

Full Length Research Paper

Construction and characterization of a partial binary bacterial artificial chromosome (BIBAC) of *Agave tequilana* var. azul (2X) and its application for gene identification

Tamayo-Ordoñez, M., Rodríguez-Zapata, L. C. and Sánchez-Teyer, L. F.*

Unidad de Biotecnología, Centro de Investigación Científica de Yucatán A.C., C43 #130 Col. Chuburná de Hidalgo. Mérida, Yucatán, C. P. 97200, México.

Accepted 24 October, 2012

The structure and organization of the genome of *Agave* is still unknown. To provide a genomic tool for searching sequences of the genus, we built and characterized a binary (BIBAC2) genomic library of *Agave tequilana* Weber var. azul. Clones of the library had an average insert size of 170 Kb. The frequency of inserts with internal Not I sites was 30% and only 5% of the library showed organelle contamination. The library was assessed using probes with high homology to repeated regions (retroelements and rDNA regions), genes involved in the resistance to diseases (NBS-LRR) and genes related to late embryogenesis (LEA). Recombinant clones that hybridized with each of the probes were identified. Our results indicate that the obtained genomic library is suitable for the identification of sequences of interest, for genetic mapping and for studies of gene regulation and expression.

Key words: Binary bacterial artificial chromosome (BIBAC), *Agave tequilana*, genome, clones.

INTRODUCTION

Genomic libraries are derived from the cloning of DNA fragments into vectors. According to their origin, vectors differ in their properties and in the size of DNA fragments they allow to be inserted and are classified as plasmids, cosmids, phages, yeast artificial chromosomes (YAC), P1-derived artificial chromosome (PAC) and bacterial artificial chromosomes (BAC). BAC vectors are currently the most used for genomic libraries because they enable replication and integration of large fragments (Cavagnaro et al., 2009; Akhunov et al., 2005). Important features of this type of DNA delivery is the *Escherichia coli* factor F' allowing for strict control of copy number and unidirectional DNA replication, the latter property promoting plasmid stability compared to YAC having poor stability (Willems and Skurry, 1987).

BAC technology has had various applications, such as the transformation of clusters of genes of interest in plant cells (Hamilton et al., 1996, 1999; Liu et al., 1999, 2002; He et al., 2003), the construction of contigs (Marra et al., 1997; Boysen et al., 1997), development of molecular markers (Wang et al., 1996; Nakamura et al., 1997; Yang et al., 1997; Cai et al., 1998; Danesh et al., 1998) and even the isolation and molecular characterization of repetitive DNA sequences. Given the wide range of applications of the BAC vectors, these have been developed in a variety of important crops such as tomato (Choi et al., 1995), maize (Mozo et al., 1998; Folkertsma et al., 1999), soybean (Salimath and Bhattacharyya, 1999; Danesh et al., 1998; Tomkins et al., 1999b), rice (Wang et al., 1995) sugar cane (Tomkins et al., 1999a) and potato (Song et al., 2000).

The genus *Agave* comprises about 200 species with its diversification center in Mexico (García-Mendoza, 2002); a country where many species of the genus have a significant ecological and economic importance. Given

*Corresponding author. E-mail: santey@cicy.mx. Tel: (52) 9999428330. Fax: (52) 9999813900.

the importance of the genus, cytogenetic studies of *Agave* have been made since 1936, attempting to determine the size of the genome and its karyotype composition (Doughty 1936; Castorena et al., 1991; Cavallini et al., 1995; Palomino et al., 2005; Moreno-Salazar et al., 2007; Robert et al., 2008). However, decrypting the genome of *Agave* is a difficult task, due to which Martínez-Hernández et al. (2010) constructed four cDNA libraries of *Agave tequilana* var. azul from tissues with different functions and metabolic properties. These authors reported the usefulness of the libraries for the isolation of genes related to stress associated to photosynthesis (*rbcL*, encoding the small subunit of the RuBisCo), carbon metabolism (NAD-ME1), biosynthesis of oligofructans (1-SST fructosyltransferase) and abiotic stress (*LEA*). However, the application of these libraries to better understand the structure and organization of the genome in *Agave* would be impractical, since they only include expressed genes, thus excluding non coding regions (promoters and repeated regions, among others) that have an important role in DNA organization and in the regulation of gene expression (Heslop-Harrison, 2000; White et al., 1994). Hence, the construction of a genomic library in *Agave* could become a valuable tool to learn about the organization and structure of genes of interest in *Agave*.

The present study reports the construction and characterization of a BIBAC of *A. tequilana* var. azul and demonstrates its usefulness for the identification of genes of interest. In future, the development of this BIBAC library will allow us to perform studies related to: (1) search of molecular markers, (2) identification of genes involved in the resistance of biotic and abiotic factors, (3) transformation of specific genes in plants, (4) comparative genomic studies, and 5) physical mapping of regions, among other issues.

MATERIALS AND METHODS

The young leaves of *Agave tequilana* var. Azul (2X) were collected in the greenhouse of the *Centro de Investigación Científica de Yucatán* (CICY).

