Mapping the *Trypanosoma cruzi* Genome: Analyses of Representative Cosmid Libraries

BioTechniques 21:686-693 (October 1996)

J. Hanke, D.O. Sánchez[1], J. Henriksson[2], L. Åslund[2], U. Pettersson[2], A.C.C. Frasch[1] and J.D. Hoheisel

Deutsches Krebsforschungszentrum, Heidelberg, Germany, [1]Fundación Campomar, Buenos Aires, Argentina, and [2]University of Uppsala, Uppsala, Sweden

ABSTRACT

In order to generate contiguous cosmid coverage of the genome of the protozoan parasite Trypanosoma cruzi *for large-scale sequence analysis, a cosmid library of 36 864 individual, primary clones was generated. Total genomic DNA of the reference strain CL Brener was fragmented both by partial digestion with* Mbo*I and by physical shearing. For cloning, a modified cosmid vector was used that simplifies analyses such as restriction mapping. The library's representation is about 25 genome equivalents, assuming a size of 55 Mb per haploid genome. No chimerism of inserts in the clones could be detected. The colinearity between cosmid inserts and genomic DNA was verified. Also, hybridizations to the gel-separated karyotype of the organism were carried out as a quality check. Gridded onto two nylon filters, the library was analyzed with a variety of probes. Apart from being used for combined physical and transcriptional mapping of the genome, library filters and clones are also available to interested parties.*

INTRODUCTION

Trypanosoma cruzi is a protozoan parasite that causes Chagas' disease. This chronic, incapacitating disease is prevalent in most of Latin America and affects 16 to 18 million people (20). A detailed knowledge of the structure of the parasite's genome is prerequisite for disclosing key mechanisms of the parasite's survival and pathogenicity, such as the attachment to host cells, invasion, subsequent differentiation and evasion from the immune response. Hence, the aim of molecular genetic analysis must be the sequencing of at least those parts of the genome that potentially contain the basic genetic information for such crucial aspects of the pathogen's life cycle. Access to this information will facilitate the development of specific tools for diagnosis and assist in efforts aimed at an effective treatment by immunotherapy or drugs. To this end, we have embarked on the mapping of the entire genome of *T. cruzi* at the level of cosmid cloning, which will form an indispensable basis for detailed sequence analyses on a large scale. The *T. cruzi* clone CL Brener was used, since it was selected as a reference strain at a meeting in 1994 held at Fiocruz (Brazil) sponsored by the World Health Organization (WHO). In this manuscript, the techniques of library generation and essential analyses of representation, colinearity, chimerism and other quality characteristics are described—procedures that are crucial for any library. Typical results from hybridization mapping are shown, as is the utilization of a modified cosmid vector for uncomplicated and convenient generation of restriction maps.

MATERIALS AND METHODS

Cosmid Vector

To add recognition sites for the enzymes I-*Sce*I and *Asc*I to the vector cloning site, the 31-bp cassette, depicted in Figure 4A, was inserted into the unique *Bam*HI site of Lawrist7 (5). After re-ligation and transformation into *E. coli*, several clones were picked for restriction analysis and two clones (Lawrist7-Sce1 and Lawrist7-Sce2) were selected containing the cassette in either orientation.

Cosmid Cloning

Epimastigotes of the cloned CL Brener stock (2) were grown in BHT culture medium and harvested as described previously (4). DNA was isolated following the protocol of Herrmann and Frischauf (9). The bulk of the DNA was in the size range of 300 to 400 kb, as determined by pulsed-field gel (PFG) electrophoresis (run for 24 h in 0.5× TBE buffer, pH 8.3, at 6 V/cm, initial switch time 60 s, final switch time 120 s). DNA concentrations were determined in a pH 12.0 buffer by measuring the increased fluorescence of DNA-intercalated ethidium (18). Insert DNA for cloning into cosmids was prepared by using the following two procedures (10,11):

For a partial *Mbo*I digest, 0.2 µg/µL DNA was cut with 0.025 U/µL *Mbo*I (New England Biolabs, Beverly, MA, USA) at 37°C. Aliquots were taken at various time points, and the reaction was stopped by the addition of EDTA and heat-inactivation at 68°C. Subsequent to phenol and chloroform extractions and an ethanol precipitation, the DNA was dephosphorylated using calf intestine phosphatase (Boehringer Mannheim, Mannheim, Germany). Vector arms of the cosmid Lawrist7-Sce1 were prepared by linearization with *Sca*I (New England Biolabs) and dephosphorylation, followed by cleavage of the *Bam*HI cloning site. Two micrograms of *T. cruzi* DNA, digested for 5–10 min, were ligated to 2 µg vector arms in 25 µL of 25 mM Tris-HCl, pH 7.5, 10 mM dithiothreitol, 100 mM NaCl, 7 mM $MgCl_2$, 1 mM spermidine, 1 mM ATP and 2.5 U ligase (Amersham, Arlington Heights, IL, USA) at 15°C overnight and stored for a week at 4°C. The ligation products were packaged in vitro using Gigapack® extracts (Stratagene, La Jolla, CA, USA), transfected into the *E. coli* strain DH5α and grown on 30 µg/mL kanamycin agar plates. Individual clones from randomly plated colonies were picked by a robotic device (Genetix, Christchurch, England, UK) and transferred to 384-well microplates containing 2× YT growth medium supplemented with freezing medium [36 mM K_2HPO_4, 13.2 mM KH_2PO_4, 1.7 mM Na-citrate, 0.4 mM $MgSO_4$, 6.8 mM $(NH_4)_2SO_4$, 4.4% (vol/vol) glycerol] and kanamycin, grown overnight at 37°C and then stored frozen at -70°C.

