Electronic Journal of Biotechnology ISSN: 0717-3458
Vol.9 No.5, Issue of October 15, 2006

Received November 3, 2005 / Accepted April 4, 2006
DOI: 10.2225/vol9-issue5-fulltext-3
RESEARCH ARTICLE

Genome analysis of thirteen Colombian clostridial strains by pulsed field gel electrophoresis

Diana Milena Quilaguy Ayure
Bioprocesses and Bioprospecting Group
Instituto de Biotecnología
Universidad Nacional
Ciudad Universitaria, Edificio Manuel Ancizar
Bogotá, Colombia
Tel: 571 3165000. Ext.16970/71
Fax: 571 3165415
E-mail: dmquilaguya@unal.edu.co

Zulma Rocío Suárez Moreno*
Bioprocesses and Bioprospecting Group
Instituto de Biotecnología
Universidad Nacional
Ciudad Universitaria, Edificio Manuel Ancizar
Bogotá, Colombia
Tel: 571 3165000. Ext.16972/70
Fax: 571 3165415
E-mail: zrsuarezm@unal.edu.co

Fabio Ancizar Aristizábal Gutierrez
Bioprocesses and Bioprospecting Group
Instituto de Biotecnología
Universidad Nacional
Ciudad Universitaria
Bogotá, Colombia
Tel: 571 3165000. Ext.14643
Fax: 571 3165415
E-mail: faaristizabalg@unal.edu.co

Jose Mauricio Bernal Morales
Bioprocesses and Bioprospecting Group
Instituto de Biotecnología
Universidad Nacional
Ciudad Universitaria, Edificio Manuel Ancizar
Bogotá, Colombia
Tel: 571 3165000. Ext.16970/71
Fax: 571 3165415
E-mail: jmbernalmo@unal.edu.co

Dolly Montoya Castaño
Bioprocesses and Bioprospecting Group
Instituto de Biotecnología
Universidad Nacional
Ciudad Universitaria, Edificio Manuel Ancizar
Bogotá, Colombia
Tel: 571 3165000. Ext.16952
Fax: 571 3165415
E-mail: dmontoyac@unal.edu.co

Financial support: Colciencias, Universidad Nacional de Colombia.

Keywords: *Clostridium*, genome size, PFGE, pSOL1.

Abbreviations: AFLP: amplified fragment length polymorphism
ATCC: American Type Culture Collection
bp: base pairs

* Corresponding authors

(CHEF)-PFGE: contour-clamped homogeneous electric field
EC: lysis buffer
EDTA: ethylenediaminetetracetic acid
ES buffer: EDTA–sarcosine buffer
ESP: EDTA-sarcosine-proteinase buffer
ET buffer: EDTA-Tris Buffer (Tris-HCl 10 mM, EDTA 100 mM)
IBUN: Instituto de Biotecnología de la Universidad Nacional
Mbp: mega base pairs
OD: optical density
ORF: open reading frame
PCR: polymerase chain reaction
PFGE: pulsed field gel electrophoresis
PMSF: phenylmethylsulphonyl fluoride
RCM: reinforced clostridial medium
RPM: Revolutions Per Minute
TBE: Tris-borate-EDTA buffer
TE: Tris-EDTA buffer (Tris-HCl 10 mM, EDTA 1 mM)

Pulsed field gel electrophoresis was used for estimating the size of the genome and evaluating the presence of megaplasmids in 13 native Colombian solventogenic *Clostridium* strains. DNA preparation and purification were optimised for obtaining differentiated restriction fragments in electrophoresis. Genomic DNA was digested with *Apa*I, *Eco*52I, *Sma*I and *Xho*I enzymes. Estimated genome size for native strains ranged from 4.0 to 4.2 mega base pairs. Larger sized plasmids were detected and the presence of genes related to megaplasmid pSOL1 was determined by polymerase chain reaction. *adc* gene region amplification suggested that genes related to solventogenesis in native strains may be located in an extra-chromosomal element. Determining genome size provides useful information aimed at enhancing native strains' solvent production.

