Theor Appl Genet (1999) 99:419–424

J.P. Tomkins · Y. Yu · H. Miller-Smith · D.A. Frisch
S.S. Woo · R.A. Wing

A bacterial artificial chromosome library for sugarcane

Received: 12 September 1998 / Accepted: 12 March 1999

Abstract Modern cultivated sugarcane is a complex aneuploid polyploid with an estimated genome size of 3000 Mb. Although most traits in sugarcane show complex inheritance, a rust locus showing monogenic inheritance has been documented. In order to facilitate cloning of the rust locus, we have constructed a bacterial artificial chromosome (BAC) library for the cultivar R570. The library contains 103,296 clones providing 4.5 sugarcane genome equivalents. A random sampling of 240 clones indicated an average insert size of 130 kb allowing a 98% probability of recovering any specific sequence of interest. High-density filters were gridded robotically using a Genetix Q-BOT in a 4 × 4 double-spotted array on 22.5-cm^2 filters. Each set of five filters provides a genome coverage of 4x with 18,432 clones represented per filter. Screening of the library with three different barley chloroplast gene probes indicated an exceptionally low chloroplast DNA content of less than 1%. To demonstrate the library's potential for map-based cloning, single-copy RFLP sugarcane mapping probes anchored to nine different linkage groups and three different gene probes were used to screen the library. The number of positive hybridization signals resulting from each probe ranged from 8 to 60. After determining addresses of the signals, clones were evaluated for insert size and *Hin*dIII-fingerprinted. The fingerprints were then used to determine clone relationships and assemble contigs. For comparison with other monocot genomes, sugarcane RFLP probes were also used to screen a *Sorghum bicolor* BAC library and two rice BAC libraries. The rice and sorghum BAC clones were characterized for insert size and fingerprinted, and the results compared to sugarcane. The library was screened with a rust resistance RFLP marker and candidate BAC clones were subjected to RFLP fragment matching to identify those corresponding to the same genomic region as the rust gene.

Key words Sugarcane · *Saccharum* spp. · BAC library · BAC fingerprinting · Rust resistance

Communicated by B.S. Gill

J.P. Tomkins · Y. Yu · H. Miller-Smith · D.A. Frisch · S.S. Woo
R.A. Wing (✉)
Clemson University Genomics Institute Room 100 Jordan Hall
Clemson, SC 29634, USA
e-mail: www.genome.clemson.edu
Fax: +1 864 656 4293

Introduction

Modern cultivated sugarcane is a complex aneu-polyploid derived from interspecific crosses between *Saccharum officinarum* and *Saccharum spontaneum*. Using comparative genomic in situ hybridization, D'Hont et al. (1996) determined that among the chromosomes of the modern cultivar R570 (2n = 107–115) about 10% originated from *S. spontaneum* and about 10% were determined to be recombinant chromosomes between *S. officinarum* and *S. spontaneum.*

We are interested in the map-based cloning of a rust resistance gene from sugarcane in collaboration with the International Consortium for Sugarcane Biotechnology. Common rust of sugarcane is caused by the fungus *Puccinia melanocephala* Syd P. Syd. Inheritance of rust resistance in sugarcane was recently studied in the selfed progeny of cv R570 (Daugrois et al. 1996). While Mendelian segregation is difficult to observe in progeny of a complex aneuploid polyploid like cv R570, rust resistance segregated in a clear 3 (resistant) : 1 (susceptible) ratio. In fact, this was the first documented report of monogenic inheritance for disease resistance in sugarcane. The same study also reported a RFLP marker linked at 10 cM with the gene. The agronomic/economic importance of rust resistance, combined with monogenic inheritance and a linked molecular marker make this a potential candidate for map-based cloning.

The primary objective of the present study was to establish the resources and protocols necessary to facilitate map-based cloning in sugarcane. Specific objectives were: (1) develop a bacterial artificial chromosome (BAC) library providing at least four haploid genome equivalents, (2) characterize the library for insert size