As a second approach, 4 µg of CL Brener DNA was physically sheared to an average size of 80 to 100 kb using a vortex mixer, combined with the same amount of vector arm DNA and blunt-ended with 6 U T4 DNA polymerase (New England Biolabs) in 20 µL 33 mM Tris-acetate, pH 7.9, 66 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM dithiothreitol and 100 µM of each nucleotide at 12°C for 30 min. After 10 min of heat-inactivation of the T4 DNA polymerase at 75°C, the sample was cooled to room temperature. ATP was added to a concentration of 0.5 mM and T4 DNA ligase (New England Biolabs) to 0.5 U/µL followed by incubation at 12°C overnight and subsequent storage at 4°C. Packaging and transfection were carried out as described above.

Clone Filters and Probe Hybridization

High-density filter arrays of the clone library were generated as described in detail earlier (10,11). Clone material was spotted by a modified version of the robotic device used above. The material was picked onto 22 cm × 22 cm nylon membranes (Hybond™-N^+; Amersham). Each filter contained 18 432 cosmid clones, each spotted in duplicate for easy and unequivocal detection. Colony growth was done on agar plates, after which the DNA was bound to the filter by a denaturation and subsequent neutralization step, followed by a removal of proteins using proteinase K or pronase E (Boehringer Mannheim).

Radioactive probes were generated by random hexamer priming (7) incorporating [α-^{32}P]dCTP (Amersham). Detailed protocols for probe hybridization, signal recording and subsequent stripping of the filters can be found elsewhere (10,21). Briefly, hybridizations were carried out in 0.5 M Na-phosphate, pH 7.2, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA, 0.1 mg/mL yeast tRNA at 65°C usually overnight. The probe concentration was about 0.5 million cpm/mL. Subsequently, the filters were briefly rinsed twice at room temperature in 40 mM Na-phosphate, pH 7.2, 0.1% SDS and washed in the same buffer by slowly rocking in a waterbath of 65°C for 15–30 min. Only filters hybridized with the identical probe were washed together. The filters were briefly blotted dry, covered (e.g., with Saran® Wrap) and exposed to film from 2 h to overnight at 70°C using an intensifying screen. Stripping was done in 5 mM Na-phosphate, pH 7.2, 0.1% SDS at about 90°C for 30 min, after which the procedure was repeated.

Clone Karyotyping

DNA of the *T. cruzi* cloned stocks CL Brener and, as a comparison, CA I/72 was run on pulsed-field gradient gels to separate the chromosomes in the size range below 1.8 Mb (run for 34 h in 0.5× TBE buffer at 6 V/cm, initial switch time 58 s, final switch time 169 s). The two quite distantly related strains belong to either of the two identified sub-populations of *T. cruzi*, the zymodemes Z1 (CA I/72) and Z2 (CL Brener). Filter transfer was carried out as reported in detail by Henriksson et al. (8). Hybridization was done as in the clone filter experiments.

RESULTS

Library Construction

For cloning *T. cruzi* DNA, a modified version of cosmid vector Lawrist7 (5) was used. The Lawrist series is a versatile system with two cos sites. It has been proven convenient and effective for the production of libraries from many different organisms (e.g., Reference 12). The average insert size of 37 kb and the relatively small vector portion of 5.5 kb, together with high DNA stability and a high-copy number, make it an excellent DNA source for mapping and subsequent sequence analysis (e.g., References 21 and 24).

The most convenient and most effective way of cloning genomic DNA into cosmids is following partial digestion with a frequently cutting enzyme. Here, the restriction nuclease *Mbo*I was used. Total *T. cruzi* DNA, including the kinetoplast DNA, was isolated and digested for 1–30 min with a limited amount of enzyme (Figure 1). The DNA of the 4-min and 8-min digests was used for ligation, although the majority of the DNA was still much larger than the about 28–44 kb that are clonable into Lawrist-based cosmids. This was done to avoid the generation of chimeric clones containing two unrelated fragments. This is particularly important since the occurrence of a band ladder in the 30-min digest depicted in Figure 1 clearly indicates the existence of a short repeat in the *T. cruzi* genome, which was cut with *Mbo*I. The size difference of about 200 bp between the bands corresponds well to the known 196-bp satellite DNA repeat (22). Due to the very limited partial digestion used for cloning, there is only a relatively low probability for the occur-

rence of chimeric clones in the library, even for the case where a satellite repeat is one of the fragments. However, even if there were clones containing two unique but unrelated fragments, they would not seriously disturb the mapping. Since each such clone would be a unique connection between two distant probes, such data and the respective clone would be ignored in the analysis until confirmed by at least one more clone, which is extremely unlikely to happen in the case of a random coligation event (10). As a result, these clones will not be considered for subsequent analyses such as sequencing. Clones chimeric with a 196-bp fragment would not necessarily show up during the mapping. However, their chimeric portion could be easily identified at the sequencing stage, since the repeat sequence would be located in a terminal position of the insert surrounded by two *Mbo*I sites.