Construction of a genomic library of *A. tequilana* var. azul

High molecular weight DNA was obtained by the extraction of isolated nuclei according to the procedures reported by Zhang et al. (1995, 1996b). The isolated nuclei were later embedded in low melting point (LMP) agarose and arranged in molds that were subjected to analysis according to the study of Ortiz-Vázquez et al. (2005). The DNA agarose blocks were subjected to pulsed field electrophoresis (PFGE), using the conditions reported by Tao et al. (2002) to remove low molecular weight fragments. To generate fragments of 100 to 300 kb, agarose blocks (each one with approximately 0.5 µg of DNA) were digested in a reaction containing 1.6 U of Bam HI, 1 ml of 1X Bam HI buffer (50 mM Tris-HCl, 10 mM MgCl₂ and 50 mM NaCl). The reaction was incubated at 37°C for 8 min, after which a tenth of the volume of 0.5 M ethylenediaminetetraacetic acid (EDTA) at pH 8.0 was added. The

restriction of fragments was subjected to PFGE using the equipment CHEF DRII (Bio-Rad, USA) under the following conditions: 6 V/cm, 15 s time switch, 14°C, 0.5X TAE, by 10 h. The gel was stained with ethidium bromide and 100 to 300 Kb DNA fragments were cut and purified by electroelution (Strong et al., 1997) and dialysis bags with an exclusion molecular weight of 14 KDa (Sigma, St. Louis, MO, USA). Then 1 ml final volumes of electroeluate were reduced to 100 µL using Amicon NMWL 3-15 KDa columns. The DNA was quantified and used for ligations.

The binary vector BIBAC2 described by Hamilton (1997) was used for cloning the DNA fragments. DNA plasmid was extracted by the alkaline lysis method (Birnboim and Dolly, 1979) and purified by affinity chromatography using the Pure Link Kit (Invitrogen™ Corp., Carlsbad, CA, USA). Subsequently, the vector was linearized with Bam HI enzyme. The reaction contained 0.5 µg of BIBAC2 plasmid DNA, 1U of Bam HI enzyme and 1X Bam HI buffer (50 mM Tris-HCl, 10 mM MgCl₂, 50 mM NaCl) in a final volume of 15 µL. The reaction was incubated at 37°C overnight, after which it was stopped by the addition of 1.5 µL of EDTA pH 8.0, 0.5 M. Fragments were subjected to 0.8% agarose gel electrophoresis and the gels were stained with ethidium bromide. The linearized plasmid was excised from the gel, purified with β-agarase (New England BioLabs, USA) and then dephosphorylated using calf intestinal alkaline phosphatase (CIP) (New England BioLabs, USA) in a reaction mixture containing 0.5 µg of BIBAC2 plasmid DNA, 0.3U of CIP and 1X CIP buffer (50 mM Tris-HCl, 100 mM NaCl, 10 mM MgCl₂, 1 mM dTT, pH 7.9) in a final volume of 20 µL. Subsequently, the plasmid was cleaned with β-agarase.

The ligation was conducted with a 1:5 vector: insert relationship. Ligation mixtures contained 85 ng of DNA (insert + vector), 2U of T4 DNA ligase (New England BioLabs, USA), 1X buffer T4 DNA ligase (50 mM Tris-HCl, 100 mM NaCl, 10 mM MgCl₂, 1 mM dTT, pH 7.9) in a final volume of 10 µL. The reaction was incubated overnight at 16°C. Afterwards, 5 µL of the ligation mixture was used to transform *E. coli* DH10B competent cells. The transformation was carried out by electroporation at 350 V, 330 IF capacitance and 4 kW resistance. The transformation was transferred to 1 ml of SOC (2% bacto tryptone, 0.5% yeast extract, 10 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose, pH 7.0) and incubated for 1 h at 37°C and 250 rpm. The transformants were grown overnight at 37°C in selective medium (Luria broth (LB) plus kanamycin (30 mg/L) and 5% sucrose). Selected colonies were grown in medium supplemented with kanamycin LB and then stored at -60°C until use. The efficiency of transformation was calculated according to the formula: transformed UFC = (number of colonies)(original volume of transformation)(dilution ratio) / (volume used in the plate); and Efficiency of transformation = UFC transformed / plasmid DNA in µg (Sambrook et al., 1989). The reported values were obtained from the CFU count of ten independent transformation events.

Estimation of insert size, internal Not I sites, percentage of empty clones and redundant clones

One hundred clones were randomly selected, the plasmid DNA was extracted according to Birnboim and Dolly (1979) and digested with Not I enzyme (New England BioLabs, USA) incubated at 37°C overnight. The restriction profiles of clones were analyzed by PFGE in 0.8% agarose gels. The insert size was estimated by comparison with a high molecular weight ladder (NEB, USA). The percentage of empty clones was determined by the absence of extra bands besides the cloning vector. The percentage of Not I internal sites was determined by the presence of more than one band, exclusive of the insert. In order to analyze the redundancy of clones in the library, one hundred clones with insert were analyzed by incubation with the restriction enzymes Eco RV and Hind III, after which insert clones showing the same profile were tagged as redundant.

Calculation of the probability of finding a sequence of interest in the *A. tequilana* var. azul BIBAC and estimation of the representativeness in the species' genome

A calculation of the probability (P) of finding a single copy sequence in the *Agave* BIBAC was made by the formula: $N = \ln(1-P) / \ln(1-1/GS)$ (Clarke and Carbon, 1976), where N is the number of recombinants needed to find the sequence of interest, l is the insert size in base pairs (bp) and GS is the size in bp of the haploid genome. The level of coverage of the genome (W) was estimated by the formula: $W = Nl/GS$ (Paterson, 1996).

Representation in the BIBAC of regions of interest and contamination by organelle DNA

Using bioinformatics, seven pairs of primers were designed for the amplification of sequences specific to *rbcL*, 5S and 18S rDNA, *NAD4*, *NBS-LRR*, *LEA* and retroelements (Table 1). The primers corresponding to the *NBS-LRR*, *LEA* and *rbcL* genes were designed using sequences isolated from an EST library of *Agave fourcroydes* (Unpublished data). In the case of *NAD4*, the primers were designed from sequences reported in GenBank (NCBI) for *Agave parryi* and *Agave celsii*, with accessions numbers AANS2069.1 and ABB16359.1, respectively. The specific primers of retroelements used were those reported by Xion and Eickbush et al. (1990), which amplified the conserved domains of the reverse transcriptases of the retroelements DVKTFH (N) G-LLYVDDM (V) and RMCVDYR-YAKLSKS. Finally, the primers of 5S and 18S rDNA genes were designed from a consensus sequence of ribosomal regions obtained by sequencing rDNA from different *Agave* species. The software packages used for sequence alignment and analyses were BioEdit, Mega 4.1 and Dnasp5. The design of primers was performed using DNAMAN and Oligo analyzer. The primers used are positioned on conserved regions or domain features of the proteins encoded by the above-mentioned genes. The sequences of the used primers are shown in Table 1. The identities of the amplified fragments were confirmed by sequencing and the sequences were submitted to GenBank.