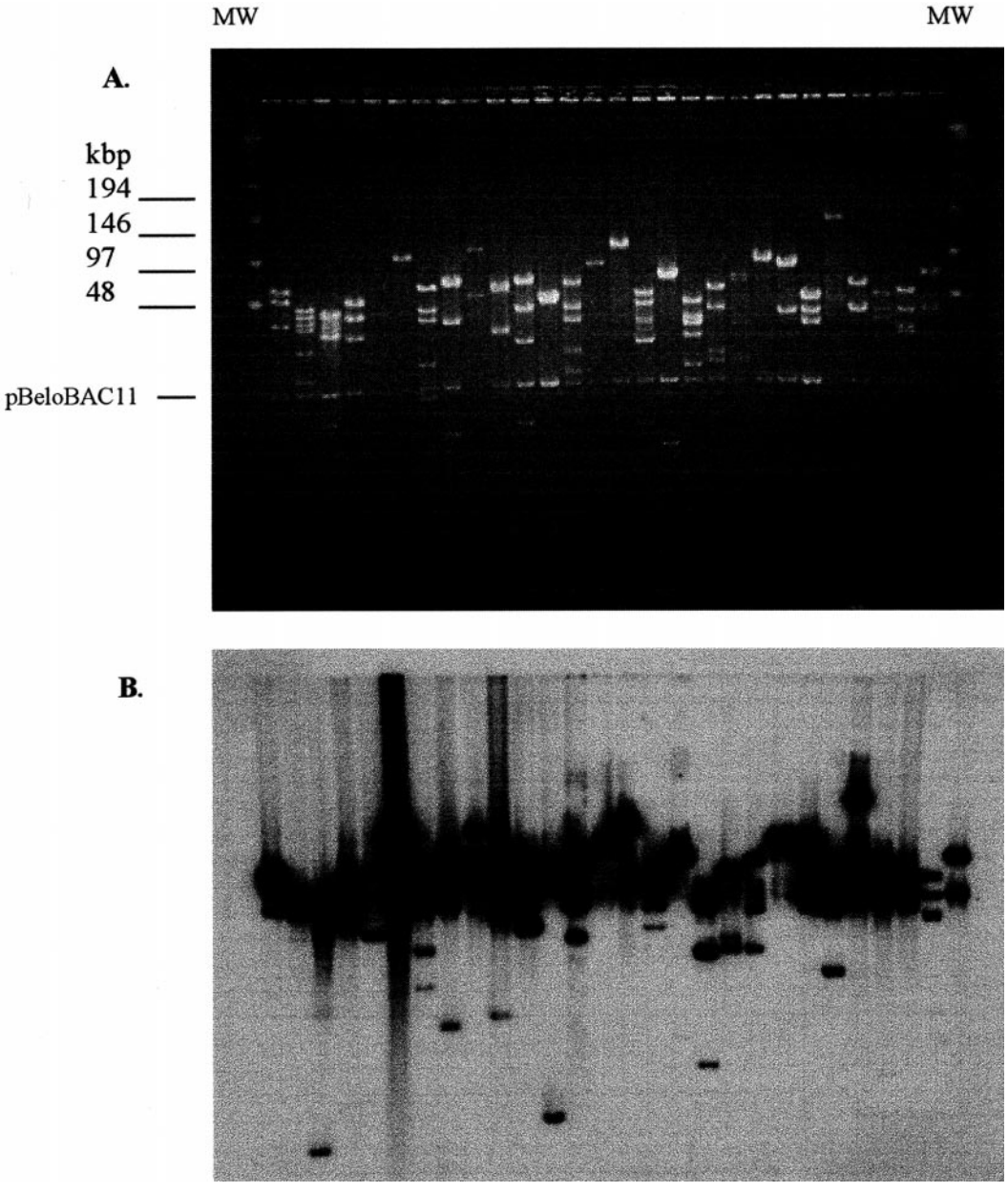

Fig. 1 A, B Analysis of 28 randomly selected sugarcane BAC clones. **A.** Ethidium bromide-stained CHEF gel (5–15°s switch time, 14°h) showing insert DNA above and below the common 7.5-kbp pBelo BAC11 vector band. **B.** Autoradiograph of gel in **A** after Southern transfer and probing with total sugarcane genomic DNA. Molecular-weight marker is a 48.5 kbp lambda concatamer (Bio-Rad)

and chloroplast DNA content, and (3) screen the library with probes distributed across the genome.