As an additional mechanism for verifying the authenticity of the cloned

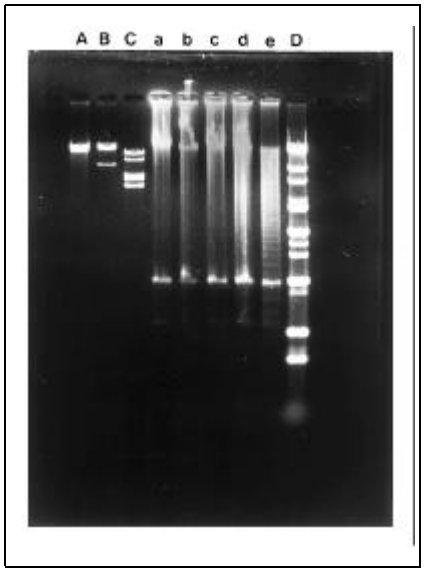

Figure 1. Partial *Mbo*I digestion of genomic *T. cruzi* DNA. DNA was cut for 1 min (a), 2 min (b), 4 min (c), 8 min (d) and 30 min (e). Standards are uncut λ DNA and λ cleaved with *Xba*I (A; bands unresolved), λ *Apa*I (B; 38 and 10 kb), λ *Bam*HI (C; 16.8, 7.2, 6.8, 6.5, 5.6, 5.5 kb) and a mixture of linear fragments (D; 23.1, 9.4, 6.5, 4.3, 3.9, 2.6, 2.3, 2.0 [undigested form V DNA of pUC18], 1.4, 0.7, 0.5, 0.4 kb). The 1.4-kb bands in lanes a–e represent linearized minicircles of the kinetoplast DNA. In the 30-min *Mbo*I digest (e), a band ladder of about 200 bp size difference can be seen, most likely being partially cut satellite DNA.

Table 1. Marker Hybridizations to the Cosmid Library

Probe	Locus	No. of Positives	Reference
Markers:			
Clone No. 1	200-kDa antigen Tc No. 1 gene	19 (+60)	13
Clone No. 10	antigen Tc No. 10 gene	10 (+21)	13
Tc36	85-kDa antigen	40	13
MUC.CA-3	mucin gene CA-3	33	6
MUC.RA-2	mucin gene RA-2	0	6
pUC18/6.4f	pyruvate dehydrogenase, α1 subunit gene	8	unpublished
pUCG18	antigen Tc No. 2 gene	26	3
TS	trans-sialidase gene	60	17
Cys.Prot	cruzipain (7:6.1)	150	1
Repeated Sequences:			
Satellite DNA	196-bp satellite repeat	3312	22
kDNA	kinetoplast minicircle	380	16
10A7	24S rDNA	21	unpublished
Homologous Sequences:			
cDNA1:31D7	*Drosophila* cDNA	25	unpublished

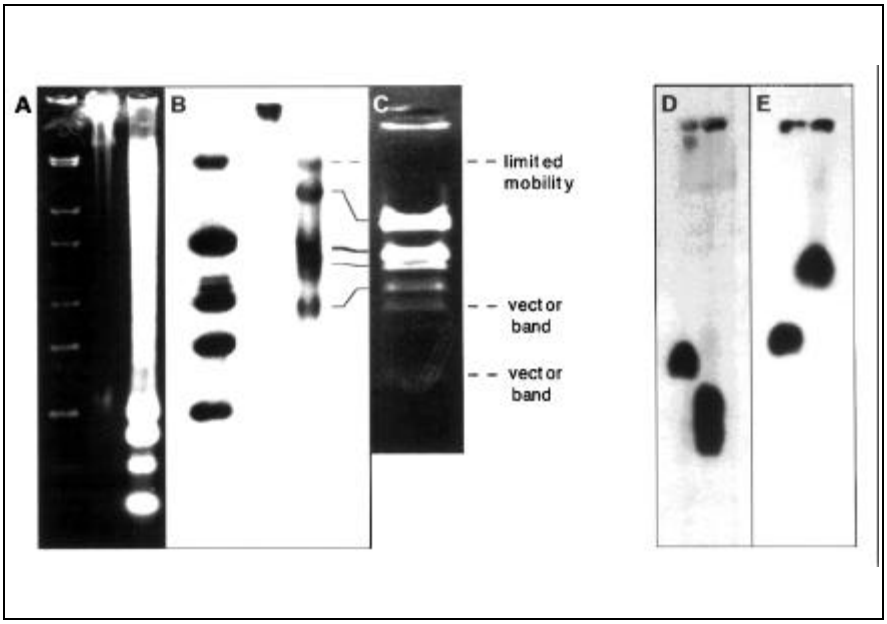

Figure 2. Quality checks. Genomic *T. cruzi* DNA was cut to completion with *Eco*RI and run on a 1% agarose gel. Panel A shows such a separation stained with ethidium bromide. From left to right, a mixture of standard fragments (λ *Hin*dIII plus fragments of 3912, 2686, 1444, 736 and 476 bp), uncut and *Eco*RI-cut genomic DNA are shown. Since the kinetoplast minicircle DNA consists of four partially conserved repeats of about 350 bp each (16), the *Eco*RI cleavage gave rise to four bands of 350, 700, 1050 and 1400 bp. Panel B: Total cosmid DNA of a selected clone was hybridized to a Southern blot made from a gel. The correspondence between positive fragments and the *Eco*RI cleavage pattern of the cosmid insert is shown. Probe hybridization to some fragments of the size standard is due to homologies between the cosmid vector and part of the λ DNA. Cosmids were also hybridized to PFG-separated chromosomes of the strains CL Brener (left) and CA I/72 (right) in Panels D and E.