Research aimed at gaining knowledge about genetic material has often relied on estimating genome size (Fonstein and Haselkorn, 1995). PFGE has been used for estimating the genome size of different *Clostridium* strains*:* 3.5-6.5 Mbp for *C. acetobutylicum* (Wilkinson and Young, 1993), 5.3 Mbp for *C. saccharobutylicum* NCP 262 (Keis et al. 2001), 4.0 Mbp for *C. botulinum* type A (Lin and Johnson, 1995) and 3.8 Mbp for group II (Hielm et al. 1998). The *C. acetobutylicum* ATCC 824 genome was sequenced and its size was determined to be 3.9 Mbp. It was also determined that the pSOL1 megaplasmid size was 192 Kbp (Nölling et al. 2001).

This work is related to 13 native bacterial strains from the *Clostridium* genus, selected from a strain-bank (consisting of 178 isolates from different Colombian soils), based on their greater ability to produce total solvents than the *Clostridium acetobutylicum* ATCC 824 strain (Montoya et al. 2000). Native strains were molecularly characterised by sequencing the 16S rRNA gene (Montoya et al. 1999), DNA-DNA hybridisation (unpublished data), plasmid profile characterisation (Arévalo et al. 2002), amplified fragment length polymorphism (AFLP) (Jaimes et al. 2005) and pulsed field gel electrophoresis (PFGE) (Montoya et al. 2001). The strains' ability to produce 1,3 propanediol was also evaluated. Multivariate data analysis was used for taxonomically correlating phenotyping and genotyping results; it suggested that 10 out of the 13 native strains could be considered as being a new specie (unpublished data).

This research was aimed at estimating the size of the genome for native strains, detecting megaplasmids and evaluating the presence of the pSOL1 plasmid containing important genes encoding enzymes involved in solventogenesis. Determining the presence of solventogenic genes is useful for enhancing native strains' solvent production by means of metabolic engineering. Determining genome size would also contribute towards native strains' taxonomy.

MATERIALS AND METHODS

Bacterial strains and culture medium

Clostridium IBUN 22A, IBUN 125C, IBUN 140B, IBUN 62F, IBUN 95B, IBUN 13A, IBUN 18A, IBUN 18S, IBUN 62B, IBUN 137K, IBUN 158B, IBUN 18Q and IBUN 64A native strains were used in this study (Montoya et al. 2000). The *Clostridium acetobutylicum* ATCC 824 strain was used as pattern as it is known to be solventogenic.

Cells were activated from strains conserved in silica gel. They were grown in RCM medium (OXOID reinforced Clostridia medium) which had been previously gassed with nitrogen to create anaerobic conditions; all cultures were incubated at 37ºC following the methodology described by Montoya et al. 2000.

DNA preparation

DNA was prepared in agarose plugs, according to a modified protocol described by Montoya et al. 2001. Bacteria were grown to 0.3 to 0.4 OD (680 nm) in 40 ml RCM medium; up to 180 µg/ml thiamphenicol was then added and bacteria were incubated at 37ºC for 1 hr. Bacterial cultures were placed in an ice bath for 30 min.

Table 1. Number and sizes of restriction fragments[a] with *C. acetobutylicum* ATCC 824 genomic DNA.

Fragment Nº	*Apa*I	*Apa*I Standardised data[b]	*Eco*52I	*Eco*52I Standardised data[d]	*Sma*I	*Xho*I
1	706 ± 21	x	681	1,355	713	692
2	657 ± 25	x	592	530	654	620
3	523 ± 22	523 ± 22	465	485	628	506
4	485 ± 12	485 ± 12	201	440	484	454
5	413 ± 18	431 ± 18	184	390	316	404
6	379 ± 19	379 ± 16	149	275	274	352
7	352 ± 15	352 ± 15	106	270	259	322
8	295 ± 22	295 ± 18	74	200	215	265
9	283 ± 16	283 ± 16	13	145	187	229
10	212 ± 11	212 ± 11[c]	10	55	150	170
11	192 ± 12	192 ± 12			102	125
12	181 ± 12	181 ± 12			71	109
13	136 ± 12	136 ± 12			56	73
14	109 ± 9	109 ± 9				13
15	88 ± 11	88 ± 11				10
16	71 ± 7	71 ± 7				
Estimated genome size (Kbp)	**5,081**	**3,931**	**2,455**	**4,145**	**4,115**	**4,351**

[a]standard deviations were obtained from three determinations. Fragment sizes were determined with Lambda molecular markers (NEB).
[b]according to Wilkinson and Young, 1993.
[c]double band.
[d]according to Cornillot et al. 1997b.

