BAC library construction

To develop a BAC library that best represents the complete quinoa genome, we partially endonuclease-digested HMW quinoa DNA with two enzymes, *Bam*HI and *Eco*RI. The DNA size-range collected and subjected to a second size selection was between 130–350 and 150–400 kb fragments for *Bam*HI and *Eco*RI, respectively. For both libraries, we divided the first size selection smear into three pieces (*Bam*HI approximately 130–195, 195–285, and 285–350 kb; *Eco*RI approximately 150–220, 220–330, and 330–400 kb). Subjecting these DNAs to a second CHEF-gel size selection resulted in a preponderance of the cloned fragments ranging near the selected size (Fig. 1). The most efficient cell transformations, in both libraries, were from the smallest size selection (*Bam*HI approximately 130–195 and *Eco*RI approximately 150–220 kb).

Although several ligation reactions were performed for both the *Bam*HI and *Eco*RI libraries only one was used to construct the *Bam*HI library and three ligation reactions were used to develop the *Eco*RI library. Our transformation efficiency for the *Bam*HI library was approximately 460 white recombinant clones per 1 µl of ligation reaction. A total of 26,880 (70×384 well microtiter plates) clones were harvested for this library. When analyzing the *Eco*RI library, we found an average of approximately 552 white recombinant clones per 1 µl of ligation reaction across the three ligation reactions. A total of 48,000 (125×384 well microtiter plates) clones were harvested for this library.

Characterization of the quinoa BAC libraries

We analyzed 200 randomly selected clones from the *Bam*HI library by digesting the DNA isolated from overnight cultures of each clone with *Not*I and electrophoresing the resulting products on a CHEF gel. These results show that the average insert size in the *Bam*HI library is approximately 113 kb, ranging in size from 21 to 288 kb (Fig. 1) with approximately 2% (4 of 200 clones) lacking an insert. The same procedure was repeated for the *Eco*RI library and we determined the average insert size as approximately 130 kb ranging from 8 to >300 kb (Fig. 1), with approximately 1% (3 of 280 clones) lacking

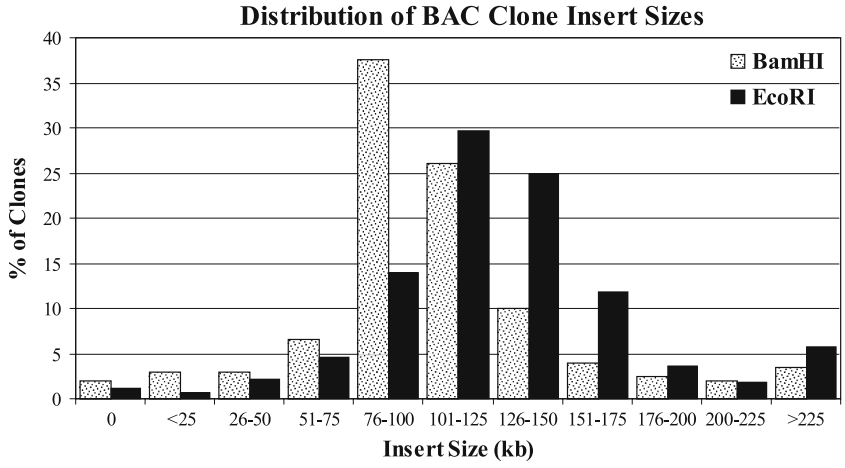

Fig. 1 Insert size distribution of 200 *Bam*HI and 280 *Eco*RI randomly selected clones of the respective quinoa BAC libraries

an insert. We found a high proportion, approximately 44 and 38% of the *Bam*HI and *Eco*RI libraries, respectively, to have two or more bands, indicating the presence of *Not*I recognition sites within the quinoa DNA insert (Fig. 2). The combined quinoa BAC libraries have 74,880 clones with approximately 1.5% lacking an insert and an overall average insert size of approximately 123 kb.