DNA, about one-fifth of the cosmid library (5 genome equivalents) was made from sheared DNA. To keep the cloning efficiency at a reasonable level, the blunt-ended insert DNA was not dephosphorylated before ligation. Although this increases the probability of the occurrence of chimeric inserts, there was subsequent selection against chimera because of the high excess of vector DNA during ligation and the size selection of the packaging of the DNA into λ phage particles. Since on the average, the broken DNA was more than 80 kb in size, where only a small portion of the fragment population was actually in the size range that could be cloned in the Lawrist-vector, it was rather unlikely that two such small fragments had been ligated to the same vector molecule. Most chimeric inserts were too big to be packaged. The cloning efficiency was lower but comparable to the one achieved with the *Mbo*I-cut DNA, with the transfection rate being about 5000 clones per microgram DNA rather than 20 000 clones. Although the overall rate of chimerism could be higher, any particular artificial cohesion of the *Mbo*I-cut 196-bp repeat fragment can be ruled out for this library. Although not adequate for mapping the entire genome on its own account, the fivefold coverage of the second library should be sufficient to enable examination of any region potentially ambiguous in the library made from the *Mbo*I partial digests.

Clone Chimerism

During the partial digestion, the presence of an *Mbo*I site in the satellite DNA 196-bp unit and its apparently preferential cleavage produce an overproduction of many DNA fragments of multiple sizes of 196 bp (Figure 1). These fragments were predestined to counteract the selection for unique inserts by the addition of an excess of vector DNA during ligation and by the size selection of the packaging process. Testing for the repeat unit's frequency in the library therefore provides a tool for quality control of chimerism in the library. Hybridization with a satellite DNA probe (22) produced positive hybridization on 9.0% of *Mbo*I-cut inserts and 9.1% of sheared inserts of the library clones (e.g., Figure 3B). Since about 9% of the *T. cruzi* genome is known to consist of satellite DNA (22) and since the percentage of positives in the two library types are the same, it can be concluded that the degree of chimerism must be small.

Colinearity of Clone Inserts

Thirty-five cosmids from the entire library were picked at random and isolated. After *Pst*I cleavage, all clones showed different restriction patterns (not shown), adding up to about the expected average insert size of 37 kb. From restriction and hybridization patterns, it was found that some 3% of the clones consisted of DNA with no insert, a value in agreement with data from other libraries (12). To confirm the colinearity of cosmid inserts, genomic DNA was cut to completion with *Eco*RI and separated on a 1% agarose gel (Figure 2A). After a Southern transfer, several cosmids were hybridized to this DNA. In all cases examined, the band patterns of genomic and insert DNA were identical (e.g., Figure 2, B and C); *Eco*RI cuts off the vector DNA very closely on either side of the cloning site. Clones were also hybridized to filters containing PFG-separated chromosomes of the *T. cruzi* strains CL Brener and CA I/72. As expected from earlier results (8), the cosmids identified homologous sites in genetically equivalent chromosomes, which, nevertheless, showed large differences in size (e.g., Figure 2, D and E).

Hybridization Mapping

Altogether, 36 864 primary cosmid clones were picked into 384-well microplates. Assuming a haploid genome