OD was measured and bacterial cells were harvested at 5,000g for 10 min at 4ºC. Bacterial pellets were washed twice with 5 mL PETT IV buffer (10 mM Tris-HCl, 1M EDTA) and suspended in the same buffer in a volume:OD ratio as follows: 1 mL was added to 0.4 OD, 1.5 mL were added to 0.5 OD and 2 mL were added to 0.6 OD. All buffers were prepared in the laboratory just before using them (*i.e.* no commercial kits were used).

Bacterial cells were counted in a Newbauer chamber to determine the quantity of available DNA in each plug, as described in Birren and Lai (1993). Cells were suspended in an equal volume of InCert (FMC) low melting-point liquid agarose (1.2% in TE buffer (10 mM Tris-HCl 1mM EDTA pH 7.6) at 40ºC. This was mixed and 50 µl aliquots were pipetted into each plug-forming mould. The mixture was cooled on ice for 20-30 min. Each plug contained an average of 10^6 cells.

200 µl EC lysis buffer (6 mM Tris-HCl pH 7.6, 1M NaCl, 100 mM EDTA pH 8.0, 0.05% Triton 100X, 0.5% wt/vol N-lauroylsarcosine, 1.0 mg/ml lisozyme and 20.0 µg/ml RNAse A) were added to each plug; these were then shaken slowly at 37ºC for 48 hrs, the buffer being changed each 24 hrs. The buffer was changed for ES buffer (500 mM EDTA, 1% wt/vol N-lauroylsarcosine) and plugs were washed twice with 200 µl of this new buffer and incubated at room temperature for 5 min each.

Buffer was changed for 200 fresh µl ESP buffer (500 mM EDTA, 1% wt/vol N-lauroylsarcosine, 2 mg/mL proteinase K) and plugs were incubated overnight at 50ºC in a water bath. Then plugs were cooled at 4ºC for 10 min. This step was repeated three times. Buffer was changed for fresh TE + 2.5 µl PMSF and slowly shaken overnight at 37ºC. This was then changed for 200 µl TE and shaken for two hours at 37ºC. There were 2 further changes of fresh TE followed by incubation at 37ºC overnight. Plugs were then stored in 200 µl ET buffer (Tris-HCl 10 mM, EDTA 100 mM) at 4ºC until needed.

Bands were seen to have the best resolution during electrophoresis at an estimated 0.7 to 3.0 µg/ml DNA concentration, equivalent to 35 to 150 ng in each 50 µl agarose insert, corresponding to 0.8 to 3.5 x 10^8 cells/ml. Agarose concentration for preparing the inserts was modified from 1.2% to 1.5% (Montoya et al. 2001), bands being resolved better at less concentration (Figure 1). Some tests and modifications were made to Montoya's protocol for DNA treatment (Montoya et al. 2001). Two treatments were made with lysozyme, each one lasting 24 hrs at 37ºC; however, they did not improve DNA diffusion from insert to gel. Three treatments were made with proteinase K allowing DNA to run from the insert to the agarose gel; each one lasted 24 hrs at 50ºC. Digestion was also done with a thermophilic proteinase (Pretaq Rt41 A, Gibco-BRL), but results were similar to those obtained with proteinase K treatment. Proteinase K treatment was thereby selected for optimising the DNA purification protocol.