The primers were used to amplify probes for screening of the clones. To do this, in a 22.5 cm nylon membrane (Hybond N⁺, Amersham Pharmacia Biotech, Arlington Heights, IL, EEUU) 2,880 BIBAC library clones were set in 96 points (Figure 1a). Each point of the membrane contained a mixture of 30 clones. The membranes were hybridized with probes corresponding to the retroelements, rDNA, *NBS-LRR*, *LEA* and organelle DNA including chloroplast and mitochondrial DNA (*rbcL* and *NAD4*, respectively). Probes were labeled with digoxigenin-11-dUTP by polymerase chain reaction (PCR) as described by Fulnecek et al. (2002). Hybridized probes were immunodetected with anti-digoxigenin-AP and revealed with the substrate CSPD. The membranes were exposed to X-ray film; clones showing a signal were analyzed by PCR at 3 different levels. Two PCR reactions, including DNA from 15 clones (equimolar concentration) were evaluated (Figure 1b). If amplification was positive, a PCR reaction was made divided in reactions including only 5 clones (Figure 1c) and, finally, if amplification was positive, individual PCR reactions were ran (Figure 1d).

Identification of clones with possible application for the genetic mapping of *Agave*

A group of 10 randomly selected clones were analyzed, each containing a region of interest (*NBS-LRR*, *LEA*, 18S rDNA, 5S rDNA and retroelements). A total of 40 clones were assessed by digestion with the enzyme Apo I and incubated at 50°C overnight. The restriction profile of these clones was analyzed by PFGE in a

0.8% agarose gel. Redundancy of clones was determined from the observed restriction profiles.

RESULTS

Characterization of the BIBAC of *A. tequilana* var. azul

A binary genomic library (BIBAC) was successfully built from nuclear DNA of *A. tequilana* var. azul, on the basis of a nuclei solution from 10 individuals of *A. tequilana* var. azul. Fragments selected for cloning were in a size range of 100 to 300 Kb. The relation vector: insert used for the construction of the BIBAC2 was 1:5. The efficiency of transformation was estimated at 3.4×10^3 clones/ μ g of DNA, using electrocompetent cells. The partial library consists of 9,800 clones.

To estimate the size of insert of the clones forming the library, 100 clones were selected at random and were digested with the enzyme Not I and analyzed by PFGE (Figure 2). The average insert size of clones was estimated at 170 Kb. More than 50% of the analyzed clones showed inserts within a size range of 100 to 200 Kb (Figure 3). Only 10% of the clones were empty and 30% contained internal Not I sites. Moreover, 15% of the analyzed clones were redundant. Considering that the 1CX genome size (not replicated nuclear DNA content) of *A. tequilana* var. azul is 4.40 pg (Bannerje and Sharma, 1987); this percentage would correspond to 4312 Mpb (Banerjee and Sharma, 1987). The clones obtained in the *A. tequilana* var. azul genomic library represented 0.38 X of the haploid genome of *Agave*. According to Clarke and Carbon (1976), the probability of obtaining a particular clone in this library using a single copy probe is around 8.3%.

BIBAC with organelle DNA contamination

To estimate the contamination of the BIBAC with organelle DNA, a membrane was analyzed containing 2,880 clones representing 30% of the library. It was estimated that 3% of clones contained mitochondrial DNA and 2% chloroplastic DNA. These data indicate that 95% of the clones contained nuclear DNA. Furthermore, to demonstrate the usefulness of the library for the identification of genes of interest related to biotic stress (*NBS-LRR*), embryogenesis (*LEA*) and highly repetitive sequences (retroelements and rDNAs), a total of 2,880 clones were analyzed for each gene of interest. We identified clones corresponding to each region of interest (Table 2). According to the results, the repetitive sequences are most represented in this library (7.5%), followed by *NBS-LRR* (0.76%) and *LEA* regions (0.28%). These results indicate that the obtained genomic library is a great tool for searching genes of interest.

In order to evaluate the presence of specific genes/

Table 1. DNA probes specific for mitochondrial, chloroplastic and nuclear genes used in the characterization of BIBAC of *A. tequilana* Weber var. azul.

Probe	Coding region	Sequence primer (5'-3')	Tm (°C)	Size (pb)
rbcL ^a	Small subunit of ribulose-1-5-bisphosphate	L-5'-ggctcttccgcgccaccg-3'	63	350
		R-5'-acgcactgcactgacggt-3'	63	
lea ^b	Abundant protein during embryogenesis	L-5'-gcagcagctccagccaatac-3'	63	350
		R-5'-ggctgtagtccgtaactgg-3'	63	
nbs-lrr ^b	Disease resistance protein	L-5'-tggttacaccgttgaggact-3'	62	400
		R-5'-ctggtctcc ccagtccac-3'	62	
Nad4 ^c	Subunit of dehydrogenase 4	L-5'-ggcagcaagtgtgtcaaa-3'	62	1000
		R-5'-ctattgactg cttaggta-3'	62	
18S ^b	Ribosomal 18S rDNA	L-5'-cggctaccacatccaaggaa-3'	63	180
		R-5'-gctggaattaccgcggt-3'	63	
5S ^b	Ribosomal 5S rDNA	L-5'-cgatcataccagcactaaagcacc-3'	64	120
		R-5'-atgcaacacgaggacttcccag-3'	64	

^arbcL is a chloroplastic sequence; ^blea, nbs-lrr, 18s and 5s rDNA are nuclear sequences; ^cnad4 is a mitochondrial sequence.