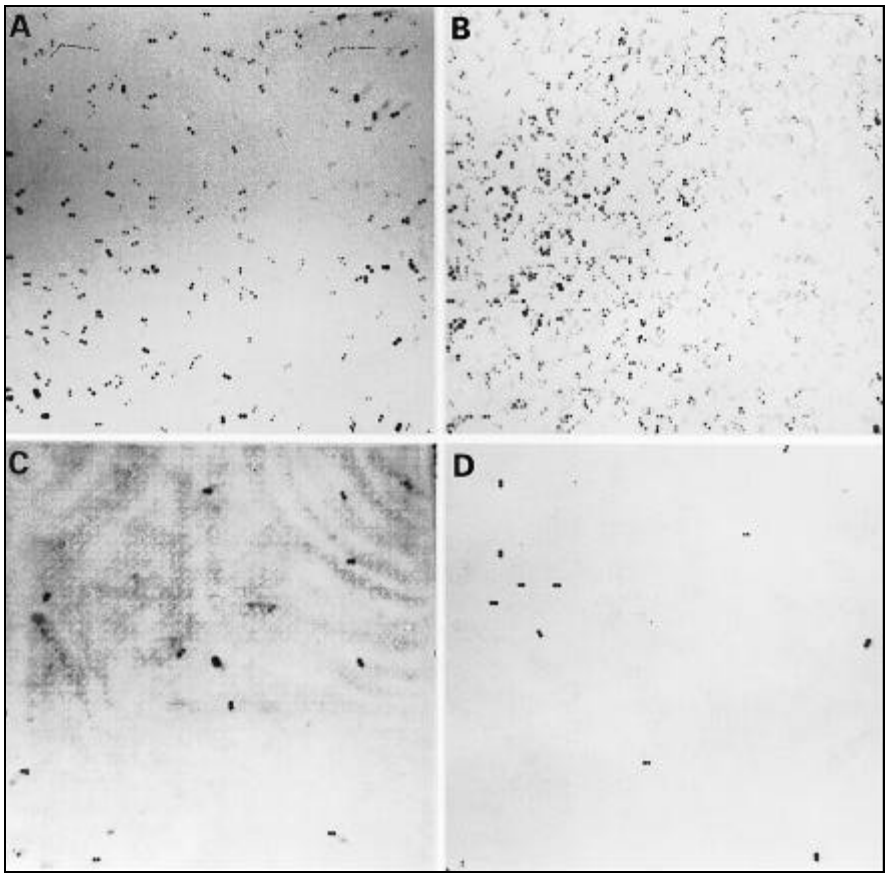

Figure 3. Probe hybridizations to cosmid library filters. Each filter represents half the library. Probes specific for the *T. cruzi* kinetoplast minicircle (A) and the 196-bp satellite repeat (B) were hybridized. Also shown are hybridizations of a cosmid clone taken from the library (C) and a unique marker-DNA for antigen Tc No. 2 (D). On the filter, the cosmid DNA is arranged in blocks of 4×4 clones. Within each block, each clone is present twice. Apart from serving as an internal control of positive hybridization, the orientation of the related clones within the box reveals the clone's identity.