The quality of immobilised DNA in the agarose inserts was tested using 3 parameters: verifying lysis, endonuclease accessibility and DNA integrity. Effective lysis was deemed to have taken place when the inserts were translucent. Endonuclease accessibility was tested by DNA digestion with *Eco*RI. Agarose inserts were placed in

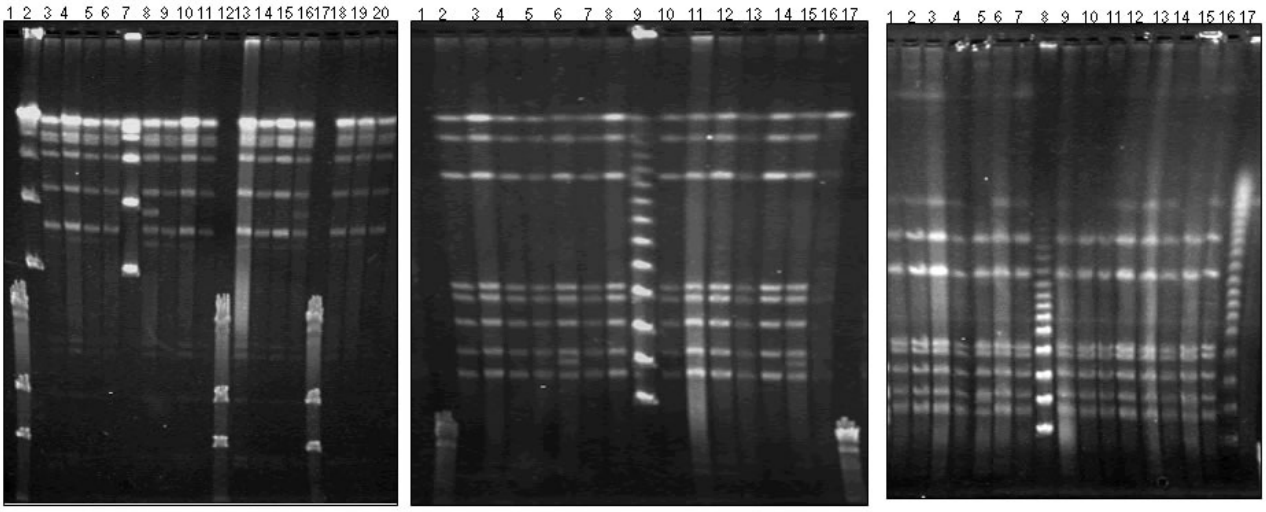

Figure 1. Native strain and *C. acetobutylicum* ATCC 824 DNA restriction fragments digested with *Eco*52I separated by PFGE.
(a) Lane 1: Lambda–*Hind*III marker; 2: Lambda concatamers (NEB); 3: *C. acetobutylicum* ATCC; 4: IBUN 22A; 5: IBUN 125C; 6: IBUN 140B; 7: Lambda concatamers (NEB); 8: IBUN 62F; 9: IBUN 95B; 10: IBUN 13A; 11: IBUN 18A; 12: Lambda–*Hind*III marker; 13: IBUN 18S; 14: IBUN 62B; 15: IBUN 158B; 16: IBUN 18Q; 17: Lambda–*Hind*III marker, 18: IBUN 64A; 19: IBUN 137K; 20: IBUN 137K Pulse time of 1 to 8 sec for 16 hrs.
(b) Lane 1: Lambda-*Hind*III marker; 2: *C. acetobutylicum* ATCC; 3: IBUN 22A; 4: IBUN 125C; 5: IBUN 140B; 6: IBUN 62F; 7: IBUN 95B; 8: IBUN 13A; 9: Lambda concatamers (NEB); 10: IBUN 18A; 11: IBUN 18S; 12: IBUN 62B; 13: IBUN 137K; 14: IBUN 158B; 15: IBUN 18Q; 16: IBUN 64A; 17: Lambda *Hind*III marker.Pulse times of 1 to 40 sec for 22 hrs.
(c) Lane 1: IBUN 125C; 2: *C. acetobutylicum* ATCC 824; 3: IBUN 22A; 4: IBUN 140B; 5: IBUN 62F; 6: IBUN 13A; 7: IBUN 18A; 8: Lambda concatamers (NEB); 9: IBUN 18S; 10: IBUN 62B; 11: IBUN 137K; 12: IBUN 158B; 13: IBUN 64A; 14: IBUN 965B; 15: IBUN 18Q; 16: IBUN 125C; 17: Lambda concatamers (Promega). Pulse time of 10 to 120 sec for 22 hrs.