To characterize the combined BAC libraries and facilitate clone identification, we prepared high-density arrays of 73,728 clones on nylon membranes. We hybridized the membranes to a series of probes representing the chloroplast genome, rDNA and low-copy DNA sequences as summarized in Table 1. Hybridization with a probe representing three barley plastid genes gave 2,035 positive hits, which represents 2.8% of the library and is consistent with chloroplast DNA content in libraries reported for other plant species (Zhang et al. 1996; Tomkins et al. 1999b; Luo et al. 2001; Nilmalgoda et al. 2003). Next, we hybridized the blots with a series of probes representing the 5S rRNA gene, the 45S rDNA IGS. The high number of positive clones on the BAC arrays reflects the expected repetitive nature of the targets for these probes in the quinoa genome. Finally, we hybridized the blots using DNA probes representing sequences predicted to be low-copy within the quinoa genome (probes 4–16 in Table 1). The average number of positive hits for these hybridizations was 12.2.

Isolating candidate BAC clones for the 11S protein

An important agronomic feature of quinoa grain is its high nutritive value, including a high level of protein and an excellent balance of amino acids. We reasoned that identification and characterization of seed storage protein genes from the quinoa genome would be an important initial step toward understanding how protein is stored in the seed. A clone with homology to published 11S seed storage protein genes from other plant species was obtained from a developing quinoa seed cDNA library (C. Coleman, unpublished results) and used as a probe to hybridize the high-density array of quinoa BAC clones. From this hybridization, we identified ten positive clones and extracted BAC DNA from four of them.

The BAC DNA from the four positive clones and quinoa genomic DNA were endonuclease-digested with either *Eco*RI or *Hin*dIII and electrophoresed on an agarose gel for Southern blotting as shown in Fig. 3. The probe for the blot was prepared by amplifying a portion of the quinoa 11S gene using gene-specific primers and one of the BAC clones as template. The genomic DNA

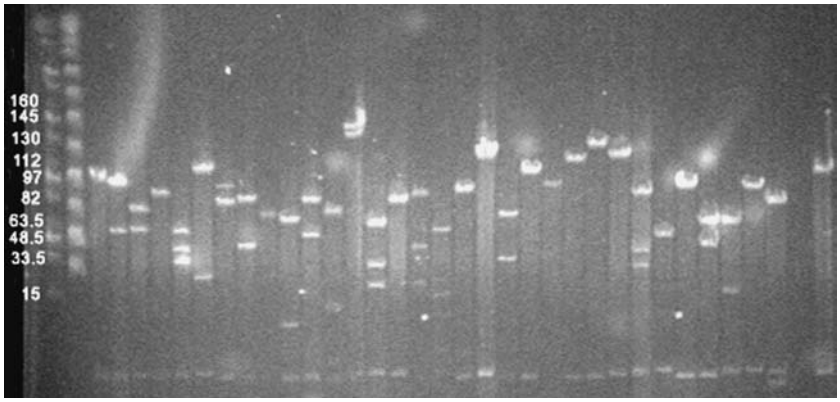

Fig. 2 An ethidium bromide-stained CHEF gel of 40 random clones from the quinoa *Bam*HI library. The inserts were released from the 7.5-kb pECBACI plasmid (noted as common band at the base of the gel) by digestion with *Not*I. In this gel example, 15 clones have clear internal *Not*I cut sites. *Lanes 1* and *2* of the gel are New England BioLabs Mid-Range I and II size markers, respectively. The sizes of the markers are given in kilobase

Table 1 Summary of the number of positive clones observed following hybridization of a high-density array of quinoa BAC clones with various probes

Probe	GenBank accession no.	Probe description or putative function	Number of positive signals
1	NA	Barley *ndhA, rbcL, psbA*	2,035
2	NA	5S rRNA	84
3	NA	45S rDNA IGS	907
4	AY562550	11S seed storage protein	10
5	CN782308	Translationally controlled tumor family protein	13
6	CN782300	Sucrose synthase	18
7	CN782239	Shepherd protein (SHD)/ clavata formation protein	14
8	CN782189	Osmotin-like protein	8
9	CN782146	Calcium-transporting ATPase	20
10	CN781947	Mago nashi family protein	10
11	CN782099	ATP citrate lyase	11
12	CN782063	Cinnamyl alcohol dehydrogenase	7
13	CN782259	Early nodulin-related protein	10
14	CN782051	UDP-glucose 4-epimerase	15
15	CN782305	Beta-ketoacyl-CoA synthase	13
16	CN781953	Phototropic-responsive NPH3 family protein	10