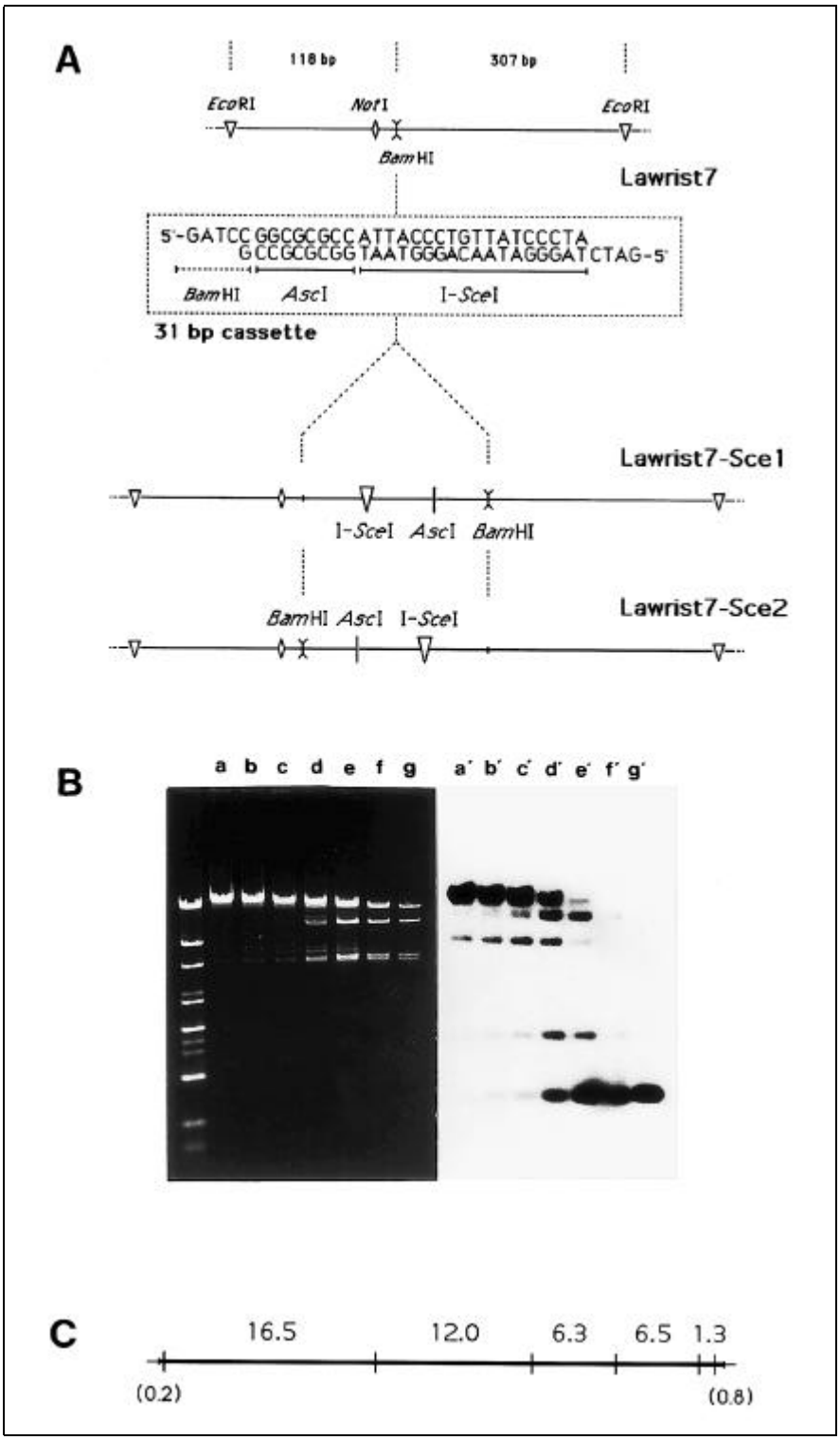

Figure 4. Restriction mapping of *T. cruzi* cosmids. Panel A: Extension of the cloning site of Lawrist7 by insertion of a 31-bp DNA cassette with the restriction sites of I-*Sce*I and *Asc*I. Panel B: Partial digest of 0.2 µg/µL I-*Sce*I-linearized cosmid DNA with 0.12 U/µL *Eco*RI. To determine the kinetics of cleavage for future digests, aliquots were taken from the reaction after 1-, 2-, 4-, 10-, 20-, 30- and 120-min intervals and run on an agarose gel (lanes a–g). After transfer to a nylon membrane, radioactive 30-mer d(CATACACATACGATTTAGGTGACACTATAG) was hybridized, binding to the vector adjacent to the right terminus of the insert (lanes a′–g′). From the bands in lane d′ (the top band containing two fragments) and a redundant experiment with a probe for the left end (not shown), the fragment sizes (in kb) and order could be deduced as shown in Panel C.

size of 55 Mb (14,23) and complete diploidy of the *T. cruzi* genome, the clones represent about 25 genome equivalents. For genome mapping, the entire library was gridded on high-density filters for hybridization (Figure 3). Mapping is currently under way following the strategy that was successfully used for completely covering the 14-Mb genome of *Schizosaccharomyces pombe* (11). However, rather than using cosmids from the library itself as probes, mostly cDNAs will be hybridized, thereby not only sorting the cosmids but also simultaneously locating the relevant genes on the map. In addition, cDNAs show a much lower proportion of repeat sequences, thus allowing the use of pooling schemes that will significantly reduce the number of experiments. Figure 3 shows some typical hybridization results. No competition was used for the suppression of cross-hybridization caused by repeat or vector sequences. Hybridizations of randomly selected cosmids back to the library yielded on average 56 (±39) clones per experiment, a number that is fairly close to the theoretical value of 50 that one would predict from the library's genome coverage.

In Table 1 the results of hybridizations with unique markers are listed. In addition, a few repeat sequences were checked for their occurrence in the library. The probe for the antigens Tc No. 1 and Tc No. 10 (13) produced two distinct sets of positive signals, strong hybridization to 19 and 10 clones, respectively, and weaker hybridization to others, the latter of varying intensity for Tc No. 1. A similar result with a *Drosophila* library could be explained by the existence of other genes in the genome homologous for at least part of the sequence (unpublished results). The variation in signal intensities suggests a cross-hybridization of the repeat unit known to be part of gene No. 1 (13). The 150 positive cosmids identified by the cysteine protease probe confirms the data from hybridization experiments on PFG-separated chromosomes that there are fewer copies in strain CL Brener than in some others. Of all probes used, only the hybridization of a DNA of a mucin gene from *T. cruzi* strain RA112 did not produce any signal on the filters at high stringency.

However, a DNA of a related mucin gene from *T. cruzi* strain CA I/72 did identify cosmids (Table 1). The mucin genes are known to differ in their central domain, while their 5′ and 3′ ends show a higher degree of homology (6). The results of the hybridizations show that the variation at the DNA level is substantial enough to cause a difference in hybridization behavior. This conclusion is supported by a second hybridization with the Muc.RA-2 probe at much lower stringency. Besides producing higher background, clones positive with the mucin probe from CA I/72 showed up in such an experiment.

Reference System

In a system first proposed and set up by Lehrach et al. (15), filters containing the complete library are distributed as a resource, and identified clones are provided. Clone filters have been used for more than 60 subsequent hybridization experiments. By using this reference system, other laboratories can work on a more detailed (e.g., sequence) analysis of a genomic region of their interest even before the physical map of the region is completed, apart from linking probes to the developing physical map, and vice versa. Eventually, their markers will be positioned in the *T. cruzi* genome without any additional effort on their part. Since one particular library is being used as a common tool, the data generated by different groups will be more easily compared.

Restriction Mapping

Once clone overlaps have been identified and formed into contiguous clone arrays, a restriction map made from a minimal clone set is, for two reasons, the next step toward a sequence analysis. First, such an analysis was found crucial in large-scale analyses, such as the European Yeast Sequencing Project (25), for example, in order to avoid much redundant sequencing of DNA regions present in overlapping cosmids. Second, restriction data are control instruments; comparison of the sequence and the restriction map allows an initial evaluation of sequence accuracy.

The extended version of Lawrist7 allows a simple restriction mapping strategy for the *T. cruzi* cosmids. The insertion of the 31-bp DNA cassette into the unique *Bam*HI cloning site of Lawrist7 introduced the 18-mer and 8-mer recognition sequences of the enzymes I-*Sce*I and *Asc*I (Figure 4A). Insertion was such that, only at one end of the cassette, a new and, hence, unique *Bam*HI site was re-created. I-*Sce*I is extremely unlikely to cut any cloned insert DNA. It will therefore linearize all recombinant clones. The enzyme is much less affected by the DNA quality than λ terminase, which requires highly purified DNA and, in our hands, gave varying results with DNA prepared by alkaline lysis. Partial cleavage of the linear DNA and a subsequent hybridization with a probe specific for either end of the vector DNA (e.g., Figure 4B) will instantly produce a restriction map (Figure 4C) as described (19). The I-*Sce*I site can also be useful for other aspects of DNA handling. For example, cleavage with I-*Sce*I produces non-palindromic 4 nucleotide 3′ protrusions, which can be utilized for strand protection in unidirectional deletion experiments. The termini generated by *Not*I or *Asc*I can be used as substrates for exonuclease III.