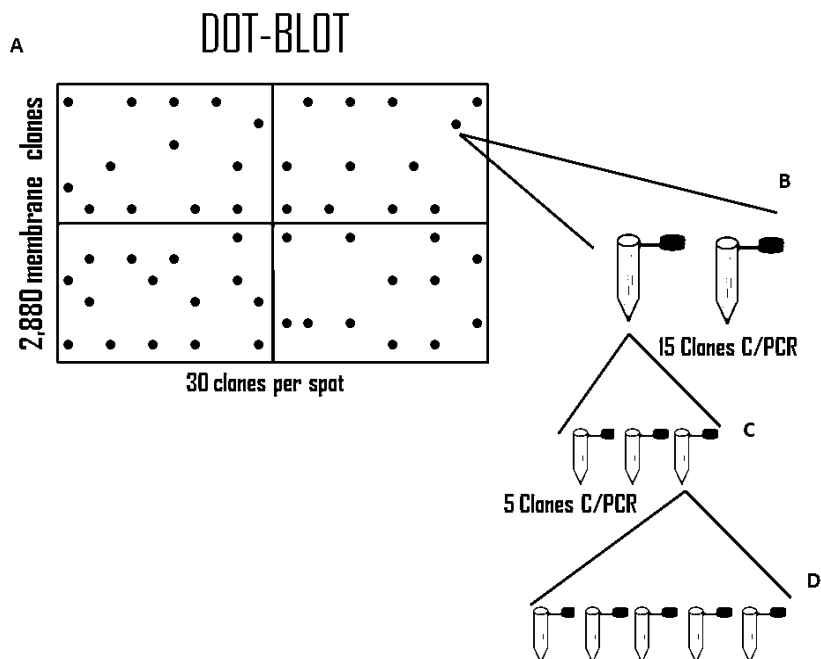


Figure 1. Strategy used for the search of repetitive elements in the BIBAC2 library of *A. tequilana* Weber var. azul. (a) 2880 clones are laid down in a Hybond N⁺ nylon membrane at 96 spots, each spot containing 30 clones. (b) Mixtures giving signs of hybridization with a particular gene area analyzed in two PCR reactions, which include 15 DNA plasmids of different clones. (c) PCR to amplify the gene of interest was made by analyzing three PCR reactions containing five DNA plasmids from different clones. (d) Finally, the studied fragment is amplified in five individual PCR reactions.

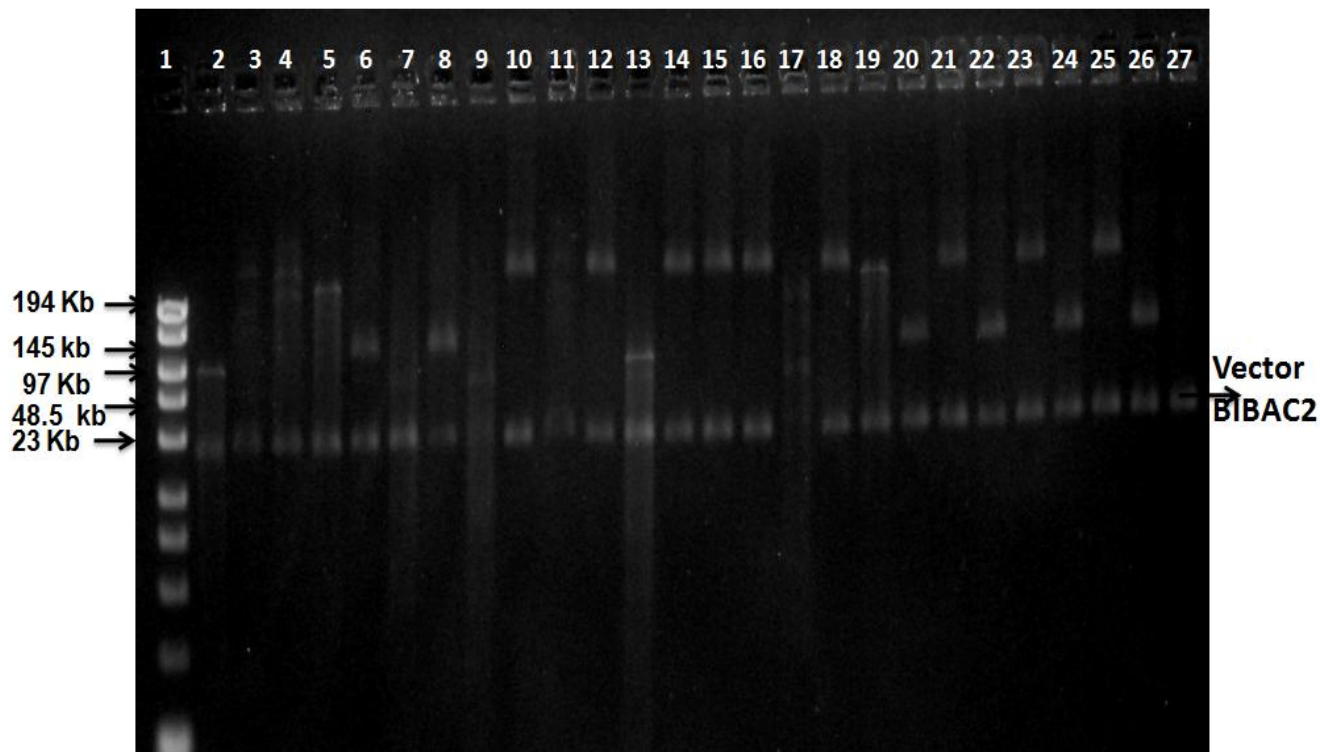


Figure 2. Analysis of 24 randomly selected clones belonging to the BIBAC of *A. tequilana* Weber var. azul. Clones were digested with the enzyme Not I and separated by pulsating field gel electrophoresis (PFGE). Lane 1, Molecular weight marker (New England); lanes 2 to 26, library clones; lane 27, linearized BIBAC without insert.