Materials and methods

BAC library construction

BAC library construction was essentially the same as that described by Zhang et al. (1996) with the following modifications. The first size selection used switch times of either a constant 20-s interval or a 1–40-s linear ramp. Fractions between 100 and 300 kbp were cut from the gel and inserted into a second gel and run at a constant 3-s switch time to remove small trapped DNA fragments. After removing appropriate fractions from the second size selection, DNA was either removed from the agarose by Gelase (Epicentre) or electroelution (Model 422 Electro-Eluter, Bio-Rad). Transformed cells were plated on 200 ml of selective medium (LB, Luria-Bertani medium) in 24×24 cm plates (Genetix) with 12.5 μg/ml of chloramphenicol, 0.55 mM IPTG, and 80 μg/ml of X-Gal. After a 20-h incubation at 37°C, the plates were placed at room temperature in the dark for an additional 20 h to allow stronger color development of non-recombinant colonies. Plates were either stored at 4°C or used immediately for picking. Recombinant white colonies were picked robotically using the Genetix Q-BOT and arrayed as individual clones in 384-well microtiter plates (Genetix) containing 50 μl of freezing broth (Woo et al. 1994). After incubation overnight, microtiter plates were stored at – 80°C. Two copies of the library were made using the replicating function of the Genetix Q-BOT and stored in separate – 80°C freezers. BAC clone characterization has been described previously by Woo et al. (1994).

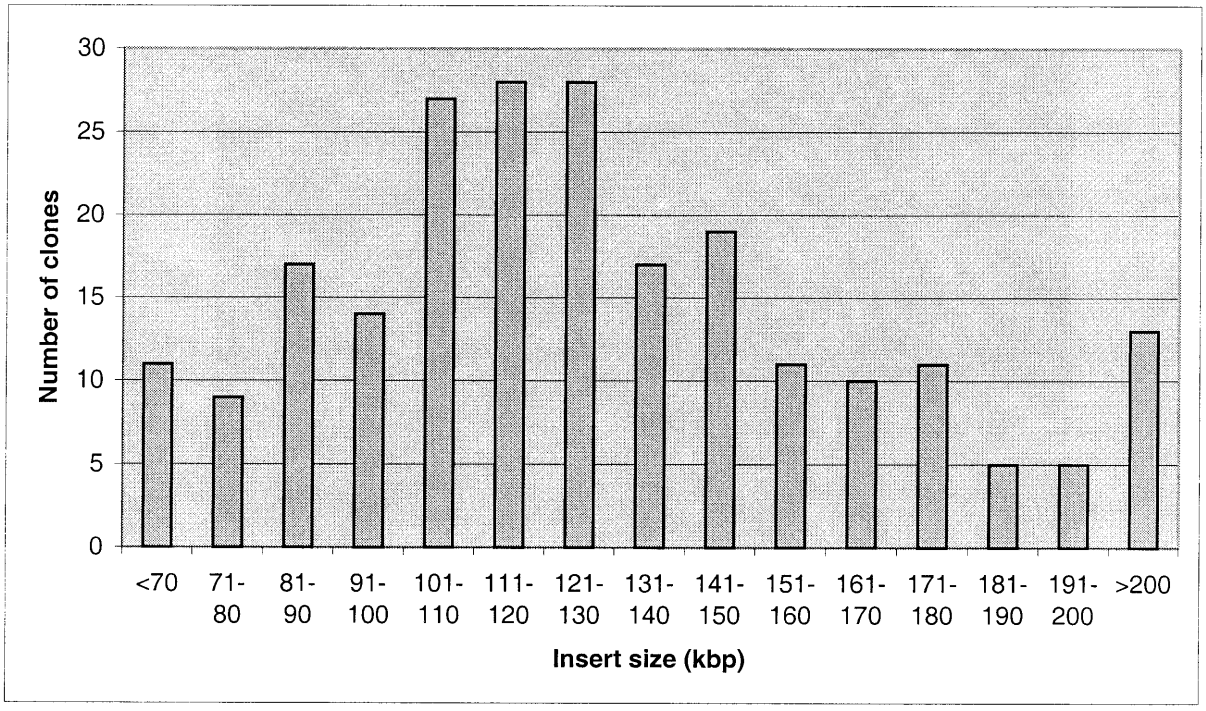

Fig. 2 Insert size distribution of BAC clones in the sugarcane BAC library. To estimate the insert size range, BAC DNA from 240 randomly selected clones was analyzed, as shown in Fig. 1 A. Results indicate that the average insert size is 130 kbp with over 75% of the clones >100 kbp