Eppendorf tubes and washed 3 times with 100 µl TE buffer. They were incubated at room temperature for 15 min and then TE buffer was replaced by 50 µl 2X enzyme restriction buffer; the mixture was left at room temperature for 1 hr. 100 µl 1X enzyme digestion buffer and 10 U *Eco*RI (Promega) were then added. This was incubated overnight at 37ºC. They were then placed on agarose gel and conventional electrophoresis was performed (0.7% agarose gel, 4.2 V/cm, 0.5X TBE buffer, pH 8.0 [Tris-Borate and EDTA]) where smearing was observed, proving that the enzyme had digested the DNA. Conventional electrophoresis was then performed to verify that DNA had remained intact (0.8% agarose gel, 100 volts, 0.5X TBE buffer); this was done with DNA but without digestion and a single band was observed, confirming DNA quality.

Enzyme digestion

Our selection of appropriate rare-cutting endonucleases was based on low *Clostridium* G+C content (22% mol to 55% mol) (Andreesen et al. 1989). The following enzymes were selected: *Apa*I (Promega), *Sma*I (Promega), *Xho*I (Fermentas), *Sda*I (Fermentas) and *Eco*52I (Fermentas). Agarose plugs containing intact chromosomal DNA (50 µL) were washed with 200 µL TE buffer and incubated at 37ºC for 1 hr. This step was repeated twice more; however, incubation only took place for 15 min. A last wash was done at room temperature for 15 min. DNA plugs were equilibrated at room temperature for 15 min in 100 µL of the restriction enzyme buffer recommended by the manufacturer before pre-incubation at 4ºC for 6 hrs with 5 to 10 U restriction enzyme per plug and 100 µL of new buffer

Pulsed field gel electrophoresis (PFGE) conditions

DNA fragments were separated in 1% (wt/vol) agarose gels (Gibco BRL) at 200 V constant voltage in 0.5X TBE buffer at 10ºC by using contour clamped homogenous field electrophoresis (PFGE-CHEF system: Gene Navigator System TM, Pharmacia LKB, Uppsala, Sweden). Three different electrophoretic running conditions were tested for determining genome size (Hielm et al. 1998). Condition one specified 16 hrs run time with 1 to 8 second pulse time, condition two required 22 hrs and 1 to 40 sec pulses and condition three specified 22 hrs and 10 to 120 sec pulses. DNA fragments were visualised after staining with ethidium bromide (5 µg/ml) for 30 min using a UV transilluminator; excess ethidium bromide was removed by submerging the gels in 0.5X TBE buffer for 20 min. Condition two was chosen as most bands were seen and had good resolution in this restriction profile.

Genome size measurement

Electrophoresis employed condition two for estimating genome size for *Clostridium* native strains, using DNA digested with *Apa*I, *Eco*52I, *Sma*I and *Xho*I. Gel images were acquired with Gel Doc (BioRad); Quantity One software was employed for analysing the images and

estimating each fragment's size compared to the lambda concatemers (New England Biolabs). GeneTools software (HITACHI) was then used for confirming fragment sizes; similar results were obtained when comparing those obtained using Quantity One. Gels were run in triplicate. Standard deviations were determined for each fragment size and genome size was estimated by adding up the sizes of the restriction fragments. Comigrating fragments were resolved by staining intensity since ethidium bromide binds stoichiometrically to DNA. Densitometric tracings of photographs for each gel were used for determining the number of comigrating fragments contained in bands having unusual intensity (Bergthorsson and Ochman, 1995).

Detecting megaplasmids by PFGE

Megaplasmids were visualised when using the PFGE conditions described by Cornillot and Soucaille (1996) using both agarose inserts containing 3.0 µg/ml genomic DNA, and genomic DNA previously prepared with solvents as described by Jaimes et al. 2005. DNA from all strains were separated on 1% agarose gel (wt/vol), in a PFGE-CHEF system, at 200 V constant voltage for 12 hrs at 10ºC using 6 to 16 sec pulses and 0.5X TBE buffer. Gels were visualised as described before.