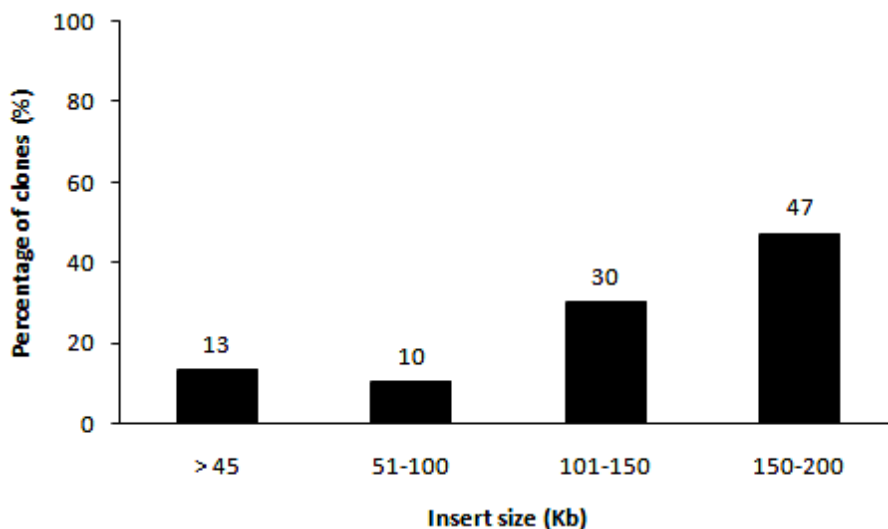


Figure 3. Insert size distribution in the BIBAC2 library clones. Based on 100 clones, analysis were digested with the enzyme Not I and separated by pulsating field gel electrophoresis (PFGE).

regions in a certain clone, individual clones containing genes such as *NBS-LRR*, *LEA*, rDNAs and retroelements were compared by restriction profiles, obtaining several different clones for each gene/region (Figure 4). The

profiles observed for each analyzed clone showed different patterns, which indicates that each one of these clones is different and that their corresponding inserts are possibly located in different positions within the genome

Table 2. Composition of sequences of interest in the genome of *Agave* estimated by the representativeness of these regions in the BIBAC of *A. tequilana* Weber var. azul.

Regions	Number of clones ^a	Proportion of clones in BIBAC (%)
<i>NBS-LRR</i>	75	0.72
<i>LEAS</i>	28	0.28
Retroelements	738	7.5
Tandem repeats		
18S rDNA	56	0.57
5S rDNA	30	0.37
Other sequences		90.52

^aDetermined through the analysis of 2880 clones.

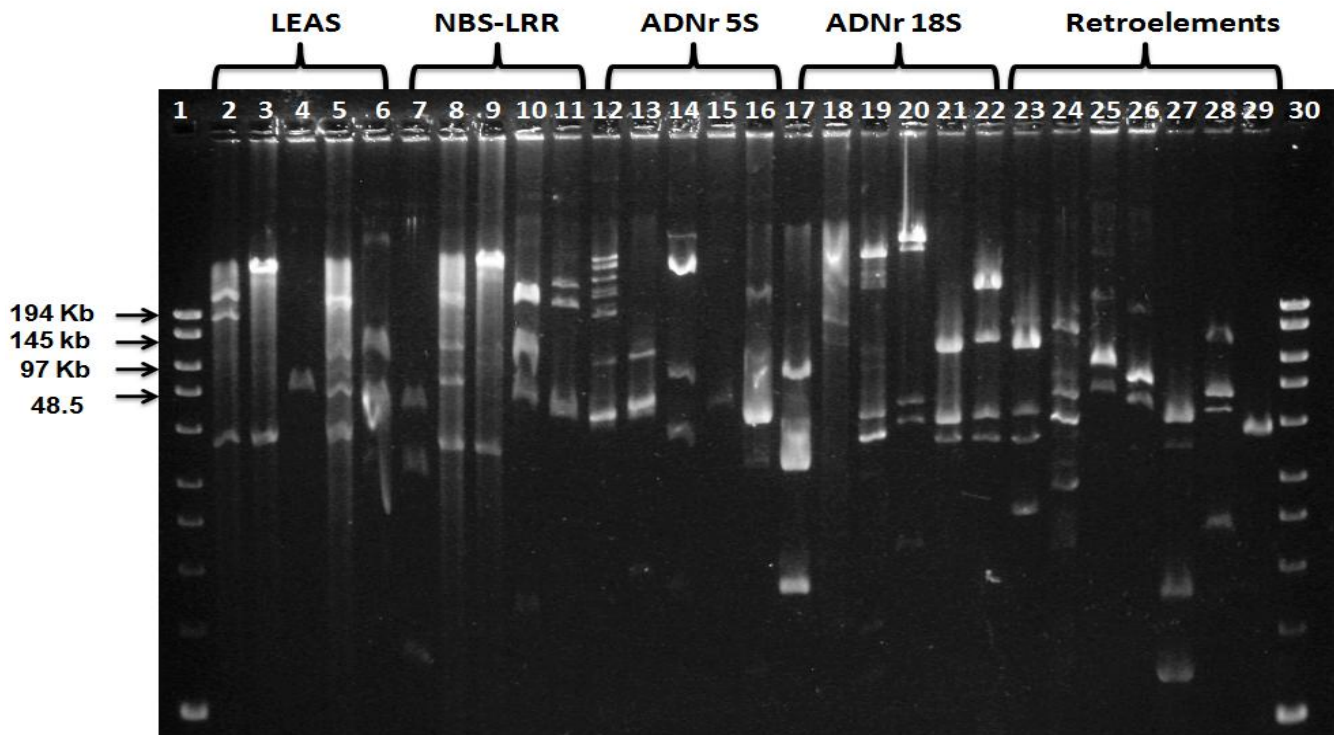


Figure 4. Analysis of 28 clones positive to genes of interest. The clones were digested with Apo I. Lanes 1 and 30) molecular weight marker (New England); lanes 2 to 6, *LEAS* clones; lanes 7 to 11, *NBS-LRR* clones; lanes 12 to 16, 5S rDNA clones; lanes 17 to 22, 18S rDNA clones and lanes 23 to 29, retroelement clones.

of *A. tequilana* var. azul. The future identification of these clones could be used for genetic mapping studies in the species.