BAC library screening and BAC fingerprinting

High-density colony filters for hybridization based-screening of the library were prepared using the Genetix Q-BOT. Clones were gridded in double spots using a 4 × 4 array with six fields per 22.5×22.5 cm of nitrocellulose (Hybond NT) filter. This gridding pattern allows 18,432 clones to be represented per filter. Library screening was performed using five filters (labeled A–E) such that one filter set represented 92,160 clones. Colony filters were processed and hybridized using standard techniques (Sambrook et al. 1989). Screening for chloroplast DNA content was performed as described by Zhang et al. (1996). Screening with 11 single-copy sugarcane RFLP mapping probes (Grivet et al. 1996) and three different genomic clones for genes from other grass species was performed to evaluate sequence representation per amount of genome coverage. Five of the same probes used for screening the sugarcane filters were also used to screen two rice BAC libraries (Zhang et al. 1996) and one sorghum library (Woo et al. 1994) for comparison. After determining the addresses of each hit, clones were fingerprinted and analyzed as described previously by Marra et al. (1997). Fingerprint gel analyses were performed to determine relatedness of the clones.

Results and discussion

BAC library construction and characterization

We have constructed a sugarcane BAC library using the cultivar R570 which is suitable for physical mapping and map-based cloning. The library consists of 103,296 clones stored in 269 384-well microtiter plates. Less than 2% of the clones do not contain inserts as judged by random analysis of BACs sampled from the library. A random sampling of 240 BACs taken from the library during the course of library construction indicated an average insert size of 130 kbp with a range of 40–280 kbp. Figure 1 A shows 28 randomly selected clones digested with *Not*I to release the insert. The two *Not*I sites in pBeloBAC11 flank the multicloning site. Because *Not*I recognizes an 8-bp GC sequence and the sugarcane genome is relatively GC rich, digestion typically generates vector plus 1–8 insert bands per BAC clone based on our data. Fig. 1 B shows a Southern blot of the gel in Fig. 1 A probed with total sugarcane genomic DNA. As indicated by the strongly hybridizing lanes, many of the BACs contain highly repetitive DNA. However, 50% of the BACs contain low-copy DNA as indicated by insert bands not hybridizing, or else only weakly hybridizing, on the Southern blot.

To determine the size distribution of BAC clones, the 240 BACs analyzed with *Not*I digests were grouped by insert size and the insert size of each clone was plotted against the frequency of each group of clones represented in the library (Fig. 2). Based on this analysis, over 80% of the clones in the library have an average insert size greater than 100 kbp. Of the clones larger than 100 kbp, 40% are between 100 and 130 kbp and 40% are greater than 130 kbp. Based on the average insert size and a haploid genome size of 3 000 Mb (Arumuganthan and Earle 1991) the coverage of the library is about 4.5 genome-equivalents, resulting in a 98% probability of recovering any specific sequence of interest.

To obtain an estimate of the representation of chloroplast DNA in the library, the filters were screened with three different chloroplast genes spaced equidistantly around the 133-kbp barley chloroplast genome. Results from this screening showed that approximately 0.29% of

Table 1 Sugarcane BAC library filter-hybridization results using 11 single-copy sugarcane RFLP mapping probes anchored to nine different linkage groups (LG). Also included are hybridization results for three gene probes