Evaluating the genes related to pSOL1 by PCR

Megaplasmids were purified from genomic DNA which had been previously prepared with solvents as described by Jaimes et al. (2005). 200 µL NaOH 0.2N were added to 500 µL of genomic DNA for 5 min in an ice bath. 150 µL potassium acetate solution (60% vol/vol potassium acetate, 11.5% vol/vol glacial acetic acid, pH5.5) were added to the mixture and completely mixed during 10 min. It was centrifuged at 14,500 rpm for 10 min and supernatant was precipitated with one volume of pre-chilled absolute ethanol. The supernatant was skimmed off and the pellet was dried at 37ºC for 1 hr. The pellet was then suspended in 50 µl TE buffer and stored at -20ºC. Megaplasmid quality was verified as described before.

DNA fragments were amplified by PCR according to Nair et al. 1994 and Cornillot et al. 1997a. One pair of primers was selected for amplifying the *adc* gene region: PADC-UP (5'-GAATTCATAAAAACACCTCCACATAAGT-3') and PADC-DN (5'-TTACTTAAGATAATCATATATAACTTCA-3'). The following modifications were made to their protocols. Both chromosomal and purified plasmidic DNA from *C. acetobutylicum* ATCC 824 and IBUN 18 A (25 ng/µL) were used as templates to establish in which of them the *adc* gene could be found. Chromosomal DNA from *Escherichia coli* DH5α was used as negative control due to low probability of annealing previously determined by primers alignment against complete genome sequences of that species (high E values).

The PCR reaction mixture contained template DNA (25 ng/µL), 200 µM DNTPs, 2.0 mM $MgCl_2$, 0.2 µM each one PADC-UP and PADC-DN primers, 1X PCR buffer and 1.25 units *Taq* polymerase (Promega) in a final 25 µL volume. Each of the 30 amplification cycles employed a denaturing step at 94ºC for 1 min followed by alignment at 40ºC for 1 min and an extension step at 72ºC for 1 min. Conventional electrophoresis was done in 2% agarose gel at 4.2 V/cm with 0.5X TBE buffer.

RESULTS

Preparing and purifying DNA, enzyme digestion and PFGE conditions

The three parameters employed for testing DNA quality were appropriate and it was possible to resolve the macrofragments in electrophoresis by fulfilling these criteria and following the standardised methodology for preparing and treating DNA.

Only four (*Apa*I, *Eco*52I, *Sma*I and *Xho*I) of the five enzymes tested for DNA cleavage and determining *Clostridium* native strain genome size produced a convenient number of fragments (between 10 and 17) (Figure 1) (Fonstein and Haselkorn, 1995). *Sda*I patterns

Table 2. Estimated genome size for *C. acetobutylicum* ATCC 824 and *Clostridium spp.* Colombian native strains.

Strain	*Apa*I	*Eco*52I	*Sma*I	*Xho*I	Mean genome size (Kbp)
ATCC 824	3,931	4,145	4,115	4,351	4,136 ± 169
IBUN 22A	3,901	4,090	4,368	4,415	4,194 ± 237
IBUN 125C	3,822	4,096	4,305	4,345	4,142 ± 235
IBUN 140B	3,918	3,999	4,321		4,079 ± 241
IBUN 62F	3,862	4,331	4,393	4,229	4,204 ± 233
IBUN 95B	4,071	3,948	4,313		4,111 ± 210
IBUN 13A	4,043	4,137	4,371	4,371	4,231 ± 163
IBUN 18A	4,043	3,855	4,345	4,398	4,160 ± 252
IBUN 18S	4,074	4,141	4,371		4,195 ± 176
IBUN 62B	3,920	4,103	4,342		4,122 ± 239
IBUN 137K	4,061	4,119	4,372	4,318	4,218 ± 148
IBUN 158B	4,091	4,085	4,097		4,091 ± 7
IBUN 18Q	3,791	4,337	4,345	4,182	4,164 ± 254
IBUN 64A	3,853	4,096	4,361	4,381	4,173 ± 245

were unsuitable for basic PFGE interpretation. Reproducibility of banding patterns between different batches of DNA was observed with all enzymes selected. Visual inspection indicated that some bands in the gel electrophoretograms were more intense than others. Two additional bands were observed in the *Apa*I pattern for *C. acetobutylicum* ATCC 824, compared to the results reported by Wilkinson and Young (1993).