NA not available

used on the Southern blot was extracted from quinoa cultivar 'Chucapaca' (lanes 3 and 4), land race KU−2 (lanes 5 and 6), land race NL-6 (lanes 7 and 8) and breeding line 0654 (lanes 9 and 10). These accessions were chosen for this assay because of their use as parents of populations used to generate a genetic map of quinoa (Maughan et al. 2004). Although only one of the BAC clones is shown on the blot, the four BAC clones from which DNA was obtained showed identical banding patterns on Southern blots. The blot shows a single band hybridizing to the 11S probe in the lane containing BAC DNA digested with either *Eco*RI or *Hin*dIII (lanes 1 and 2). When digested with *Eco*RI, the size of the single hybridized band in the BAC DNA lane is similar to hybridized bands of genomic DNA from Chucapaca, KU-2 and 0654 (compare lane 1 with lanes 3, 5 and 9). Likewise, the single hybridized BAC DNA band from the *Hin*dIII digestion is similar in size to hybridized genomic bands from Chucapaca and 0654 (compare lane 2 with lanes 4 and 10). At least two bands are present in each of the lanes containing genomic DNA and they are polymorphic with respect to the the four quinoa accessions. A banding pattern similar to Chucapaca and 0654 was seen on a Southern blot containing genomic DNA extracted from the landrace 'Real', which was used in the construction of the BAC library.

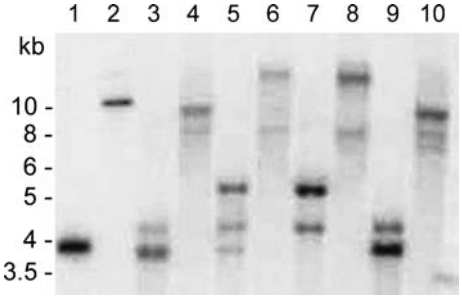

Fig. 3 Southern blot of quinoa BAC DNA containing an 11S seed storage protein gene compared to genomic DNA from four quinoa accessions. *Lanes 1*, *2* BAC clone 82M10; *lanes 3*, *4* quinoa 'Chucapaca'; *lanes 5*, *6* quinoa land race KU-2; *lanes 7*, *8* quinoa land race NL-6; *lanes 9*, *10* quinoa breeding line 0654. DNA in *lanes 1*, *3*, *5*, *7*, and *9* were digested with *Eco*RI. DNA in *lanes 2*, *4*, *6*, *8*, and *10* were digested with *Hin*dIII. The blot was probed with a fragment of DNA containing the quinoa 11S seed storage protein gene amplified from BAC clone 77L9. The sizes of DNA markers given in kilobase (kb) are shown on the *left*

Discussion

We have created a BAC library representing the genome of the Andean crop plant *C. quinoa* and made it publically available through the Arizona Genomics Institute (Tucson, AZ, USA). The library contains a total of 74,880 clones of which approximately 70,955 contain nuclear-genome inserts based on our estimation of 1.5% clones lacking an insert and 2.8% containing plastid DNA. With an average insert size of approximately 123 kb, it is estimated that the combined quinoa library represents 9.0 di-haploid equivalents of the 967 Mbp quinoa genome. The quinoa BAC library consists of two separate libraries constructed using DNA fragments generated from different restriction enzyme digests. It has been shown that the use of restriction endonucleases which differ in the GC content of their recognition sequences, such as *Eco*RI (GAATTC) and *Bam*HI (GGATCC), can improve the overall representation of the genome in the library (Chen et al. 2004).