Probe	LG	No. hits	Hit addresses
SSCIR 73	I	19	32N13, 35H11, 40O1, 62F1, 62I5, 104F12, 106I9, 113P18, 141L17, 149M23, 157E12, 158L12, 170J4, 172L3, 181H5, 188O23, 193F21, 213J21, 29O18
SSCIR74	V	31	27I7, 28C7, 28N13, 36F18, 36F20, 36F24, 36L10, 37P21, 61K19, 65O22, 88J6, 92C7, 98F18, 102L20, 106D17, 107F19, 119O10, 121A22, 163A14, 167G22, 176M6, 180M19, 187P22, 188O22, 193L20, 197G9, 202B21, 203J22, 213A13, 216D19, 219M14
SSCIR 78	III	42	17H2, 26B5, 26F17, 37E5, 39P17, 50A2, 50A17, 52N23, 54K11, 65O2, 68B14, 81L11, 89I9, 91L22, 98A5, 98G3, 114B8, 114O8, 121C20, 124F3, 132I11, 133N7, 137G2, 139H17, 142L24, 148B23, 149I5, 156K19, 170C22, 171E4, 179F16, 181B24, 184H21, 185L4, 190J3, 195D22, 198L6, 211M1, 216E16, 216G21, 232E1, 234G3
SSCIR 79	II	8	30N11, 35D21, 67E24, 143H21, 170F15, 189K21, 197M22, 230G22
SSCIR 101	IV	60	8G8, 11M10, 29J2, 47N5, 50D15, 50H11, 53M3, 57M8, 65P5, 73J9, 83L6, 91B19, 104G13, 104G14, 112B2, 112F24, 125E16, 135K19, 139N24, 145I9, 146O21, 157C15, 163H19, 167F4, 171J9, 171K18, 173L1, 173M16, 174H9, 179N21, 181M17, 184I18, 185G18, 185G21, 188B15, 188L9, 191J21, 194C1, 195I8, 195N7, 196N11, 196E14, 198M4, 200J5, 200O19, 203H16, 204I3, 204D12, 208G17, 212D21, 214O13, 215B20, 215O12, 216H16, 226G17, 228B16, 228P17, 230J20, 238N9
SSCIR60	III +U2/U3	23	4G10, 9E1, 10J15, 14M16, 19O15, 21E20, 24B14, 31L21, 46D8, 86L20, 87C21, 141L1, 142O14, 149M21, 151F12, 154D4, 207E12, 212L21, 214A12, 218E9, 230L20, 232M17, 238B16
SSCIR92	VIII	8	20D23, 22D13, 82J13, 144K17, 154P17, 204H22, 221M13, 224M9
SSCIR 194	VII	19	11B9, 35J13, 43A3, 50D6, 69I18, 69O23, 73P5, 74F24, 75M3, 112O15, 126D9, 134O2, 135N4, 146N19, 158I6, 164L16, 198A9, 207K20, 227L11
SSCIR 256	VI	14	1G20, 1L5, 37E13, 49K23, 56P15, 73J10, 81P17, 105C9, 136M13, 157P2, 164I4, 195N23, 227N5, 240M22
SSCIR257	X	20	14A20, 14K17, 29O19, 31K9, 32C6, 48E15, 49E5, 51B24, 85I15, 114B9, 114I24, 142P9, 165D24, 165J19, 177M3, 183K21, 184C5, 187B7, 228A21, 233B18
Adh1[1]	–	13	24F15, 35F1, 38A17, 39A16, 51L1, 102M23, 107J17, 120C21, 126J21, 127F10, 147G11, 169L11, 172H13
Aga5	–	14	7O3, 22C12, 27I16, 33L20, 119J13, 137B4, 143D5, 161H15, 169L2, 179H18, 204N3, 206H2, 209H16, 228K24
Sh2	–	45	3M6, 7B15, 7B16, 7C7, 9B1, 12A1, 13A18, 22D5, 22M20, 26K6, 29G4, 38I8, 38K14, 38M8, 51N2, 52I3, 54I3, 54N6, 63G10, 65G1, 66E14, 87I1, 91D11, 94E5, 99D20, 130F11, 133J7, 143N19, 146B3, 148B7, 152G10, 160O6, 161B11, 174G16, 177O18, 182N4, 198L13, 199A10, 212D23, 216J5, 230F15, 235B15, 238M13, 240H12, 240M3
CDSR029[2]	VII	13	4O5, 5C19, 5C20, 17P9, 19K12, 62L4, 68H8, 107M24, 140E14, 161B16, 166O19, 179N14, 187M21

[1] The *Adh1* (maize), *Aga5* (barley), and *Sh2* (maize) probes have not been mapped in sugarcane cv R570
[2] The CDSR029 probe is linked to the rust resistance gene

library sequences are chloroplast DNA (data not shown). The exceptionally low chloroplast DNA content of the library was most likely due to the use of the innermost layers of young shoots as a tissue source for nuclei isolation.

BAC library screening

In order to evaluate various aspects associated with map-based cloning in sugarcane, library screening was performed using ten different sugarcane RFLP mapping probes anchored to nine different linkage groups. These probes were a subset of those developed and used to generate a genetic map of sugarcane (Grivet et al. 1996). Screening with these probes resulted in a range of 8–60 hits (Table 1). The data in Table 1 describe probe size, linkage group, and addresses for all positive clones. Five of the sugarcane RFLP probes were also hybridized to a sorghum library (Woo et al. 1994) and two rice libraries (Zhang et al. 1996) for comparative study. In comparing the three libraries using these five probes, the average number of hits per probe per genome equivalent was 1.3, 1.4, and 6.1 for rice, sorghum, and sugarcane, respectively. Due to the polyploid nature of the sugarcane genome, there was nearly a 4.5-fold increase in the number of hits per genome equivalent as compared to rice and sorghum.