DISCUSSION

This study reports the construction and characterization of a partial BIBAC of *A. tequilana* var. azul with an average size of inserted fragment of 170 Kb, greater than that reported for other species' BACs in *Arabidopsis*, tomatoes and maize having an insert size of 100 kb (Choi et al., 1995; Mozo et al., 1998; Folkertsma et al., 1999).

Genomic libraries of soybean have an average insert size of 105 Kb (Salimath and Bhattacharyya, 1999), 120 Kb (Danesh et al., 1998) and 136 Kb (Tomkins et al., 1999b). The average insert size estimated in the genomic library of rice was found to be of 125 Kb (Wang et al., 1995), 130 Kb in sugar cane (Tomkins et al., 1999a) and of 155 Kb in potato (Song et al., 2000). The cloning in the present work of fragments greater than those reported for other BACs, suggests that this is related to DNA integrity and to the enrichment of large fragments (100-300 kb). In addition, 30% of the clones contain internal Not I sites, suggesting a low GC content or CpG Islands in the genome of *Agave*. He et al. (2003) also reported that the

G+C content in monocotyledon plants, such as *Agave*, is higher than the same in dicots. This could reduce the frequency of finding the recognition site for the enzyme Not I.

The high quality of the library reported herein is demonstrated both by the low percentage of empty clones found (10%) –similar to that in the BIBACs of *Lysopersicon esculentum* and *L. pennellii* (Hamilton et al., 1999)– and by the low percentage of organelle DNA (3%) contamination –which was comparable to that reported for other BACs. An important aspect to highlight is the high representation of retroelements found in this library of *A. tequilana* Weber var. azul. The retroelements are considered to be a mutation source having an evolutionary importance due to its insertion mechanism (Xion and Eickbush, 1990). The genetic variability resulting from the insertion of retroelements in the genome contributes to genomic rearrangements and to the increase in size of the genomes of plants (SanMiguel and Bennetzen, 1998; Wang et al., 1999). It is suggested that retroelements represent more than 50% of the genome of higher plants (SanMiguel et al., 1996), so, it is not surprising to find that 7.5% of the *Agave* library contains retroelements. Sequencing projects of *Daucus carota* L. (Cavagnaro et al., 2009), *Zea mays* (Meyers et al., 2001), *Triticum urartu* (Akhunov et al., 2005), *Pisum sativum* (Blemishes et al., 2007) and *Secale cereale* (Bartos et al., 2008) among others, demonstrated that in these species the retroelements are represented by 56-88%.

It is also important to mention that *Agave* belongs to a genus of recent origin, with a rate of species diversification varying between 0.32 to 0.56 species per one million years (Good-Avila et al., 2006), a high value in comparison with the same rate reported for angiosperms (0.077 to 0.089) (Magañon and Sanderson, 2001). So far, the events that cause this diversification are not known with enough accuracy and the retroelements could be playing an important role in this genus. Also it has been reported that the variation in the size of plant genomes is in part caused by the accumulation of repetitive elements (Flavell et al., 1974; Flavell, 1980); in species of *Agave* in particular, the retroelements may also be involved in the size of this species genome.

Concerning the representativeness of the obtained library, *Agave* presents a medium size genome of 4312 Mpb (Banerjee and Sharma, 1987) compared to other plant species, such as those of *Citrus clementina* (367 Mpb) (Terol et al., 2008), *Musa acuminata* (600 Mpb) (Ortiz and Vazquez et al., 2005), *Oryza nivara* (760 Mpb) (Ammiraju et al., 2006) and *Pinus pinaster* (23 000 Mpb) (Bautista et al., 2008). High representativeness in the construction of genomic libraries has been reported for several species of plants, for which automatic procedures and equipment has been used. However, the larger the genome the studied subjects is, the lower will be the

representativeness of the genome in the libraries. Even in *P. pinaster*, a species with a large genome size (23 000 Mpb) compared to that of *Agave*, until date the total representation of 1CX value has not been achieved.

The obtained library for *A. tequilana* represents 0.38X of the haploid genome of *Agave*; this level of representation is significant considering that the genome of this species is 33 times larger than that of *Arabidopsis* and that the BIBAC was not built by automated processes. In addition, clones positive to genes of interest has proven not to be redundant and possibly have different locations in the genome of *Agave*. These clones could be used for future physical and genetic mapping in this species. This is the first reported genomic library of *A. tequilana* var. azul, which can be used as a powerful tool for studying the genome of *Agave* and for studies of functional genomics and genetic engineering in this species.

ACKNOWLEDGEMENTS

The authors thank Matilde Ortiz García and Adriana Quiroz Moreno for technical their support. This project was supported by the National Council of Science and Technology of Mexico by the project (50268).