When we probed high-density arrays of the library with low-copy DNA sequences, the average number of positive clones identified was 12.2, which is slightly higher than the estimated genome coverage of the library. This discrepancy may be due to an underestimation of the genome coverage of the library or it may reflect the presence of duplicate genetic loci for some of the genes in the quinoa genome, consistent with its allotetraploid genome (Simmonds 1971; Risi and Galwey 1984; Ward 2000). The probes we used for the array hybridizations were chosen based on our ability to obtain PCR amplification products that appeared as single bands following agarose gel electrophoresis. Furthermore, DNA sequencing of the PCR products used as probes yielded no ambiguities, suggesting that each represents a single genetic locus. Although this procedure does not preclude cross-hybridization of our probes to DNA targets representing a duplicated locus, we believe that it significantly reduces the probability of such. Only two of the thirteen low-copy probes hybridized to 18 or more clones on the arrays (Table 1, probes 6 and 9) and the number of positive hits for just two other probes (7 and 14) was closer to

eighteen than to nine. Thus, we believe that 9.0 genome equivalents for the combined quinoa BAC libraries is a conservative estimate.

When probed with 45S and 5S rDNA sequences from quinoa, the library was found to contain 907 and 84 clones, respectively (Table 1), which represent the single 45S rDNA locus and two 5S rDNA loci in the quinoa genome (Kolano et al. 2001). Because each monomeric 45S rDNA repeat is approximately 9 kb long and each monomeric 5S rDNA repeat averages slightly over 300 bp (unpublished data), we conclude that there are approximately 1,500 tandem repeats at the 45S rDNA locus and approximately 4,000 tandem repeats at the two 5S rDNA loci.

A surprising feature of the quinoa library is the relatively high number of cloned inserts that contain *Not*I restriction enzyme recognition sites. Our finding of one or more internal *Not*I sites in 44% of the clones from the *Bam*HI library and 38% of the clones from the *Eco*RI library differs from the relatively low number of clones containing internal *Not*I sites reported for BAC libraries prepared from other dicotyledonous species. An examination of dicot BAC library literature using ether gel photos and/or reported data suggest that the average dicot internal *Not*I site ranges between > 0 and 11%. These data are from nine BAC libraries; *Arabidopsis* (Choi et al. 1995), coffee (Noir et al. 2004), cucumber (Nam et al. 2005), ginseng (Hong et al. 2004), melon (Luo et al. 2001), *Medicago truncatula* (Nam et al. 1999), pepper (Yoo et al. 2003), potato (Chen et al. 2004), and soybean (Tomkins et al. 1999b; Danesh et al. 1998). The 14–26% of clones containing internal *Not*I sites in a citrus and a lettuce BAC library are the highest reported for any dicotyledonous species (Deng et al. 2001; Frijters et al. 1997). For libraries of monocotyledonous *Poaceae* species, including wheat (Cenci et al. 2003) the D genome of durum wheat (Moullet et al. 1999), maize (Tomkins et al. 2002), rice (Zhang et al. 1996; Yang et al. 1997), sorghum (Woo et al. 1994), and sugarcane (Tomkins et al. 1999a) the proportion of BAC clones containing inserts with internal *Not*I sites ranged from 72 to 91%. Two exceptions are for a wheat library reported by Nilmalgoda et al. (2003) to have 41% of its clones containing internal *Not*I sites and a barley library reported by Yu et al. (2000) which shows only 54% of clones with internal *Not*I sites. Thus, the level of internal *Not*I sites in the quinoa library reported here is higher than that reported for other dicotyledonous species but does not reach the level reported for many of the monocotyledonous cereal grains.

The difference in the frequency of the GC-rich *Not*I recognition site (GCGGCCGC) between monocot and dicot species is commonly attributed to the difference in the overal GC content of the genome, which is reportedly higher for monocot species than for dicot species (Salinas et al. 1988; Kawabe and Miyashita 2003). However, a survey of 100 quinoa EST sequences and 35 quinoa genomic sequences reveals that the overall GC content for these quinoa sequences is approximately 43% (data not shown), which is similar to that reported for other dicot species (Kawabe and Miyashita 2003). Thus, the elevated frequency of *Not*I recognition sites in the quinoa genome must be attributed to a cause other than high GC content in the genome, such as the presence of a highly repetitive element containing the *Not*I recognition motif.