BAC fingerprinting and contig analysis

In order to determine relationships between BAC clones identified with various probes, clones were *Hin*dIII fingerprinted, gel images scanned in a computer, digitized

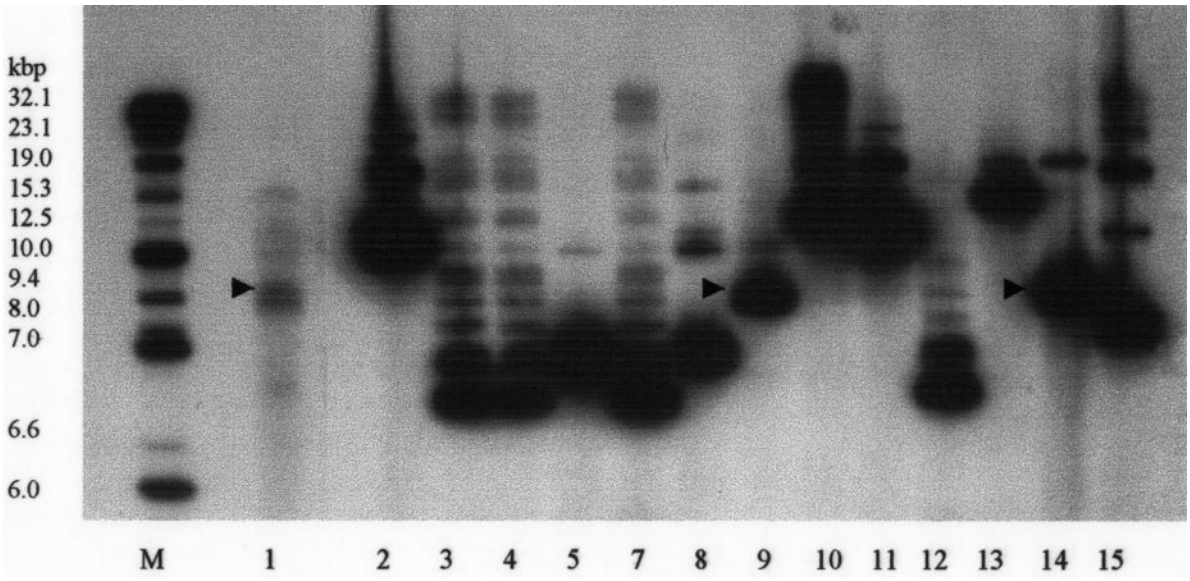

Fig. 3 Comparison between the restriction patterns of BAC clones and sugarcane DNA probed with an RFLP marker (CDSR029) linked to a rust resistance gene. BAC clones and genomic DNA were digested with *Dra*1. *Lane M*: molecular-weight standard. *Lane 1*: genomic DNA. *Lanes 2–15*: BAC clones identified from the sugarcane BAC library with the CDSR029 probe. *Arrow in lane 1* indicates an RFLP band associated with the rust gene. *Arrows in lanes 9 and 14* indicate the same RFLP band present in two of the BAC clones

using IMAGE software (Sanger Center, UK) and analyzed for contig formation as a pool using Fingerprint Contig (FPC) software (Sanger Center, UK). Stringency levels for the FPC analysis were set at a tolerance value of 7 and a cut-off value of 10^{-9}. Fingerprinting of BAC clones and analysis with FPC has previously been described in detail by Marra et al. (1997). The average insert size for sugarcane clones identified by the RFLP probes was 124 kbp, slightly lower than the average insert size determined by random sampling of clones throughout the library. Likewise, the rice and sorghum clones were also smaller than those reported from a random sampling of the libraries (Woo et al. 1994; Zhang et al. 1996). These data suggest the presence of more *Hin*dIII sites in regions containing single copy sequences in monocot genomes.

Results from FPC analysis showed that contigs could be easily formed with related clones, even in a large complex genome like that of sugarcane. Presumably, each contig represents a unique genomic region identified with a specific probe. The average number of contigs identified per probe was six and two for sugarcane and rice, respectively. The low number of positive clones for the sorghum library due to the small 2.5x genome coverage resulted in a lack of contig formation. In contrast to rice and sorghum, sugarcane showed results expected from a library constructed from a polyploid genome. Further, the large amount of data generated by screening the sugarcane BAC library with RFLP markers was easily sorted out using the fingerprint contig approach.