REFERENCES

- Akhunov ED, Akhunova AR, Devorak J (2005). BAC libraries of *Triticum urartu*, *Aegilops speltoides* and *Ae. tauschii* the diploid ancestors of polyploid wheat. *Theor. Appl. Genet.* 111:1617-1622.
- Ammiraju JSS, Luo M, Goicoechea JL, Wang W, Kudrna D, Mueller C, Talag J, Kim H, Sisneros NB, Blackmon B, Fang E, Tomkins JB, Brar D, MacKill D, McCouch S, Kurata N, Lambert G, Galbraith DW, Arumuganathan K, Rao K, Walling JG, Gill N, Yu Y, San Miguel P, Soderlund C, Jackson S, Wing RA (2006). The *Oryza* bacterial artificial chromosome library resource: Construction and analysis of 12 deep-coverage large-insert BAC libraries that represent the 10 genome types of the genus *Oryza*. *Genome Res.* 16:140-147.
- Banerjee S, Sharma AK (1987). Cytophotometric estimation of nuclear DNA in different species and varieties of *Agave*. *Cytologia* 52:85-90.
- Bartos J, Paux E, Kofler R, Havrankova M, Kopecky D, Suchankova P, Safar J, Simkova H, Town CD, Lelley T, Feuillet C, Dolezel J (2008). A first survey of the rye (*secale cereale*) genome composition through BAC end-sequencing of the short arm of chromosome 1R. *BMC Plant. Biol.* 8:95.
- Bautista R, Villalobos DP, Diaz-Moreno S, Cantun FR, Canovas FM, Gonzalo Claros M (2008). Nueva estrategia para la construcción de una genoteca genómica de *Pinus pinaster* en cromosomas artificiales de bacterias. *Sistemas y Recursos Forestales* 3:238-249.
- Birboim HC, Dolly J (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acid Res.* 7:1513-1523.
- Boysen C, Simon MI, Hood L (1997). Fluorescence-based sequencing directly from bacterial and P1-derived artificial chromosomes. *Biotechniques* 23:978-982.
- Cai WW, Reneker J, Chow CW, Vaishnav M, Bradley A (1998). An anchored framework BAC map of mouse chromosome 11 assembled using multiplex oligonucleotide hybridization. *Genomics* 54:387-397.
- Castorena-Sanchez I, Escóbedo RM, Quiroz A (1991). New cytotoxicological determinants recognized in six taxa of *Agave* in the sections *Rigidae* and *Sisalanae*. *Can. J. Bot.* 69:1257-1264.
- Cavallini A, Natali L, Cionini G, Castorena-Sanchez I (1995). Cytophotometric and biochemical analyses of DNA in pentaploid and