DISCUSSION

As part of a program initiated by the Tropical Disease Research (TDR) Special Programme of the WHO, molecular aspects of the survival and contagious mechanisms of *T. cruzi* are being studied. To obtain detailed (sequence) information that is prerequisite for an understanding at this level, we report here the generation and characterization of a genomic cosmid library. The crucial importance of the quality of a representative library, and a thorough examination thereof, is still widely underestimated. Many aspects can be significant and consequential for the success of subsequent analyses. The data presented above demonstrate that the library fulfills the requirements as a useful tool for the analysis of the genome. It presents the backbone of an ongoing project that aims at the complete clone coverage of the *T. cruzi* genome and the subsequent sequencing of at least those portions potentially important for both diagnosis and treatment of Chagas' disease.

A clone coverage much lower than the 25 genome equivalents of the *T. cruzi* cosmid library is generally satisfactory for the identification and isolation of specific DNA pieces by library screening. However, it is inadequate for contiguous mapping. Not only will there be a few real gaps, but, more importantly, there will be regions of clone representation that are insufficient for unambiguous map generation. The problem is further complicated by variations in the cloning efficiency for different DNA regions. The observed standard variation value of 39 around the average of 56 clones identified by randomly selected cosmids evidently highlights the fact. The sheared library will help in this respect, because regions with few or no appropriate restriction sites are made clonable by this procedure. Calculations on the basis of the *S. pombe* mapping experiments (11) indicated that a tenfold representation was the minimum necessary for the formation of a contiguous and unambiguous map of the genome. A higher

degree of redundancy proved to be extremely helpful for the analysis and was actually essential for resolving problematic areas.

The marker hybridizations demonstrate the ease with which the genomic environment of genes can be isolated for more detailed analyses. The very same technique is being used for the mapping of the entire genome. Using hybridization rather than restriction fragment mapping, for example, has the advantage that physical clone sorting can be associated with the generation of an at least partial transcriptional map. This is in addition to the fact that, due to the highly parallel filter presentation and analysis, far less experimental effort per individual clone is required. By mainly using cDNAs as probes, genes will be located as an integral part of the map production. This correlation between the data will facilitate subsequent analyses. The approach also emphasizes the accessibility of the library as a common tool for the entire scientific community. The power of the hybridization mapping procedures is such that a near complete coverage of the 55-Mb genome should be achievable within a reasonable time frame. Sequencing analysis on library material has already been started and will continue concurrently with cosmid mapping.

ACKNOWLEDGMENTS

The skillful technical assistance of Sandra Schwarz and Maria Rydåker is gratefully acknowledged. We thank Elmar Maier and Hans Lehrach for access to their clone-picking device. We have received financial support from the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR), the Ministry of Culture and Education of Argentina and the Swedish Agency for Research Co-operation with Developing Countries (SAREC). D.O.S. and A.C.C.F. are researchers from the Consejo Nacional de Investigaciones Científicas y Técnicas and from Universidad de Buenos Aires, Argentina.