Identification of clones near a rust resistance gene

Genetic mapping was performed in progeny derived from the selfing of R570, and the RFLP marker CDRS029 was found to be linked to the rust resistance gene at a distance of 10 cM (Daugrois et al. 1996). We obtained this probe from the developers and hybridized it to the sugarcane BAC colony filters, resulting in 13 positive clones (Table 1). Corresponding BAC clones were taken from the library and fingerprinted using *Dra*I, the restriction enzyme used to map the rust resistance gene. Genomic DNA from R570 was also restricted with *Dra*I and run on the same gel as the BACs. In doing this, it was necessary to achieve relatively similar levels of stoichiometry between the genomic DNA and the BAC DNA. The BAC DNA and the genomic DNA were then run on a standard fingerprint gel. After Southern transfer, blots were probed with the CDSR029 marker and restriction fragments between DNA in the genomic lane compared with fragments of the BACs in the other 13 lanes (Fig. 3). This analysis showed that two of the BACs contained RFLP bands near 9 kbp that were similar to the fragment used to map the rust gene in the genomic lane. This indicates that these two BACs originated from the genomic segment corresponding to the RFLP marker.

Prospects for map-based cloning of the rust gene

The map-based cloning of disease resistance genes in plants is gaining momentum, even in large genomes like that of Barley (Buschges et al. 1997). We have demonstrated that BACs corresponding to specific genomic segments can be identified in a large insert library developed from a complex aneuploid polyploid genome. Therefore, the cloning of the rust resistance gene should be feasible given that a molecular marker(s) can be obtained closer to the gene. The CDSR029 RFLP marker is

still too far (10 cM) from the rust resistance gene to permit a chromosome walk. However, the identification of the two sugarcane BACs linked to the rust gene in this study has shown that, once closer markers are obtained, the identification of BAC clones in the same genomic region will be possible. At the present time, workers at CIRAD are developing new markers closer to the rust gene using a fine-scale mapping approach (Angelique D'Hont and J.C. Glaszmann, personal communication). It is anticipated that these new markers, combined with the present BAC library, will eventually lead to the cloning of the rust gene.

Acknowledgements Appreciation for technical assistance during the process of library replication and filter production is extended to Michael Atkins, John Bishop, and Jose Luis Goicoechea. Thanks to Jim Irvine (Texas A&M Univ., College Station) and Jim Miller (USDA-ARS, Homestead, Fl) for sending tissue of R570. Thanks also to Angelique D'Hont (Cirad, Montpellier, France) for sending sugarcane RFLP probes. This research was funded by a grant from the International Consortium for Sugarcane Biotechnology and by the Coker Endowment at Clemson University. Technical contribution no. 4471 from the South Carolina Agricultural Experiment Station.

References

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

Buschges R, Hollricher K, Panstruga R, Simons G, Wolter M, Frijters A, van Daelon R, van der Lee T, Diergaarde P, Groenendijk J, Topsch S, Vos P, Salamini F, Schulze-Lefert P (1997) The barley *Mlo* gene: a novel control element of plant pathogen resistance. Cell 88:695–705

Daugrois JH, Grivet L, Roques D, Hoarau JY, Lombard H, Glaszmann JC, D'Hont A (1996) A putative major gene for rust resistance linked with a RFLP marker in sugarcane cultivar 'R570". Theor Appl Genet 92:1059–1064

D,Hont A, Grivet L, Feldmann P, Rao P, Berding N, Glaszmann JC (1996) Characterisation of the double-genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. Mol Gen Genet 250:405—413

Grivet L, D'Hont A, Roques D, Feldmann P, Lanaud C, Glaszmann JC (1996) RFLP mapping in cultivated sugarcane (*Saccharum* ssp.): Genome organisation in a highly polyploid and aneuploid interspecific hybrid. Genetics 142:987–1000.

Marra MA, Kucaba TA, Dietrich NL, Green ED, Bownstein B, Wilson RK, McDonald KM, Hillier LW, McPherson JD, Waterston RW (1997) High-throughput fingerprint analysis of large-insert clones. Genome Res 7:1072–1084

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

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

Zhang HB, Choi S, Woo SS, 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