The utility of the quinoa BAC library for gene identification is demonstrated by the isolation of specific BAC clones that encode a homolog of an 11S seed storage globulin gene. Four BAC clones identified from the library were shown to have a single 4 kb *Eco*RI fragment and a single 12 kb *Hin*dIII fragment containing 11S globulin gene sequences based on Southern hybridization (Fig. 3). These 4-kb *Eco*RI and 12-kb *Hin*dIII fragments are approximately the same size as fragments of genomic DNA derived from the highland quinoa Chucapaca and 0654 as seen on the same Southern blot. The relatively small size of the *Eco*RI fragment and the absence of any other hybridizing fragments, suggests that the 11S globulin gene is present as a single copy at the locus contained within these BAC clones. Additional fragments visible in the lanes containing genomic DNA from the highland germplasm already mentioned and the lowland land races, KU-2 and NL-6, suggest that the quinoa 11S globulin gene that we have cloned is likely part of a multigene family. Mutliple copies of 11S globulin genes in other species have been reported, such as in pea (Domoney and Casey 1985; Domoney et al. 1986), bean (Talbot et al. 1984) and fava bean (Wobus et al. 1986; Heim et al. 1989). Furthermore, the pattern of hybridized bands differs between the four quinoa varieties represented on the Southern blot, suggesting that the 11S loci may be used as markers which can be placed on the emerging genetic linkage map.

The quinoa BAC library is an important genomic resource which can be utilized for the improvement of crop production for subsistence farmers on the Altiplano and commercial growers throughout the Andean region. One of our immediate goals for these libraries is the assembly of contigs for generation of a physical map. The physical map will be an important step in unraveling the allotetraploid ancestry of *C. quinoa* to better understand its relationship with wild relatives that are found throughout the region where quinoa and its domesticated tetraploid Mexican cousin, *C. berlandieri* ssp. *nuttaliae*, are cultivated. Moreover, physical mapping will allow us to do synteny studies utilizing the DNA sequence information from other crops. We envision that the genomic resources developed for quinoa may be useful in comparative analyses with closely related species of economic importance such as sugar beet, spinach, amaranth and an array of other domesticated *Chenopodium* species from Eurasia (Partap et al. 1998).

Acknowledgements We acknowledge the contribution to this work by Jared Tyler. This research was made possible by financial support from the McKnight Foundation-CCRP, Holmes Family Foundation, and the Ezra Taft Benson Agriculture and Food Institute and germplasm contributions by PROINPA (La Paz, Bolivia) and the

National University of the Altiplano/FAO (Puno, Peru). The work reported here was carried out in compliance with laws governing genetic experimentation in the United States.

References

Arumuganathan K, Earle ED (1991) Estimation of nuclear DNA content of plants by flow cytometry. Plant Mol Biol Rep 9:229–241

Cenci A, Chantret N, Kong X, Gu Y, Anderson OD, Fahima T, Distelfeld A, Dubcovsky J (2003) Construction and characterization of a half million clone BAC library of durum wheat (*Triticum turgidum* ssp. *durum*). Theor Appl Genet 107:931–939

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

Chen Q, Sun S, Ye Q, McCuine S, Huff E, Zhang H-B (2004) Construction of two BAC libraries from the wild Mexican diploid potato, *Solanum pinnatisectum*, and the identification of clones near the late blight and Colorado potato beetle resistance loci. Theor Appl Genet 108:1002–1009

Choi S, Creelman RA, Mullet JE, Wing RA (1995) Construction and characterization of a bacterial artificial chromosome library of *Arabidopsis thaliana*. Plant Mol Biol Rep 13:124–128