REFERENCES

1. **Åslund, L., J. Henriksson, O. Campetella, A.C.C. Frasch, U. Pettersson and J.J. Cazzulo**. 1991. The C-terminal extension of the major cysteine proteinase (cruzipain) from *Trypanosoma cruzi*. Mol. Biochem. Parasitol. *45*:345,347.
2. **Brener, Z.** 1965. Comparative studies of different strains of *Trypanosoma cruzi*. Ann. Trop. Med. Parasitol. *59*:19-26.
3. **Buschiazzo, A., O.E. Campetella, R.A. Macina, S. Salceda, A.C.C. Frasch and D.O. Sanchez.** 1992. Sequence of the gene for a *Trypanosoma cruzi* protein antigenic during the chronic phase of human Chagas disease. Mol. Biochem. Parasitol. *54*:125-128.
4. **Cazzulo, J.J., B.M. Franke de Cazzulo, J.C. Engel and J.J.B. Cannata.** 1985. End products and enzyme levels of aerobic glucose fermentation in Trypanomatids. Mol. Biochem. Parasitol. *16*:329-343.
5. **de Jong, P.J., C. Chen and J. Garnes.** 1989. Application of PCR for the construction of vectors and the isolation of probes, p. 205-210. *In* H.A. Erlich, R.A. Gibbs and H.H. Kazazian (Eds.), Polymerase Chain Reaction. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
6. **Di Noia, J.M., D.O. Sanchez and A.C.C. Frasch.** 1995. The protozoan *Trypanosoma cruzi* has a family of genes resembling the mucin genes of mammalian cells. J. Biol. Chem. *270*:24149-24149.
7. **Feinberg, A.P. and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. *132*:6-13.
8. **Henriksson, J., B. Porcel, M. Rydåker, A. Ruiz, V. Sabaj, N. Galanti, J.J. Cazzulo, A.C.C. Frasch and U. Pettersson.** 1995. *Trypanosoma cruzi* genome project: chromosome-specific markers reveal conserved linkage groups in spite of extensive chromosomal size variation. Mol. Biochem. Parasitol. *73*:63-74.
9. **Herrmann, B.G. and A.-M. Frischauf.** 1987. Isolation of genomic DNA. Methods Enzymol. *152*:180-183.
10. **Hoheisel, J.D., E. Maier, R. Mott and H. Lehrach.** 1996. Integrated genome mapping by hybridization techniques, p. 319-346. *In* B. Birren and E. Lai (Eds.), Analysis of Non-Mammalian Genomes—A Practical Guide. Academic Press, New York.
11. **Hoheisel, J.D., E. Maier, R. Mott, L. McCarthy, A.V. Grigoriev, L.C. Schalkwyk, D. Nizetic, F. Francis and H. Lehrach.** 1993. High-resolution cosmid and P1 maps spanning the 14-Mbp genome of the fission yeast *Schizosaccharomyces pombe*. Cell *73*:109-120.
12. **Hoheisel, J.D., M.T. Ross, G. Zehetner and H. Lehrach.** 1994. Relational genome analysis using reference libraries and hybridisation fingerprinting. J. Biotechnol. *35*:121-134.
13. **Ibañez, C.F., J.L. Affranchino, R.A. Macina, M.B. Reyes, S. Leguizamon, M.E. Camargo, L. Åslund, U. Pettersson and A.C.C. Frasch.** 1988. Multiple Trypanosoma cruzi antigens containing tandemly repeated amino acid sequence motifs. Mol. Biochem. Parasitol. *30*:27-34.
14. **Kooy, R.F., F. Ashall, M. Van-der-Ploeg and J.P. Overdulve.** 1989. On the DNA content of *Trypanosoma cruzi*. Mol. Biochem. Parasitol. 36:73-76.
15. **Lehrach, H., R. Drmanac, J.D. Hoheisel, Z. Larin, G. Lennon, A.P. Monaco, D. Nizetic, G. Zehetner and A. Poustka.** 1990. Hybridisation fingerprinting in genome mapping and sequencing, p. 39-81. *In* K.E. Davies and S. Tilghman (Eds.), Genome Analysis Volume 1: Genetic and Physical Mapping. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
16. **Macina, R.A., D.O. Sanchez, D.A. Gluschankof, O.R. Burrone and A.C. Frasch.** 1986. Sequence diversity in the kinetoplast DNA minicircles of *Trypanosoma cruzi*. Mol. Biochem. Parasitol. *21*:25-32.
17. **Parodi, A.J., G.D. Pollevick, M. Mautner, A. Buschiazzo, D.O. Sanchez and A.C.C. Frasch.** 1992. Identification of the gene(s) coding for the trans-sialidase of *Trypanosoma cruzi*. EMBO J. *11*:1705-1710.
18. **Pohl, F.M., R. Thomae and A. Karst.** 1982. Temperature dependence of the activity of DNA-modifying enzymes: endonucleases and DNA ligase. Eur. J. Biochem. *123*:141-152.
19. **Rackwitz, H.-R., G. Zehetner, H. Murialdo, H. Delius, J.H. Chai, A. Poustka, A. Frischauf and H. Lehrach.** 1985. Analysis of cosmids using linearisation by phage lambda terminase. Gene *40*:259-266.
20. **Schenone, H. and A. Rojas.** 1989. Pragmatic data and observations related to the epidemiology of Chagas disease. Boletin Chileno de Parasitologia *44*:66-86.
21. **Scholler, P., A.E. Karger, S. Meier-Ewert, H. Lehrach, H. Delius and J.D. Hoheisel.** 1995. Fine-mapping of shotgun template-libraries; an efficient strategy for the systematic sequencing of genomic DNA. Nucleic Acids Res. *23*:3842-3849.
22. **Sloof, P., J.L. Bos, A.F. Konings, H.H. Menke, P. Borst, W.E. Gutteridge and W. Leon.** 1983. Characterization of satellite DNA in *Trypanosoma brucei* and *Trypanosoma cruzi*. J. Mol. Biol. *167*:1-21.
23. **Thompson, C.T. and J.A. Dvorak.** 1989. Quantitation of total DNA per cell in an exponentially growing population using the diphenylamine reaction and flow cytometry. Anal. Biochem. *177*:353-357.
24. **Tynan, K., A. Olsen, B. Trask, P. de Jong, J. Thompson, W. Zimmermann, A. Carrano and H. Mohrenweiser.** 1992. Assembly and analysis of cosmid contigs in the CEA-gene family region of human chromosome 19. Nucleic Acids Res. *20*:1629-1636.
25. **Vassarotti, A., B. Dujon, P. Mordant, H. Feldmann, W. Mewes and A. Goffeau.** 1995. Structure and organisation of the European Yeast Genome Sequencing Network. J. Biotechnol. *41*:131-138.

Received 31 January 1996; accepted 24 June 1996.

Address correspondence to:

Jörg Hoheisel
Deutsches Krebsforschungszentrum
Im Neuenheimer Feld 506
D-69120 Heidelberg, Germany
Internet: j.hoheisel@dkfz-heidelberg.de