- diploid *Agave* species. *Genome* 39:266-271.
- Cavagnaro PF, Chung SM, Szklarczyk M, Grzebelus D, Senalik D, Atkins AE, Simon PW (2009). Characterization of a deep-coverage carrot (*Dacus carota* L.). BAC library and initial analysis of BAC-end sequences. *Mol. Genet. Genomics* 281:273-288.
- 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.
- Clarke K, Carbon J (1976). A colony bank contained syntethic *Cole1* hybrid plasmid representative of the entire *E. coli*. *Genome Cell* 11:91-100.
- Danesh D, Penuela 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 nematode resistance gene. *Theor. Appl. Genet.* 96:196-202.
- Doughty LR (1936). Chromosome behavior in relation to genetics of *Agave*. I. Seven species of fibre *Agaves*. *J. Genet.* 33:197-205.
- Flavell RB, Bennett MD, Smith JB, Smith DB (1974). Genome size and the proportion of repeated nucleotide sequence DNA in plants. *Biochem. Genet.* 12:257-269.
- Flavell R (1980). Molecular characterization and organization of plant chromosomal DNA sequences. *Annu. Rev. Plant. Physiol.* 31:569-596.
- Folkertsma RT, Spassova MI, Prins M, Stevens MR, Hille J, Goldbach RW (1999). Construction of a bacterial chromosome (BAC) library of *Lycopersicon esculentum* cv. Stevens and its application to physically map the *Sw-5* locus. *Mol. Breed.* 5:197-207.
- Fulnecek J, Lim KY, Leitch AR, Kovarick A, Matyasek R (2002). Evolution and structure of 5S rDNA loci in allotetraploid *Nicotiana tabacum* and its putative parental species. *Heredity* 88:19-25.
- Good-Avila SV, Souza V, Gaut BS, Eguiarte LE (2006). Timing and rate of speciation in *Agave* (Agavaceae). *PNAS* 103:9124-9129.
- Hamilton CM, Frary A, Lewis C, Tanksley SD (1996). Transfer of intact high molecular weight DNA into plant chromosomes. *Proc. Natl. Acad. Sci. USA*, 93:9975-9979.
- Hamilton CM (1997). A binary-BAC system for plant transformation with high-molecular-weight DNA. *Gene* 200:107-116.
- Hamilton CM, Frary A, Xu Y, Tanksley SD, Zhang HB (1999). Construction of tomato genomic DNA libraries in a binary-BAC (BIBAC) vector. *Plant J.* 18:223-229.
- He RF, Wang Y, Shi Z, Ren X, Zhu L, Whang Q, He GC (2003). Construction of a genomic library of wild rice and agrobacterium-mediated transformation of large insert DNA linked to BPH resistance loci. *Gene* 321:113-121.
- Heslop-Harrison JS (2000). Comparative genome organization in plants from sequence and markers to chromatin and chromosomes. *Plant Cell* 12:617-635.
- Liu Y, Shiramo Y, Fukaki H, Yanai Y, Tasaka M, Tabata S, Shibata D (1999). Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artificial chromosome vector accelerates positional cloning. *Proc. Natl. Acad. Sci. USA*, 96:6535-6540.
- Magañon S, Sanderson MJ (2001). Absolute diversification rates in angiosperm clades. *Evolution* 55:1762-1780.
- Martinez-Hernandez A, Mena-Espino ME, Herrera-Estrella AH, Martinez-Hernandez P (2010). Construction de bibliotecas de ADNc y analisis de expresión génica. *Rev Latinoamer Quim.* 38:21-44.
- Marra MA, Kucaba TA, Dietrich NJ, Green ED, Brownstein B, Wilson RK, McDonald KM, Hillier LW, McPherson JD, Waterson RH (1997). High throughput fingerprint analysis of large-insert clones. *Genome Res.* 7:1072-1084.
- Meyers BC, Tingey SV, Morgante M (2001). Abundance distribution and transcriptional activity of repetitive elements in the maize genome. *Genome Res.* 11:1660-1676.
- Moreno-Salazar SF, Esqueda M, Martinez J, Palomino G (2007). Tamaño del genoma y cariotipo en *Agave angustifolia* y *A. rhodacantha* de Sonora Mexico. *Rev. Fitotec. Mex.* 30:13-27.
- Mozo T, Fischer S, Shizuya H, Altmann T (1998). Construction and characterization of the IGF *Arabidopsis* BAC library. *Mol. Gen. Genet.* 258:562-570.
- Nakamura S, Asakawa S, Ohmido N, Fukui K, Shimizu N, Kawasaki S (1997). Construction of an 800-kb contig in the near-centromeric region of the rice blast resistance gene *Pi-ta2* using a highly representative rice BAC library. *Mol. Gen. Genet.* 254:611-620.
- Ortiz-Vazquez E, Kaemmer D, Zhang HB, Muth J, Rodriguez-Mendiola M, Arias-Castro AJ (2005). Construction and characterization of a plant transformation-competent BIBAC library of the black sigatoka-resistant banana *Musa Acuminata* cv Tuu Gia (AA). *Theor. Appl. Genet.* 110:706-713.
- Palomino G, Martinez J, Mendez I (2005). Citotipos en *Agave angustifolia* Haw. Determinados por citometría de flujo y análisis de sus cariotipos. *Rev. Int. Contam.* 21:49-54.
- Paterson AH (1996). *The DNA revolution*. San Diego (CA), Academic Press.
- Robert ML, Yoong LK, Hanson L, Sanchez-Teyer F, Bennett DM, Leitch AR, Leitch IL (2008). Wild and agronomically important *Agave* species (Asparagaceae) show proportional increases in chromosome number, genome size, and genetic markers with increasing ploidy. *Bot. J. Linn. Soc.* 158:215-222.
- Salimath SS, Bhattacharyya MK (1999). Generation of a soybean BAC library, and identification of DNA sequences tightly linked to the *Rps1-k* disease resistance gene. *Theor. Appl. Genet.* 98:712-720.
- Sambrook J, Fritsch EF, Maniatis T (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, Laboratory Press, NY.
- SanMiguel P, Tikhonov A, Jin YK, Motchoukkaia N, Zalharov D, Metake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z, Bennetzen JC (1996). Nested retrotransposons in the intergenic regions of the maize genome. *Science* 274:765-768.
- SanMiguel P, Bennetzen JI (1998). Evidence that a Recent Increase in Maize genome Size was Caused by the Massive Amplification of Intergene Retrotransposons. *Annals Bot.* 82:37-44.
- Song J, Dong F, Jiang J (2000). Construction of a bacterial artificial chromosome (BAC) library for potato molecular cytogenetics research. *Genome* 43:199-204.
- Strong SJ, Ohta Y, Litman GW, Amemiya CT (1997). Marked improvement of PAC and BAC cloning is achieved using electroelution of pulsed-field gel-separated partial digests of genomic DNA. *Nucleic Acids Res.* 25:3959-3961.
- Terol J, Naranjo MA, Ollitrault P, Talon M (2008). Development of genomic resources for Citrus clementina: Characterization of three deep-coverage BAC libraries and analysis of 46 000 BAC-end sequences. *BMC Genomics* 9:423.
- 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.
- Wang GL, Holsten TE, Sung WY, wang HP, Ronald PC (1995). Construction of a rice bacterial artificial chromosome library and identification of clones linked to the Xa-21 disease resistance locus. *Plant J.* 7:523-533
- Wang GL, Warren R, Innes R, Osborne B, Baker B, Ronald P (1996). Construction of an *Arabidopsis* BAC library and isolation of clones hybridizing with disease-resistance, gene-like sequences. *Plant. Mol. Biol. Rep.* 14:107-114.
- Wang SP, Liu N, Peng KM, Zhang QF (1999). The distribution and copy number of copia-like retrotransposons in rice (*Oryza sativa*) and their implications in the organization and evolution of the rice genome. *Proc. Natl. Acad. Sci. USA.* 96:6824-6828.
- White SE, Habera LF, Wessler SR (1994). Retrotransposons in the flanking regions of normal plant genes: a role for copia-like elements in the evolution of gene structure and expression. *Proc Natl Acad Sci USA* 25:11792-11796.
- Willets N, Skurry R (1987). Structure and function of the F factor and mechanism of conjugation. In *Escherichia coli* and *Salmonella typhirium*: cellular and Molecular Biology. Am. Soc. Microbiol. 2:1110-1133.
- Xion Y, Eickbush TH (1990). Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* 9:3353-3362.
- Yang D, Parco A, Nandi S, Subudhi P, Zhu Y, Wang G, Huang N (1997). Construction of a bacterial artificial chromosome (BAC) library

and identification of overlapping BAC clones with chromosome 4-specific RFLP markers in rice. *Theor. Appl. Genet.* 95:1147-1154.

Zhang HB, Zhao XP, Ding XL, Paterson AH, Wing R (1995). Preparation of megabase-size DNA from plant nuclei. *Plant J.* 7:175-184.

Zhang HB, Woo SS, Wing RA (1996b). BAC, YAC and cosmid library construction. In Foster G, Twell D (eds) *plant gene isolation: principles and practice*. Wiley, England.