Coles ND, Coleman CE, Christensen SA, Jellen EN, Stevens MR, Bonifacio A, Rojas-Beltran JA, Fairbanks DJ, Maughan PJ (2005) Development and use of an expressed sequenced tag library in quinoa (*Chenopodium quinoa* Willd.) for the discovery of single nucleotide polymorphisms. Plant Sci 168:439–447

Coulter L, Lorenz K (1990) Quinoa—composition, nutritional value, food applications. Food Sci Technol 23:203–207

Danesh D, Peñuela S, Mudge J, Denny RL, Nordstrom H, Martinez JP, Young ND, (1998) A bacterial artificial chromosome library for soybean and identification of clones near a major cyst resistance gene. Theor Appl Genet 96:196–202

Deng Z, Tao Q, Chang Y-L, Huang S, Ling P, Yu C, Chen C, Gmitter Jr FG, Zhang H-B (2001) Construction of a bacterial artificial chromosome (BAC) library for citrus and identification of BAC contigs containing resistance gene candidates. Theor Appl Genet 102:1177–1184

Domoney C, Casey R (1985) Measurement of gene number for seed storage proteins in *Pisum*. Nucl Acids Res 13:687–699

Domoney C, Ellis THN, Davies DR (1986) Organization and mapping of legumin genes in *Pisum*. Mol Gen Genet 202:280–285

Frijters ACJ, Zhang Z, van-Damme M, Wang G-L, Ronald PC, Michelmore RW (1997) Construction of a bacterial artificial chromosome library containing large *Eco*RI and *Hin*dIII genomic fragments of lettuce. Theor Appl Genet 94:390–399

Giusti L (1970) El género *Chenopodium* en Argentina: I. Números de cromosomas. Darwiniana 16:98–105

Heim U, Schubert R, Bäumlein H, Wobus U (1989) The legumin gene family: structure and evolutionary implications of *Vicia faba* B-type genes and pseudogenes. Plant Mol Biol 13:653–663

Hong CP, Lee SJ, Park JY, Plaha P, Park YS, Lee YK, Choi JE, Kim KY, Lee JH, Lee J, Jin H, Choi SR, Lim YP (2004) Construction of a BAC library of Korean ginseng and initial analysis of BAC-end sequences. Mol Genet Genomics 271:709–716

Kawabe A, Miyashita NT (2003) Patterns of codon usage bias in three dicot and four monocot plant species. Genes Genet Syst 78:343–352

Kolano B, Pando LG, Maluszynska J (2001) Molecular cytogenetic studies in *Chenopodium quinoa* and *Amaranthus caudatus*. Acta Soc Bot Pol 70:85–90

Luo M, Wang Y-H, Frisch D, Joobeur T, Wing RA, Dean RA (2001) Melon bacterial artificial chromosome (BAC) library construction using improved methods and identification of clones linked to the locus conferring resistance to melon Fusarium wilt (*Fom-2*). Genome 44:154–162

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

Maughan PJ, Bonifacio A, Jellen EN, Stevens MR, Coleman CE, Ricks M, Mason SL, Jarvis DE, Gardunia BW, Fairbanks DJ (2004) A genetic linkage map of quinoa (*Chenopodium quinoa*) based on AFLP, RAPD, and SSR markers. Theor Appl Genet 109:1188–1195

Monaco AP, Larin Z (1994) YACs, BACs, PACs and MACs: artificial chromosomes as research tools. Trends Biotechnol 12:280–286

Moullet O, Zhang H-B, Lagudah ES (1999) Construction and characterization of a large DNA insert library from the D genome of wheat. Theor Appl Genet 99:305–313

Nam Y-W, Penmetsa RV, Endre G, Uribe P, Kim D, Cook DR (1999) Construction of a bacterial artificial chromosome library of *Medicago truncatula* and identification of clones containing ethylene-response genes. Theor Appl Genet 98:638–646

Nam Y-W, Lee J-R, Song K-H, Lee M-K, Robbins MD, Chung S-M, Staub JE, Zhang H-B (2005) Construction of two BAC libraries from cucumber (*Cucumis sativus* L.) and identification of clones linked to yield component quantitative trait loci. Theor Appl Genet 111:150–161

Nilmalgoda SD, Cloutier S, Walichnowski AZ (2003) Construction and characterization of a bacterial artificial chromosome (BAC) library of hexaploid wheat (*Triticum aestivum* L.) and validation of genome coverage using locus-specific primers. Genome 46:870–878

Noir S, Patheyron S, Combes M-C, Lashermes P, Chalhoub B (2004) Construction and characterisation of a BAC library for genome analysis of the allotetraploid coffee species (*Coffea arabica* L.). Theor Appl Genet 109:225–230

Parkinson SE (2001) Cytogenetic studies and construction of a bacterial artificial chromosome library from *Chenopodium quinoa* (Willd.). MS thesis. Agronomy and Horticulture, Provo, UT

Partap T, Joshi BD, Galwey NW (1998) Chenopods. *Chenopodium* spp. Promoting the conservation and use of underutilized and neglected crops. 22. Institute of Plant Genetics and Crop Plant Research/International Plant Genetic Resources Institute, Gatersleben, Germany/Rome, Italy

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

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

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

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

Salinas J, Matassi G, Montero LM, Bernardi G (1988) Compositional compartmentalization and compositional patterns in the nuclear genomes of plants. Nucl Acids Res 16:4269–4285

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

Talbot DR, Adang MJ, Slightom JL, Hall TC (1984) Size and organization of a multigene family encoding phaseolin, the major seed storage protein of *Phaseolus vulgaris* L. Mol Gen Genet 198:42–49

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, Columbia

Tomkins JP, Yu Y, Miller-Smith H, Frisch DA, Woo SS, Wing RA (1999a) A bacterial artificial chromosome library for sugarcane. Theor Appl Genet 99:419–424

Tomkins JP, Mahalingam R, Smith H, Goicoechea JL, Knap HT, Wing RA (1999b) A bacterial artificial chromosome library for

soybean PI 437654 and identification of clones associated with cyst nematode resistance. Plant Mol Biol 41:25–32

Tomkins JP, Davis G, Main D, Yim Y, Duru N, Musket T, Goicoechea JL, Frisch DA, Coe EH Jr, Wing RA (2002) Construction and characterization of a deep-coverage bacterial artificial chromosome library for maize. Crop Sci 42:928–933

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

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

Wobus U, Bäumlein H, Bassüner R, Heim U, Jung R, Müntz K, Saalbach G, Weschke W (1986) Characteristics of two types of legumin genes in the field bean (*Vicia faba* L. var. minor) genome as revealed by cDNA analysis. FEBS Lett 201:74–80

Woo S-S, Jiang J, Gill BS, Paterson AH, Wing RA (1994) Construction and characterization of a bacterial artificial chromosome library of *Sorghum bicolor*. Nucl Acids Res 22:4922–4931

Yang D, Parco A, Nandi S, Subudhi P, Zhu Y, Wang G, Huang N (1997) Construction of a bacterial artifical chromosome (BAC) library and identification of overlapping BAC clones with chromosome 4-specific RFLP markers in rice. Theor Appl Genet 95:1147–1154

Yoo EY, Kim S, Kim YH, Lee CJ, Kim B-D (2003) Construction of a deep coverage BAC library from *Capsicum annuum*, 'CM334'. Theor Appl Genet 107:540–543

Yu Y, Tomkins JP, Waugh R, Frisch DA, Kudrna D, Kleinhofs A, Brueggeman RS, Muehlbauer GJ, Wise RP, Wing RA (2000) A bacterial artificial chromosome library for barley (*Hordeum vulgare* L.) and the identification of clones containing putative resistance genes. Theor Appl Genet 101:1093–1099

Zhang H-B (2000) Construction and manipulation of large-insert bacterial clone libraries manual. Texas A&M University, College Station, TX

Zhang H-B, Choi S, Woo S-S, Li Z, Wing RA (1996) Construction and characterization of two rice bacterial artificial chromosome libraries from the parents of a permanent recombinant inbred mapping population. Mol Breed 2:11–24