DNA fragments from different size ranges were separated by varying the pulse time; sizes were estimated by comparing the fragments to appropriate molecular weight markers.

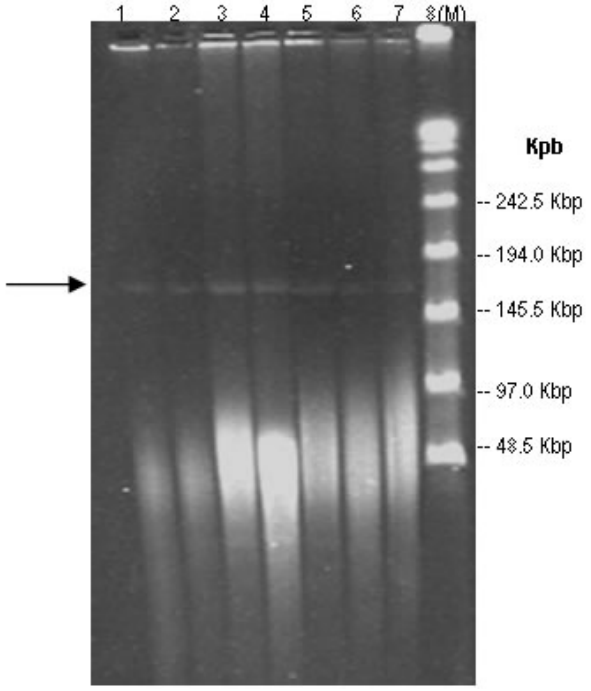

Figure 2. PFGE with total extracted DNA. Lane 1: *C. acetobutylicum* ATCC 824; 2: *C. acetobutylicum* ATCC 824; 3: IBUN 22A; 4: IBUN 158B; 5: IBUN 13A; 6: IBUN 18A; 7: IBUN 137K; 8: Lambda concatemers. The arrow shows: extrachromosomal DNA. Separated by PFGE, pulse times were ramped from 6 to 16 sec for 22 hrs at 200 volts.

Genome size measurement

Figure 1 shows the gels used for estimating genome size; these corresponded to DNA digested with the *Eco*52I enzyme in three different conditions. Most bands were better resolved in condition two. Table 1 presents the results obtained with *C. acetobutylicum* ATCC 824 DNA digested with *Apa*I, *Eco*52I, *Sma*I, and *Xho*I enzymes.

A band having greater intensity (212 Kbp) was detected in the *Apa*I pattern and doubled for determining all sizes. The two first bands were not taken into account as they were possibly related to partial digestion. These considerations were used for estimating *C. acetobutylicum* ATCC 824 genome size as being 3,931 Kbp, comparable to its real 3,940 Kbp size (Table 1). *C. acetobutylicum* ATCC 824 strain genome size was 4,115 Kbp when estimated with the *Sma*I enzyme and 4,351 Kbp with the *Xho*I enzyme. Table 2 summarises mean estimated genome size for each strain analysed with *Apa*I, *Eco*52I, *Sma*I and *Xho*I enzymes. Genome size for native strains was estimated as being 4,079 to 4,231 Mbp.

Detecting megaplasmids by PFGE and evaluating the pSOL1 plasmid

A band having little intensity (plus the chromosomal DNA band) was observed when employing the running conditions used by Cornillot and Soucaille in 1996 and Cornillot et al. in 1997a (using non-digested DNA inserts). Running times and pulses were varied for improving this fragment's visualisation. The observed band migrated above the lambda concatemer marker 145.5 Kbp fragment. This band was detected in the *C. acetobutylicum* ATCC 824 pattern strain and all native strains. Bands having greater intensity were visualised when solvent-purified DNA was run (Figure 2).

PCR amplification was employed to find out whether the extra-chromosomal DNA band observed in PFGE was related to the pSOL1 plasmid. A 900 to 1,000 bp band was amplified using the PADC-UP and PADC-DN pair of primers. This band was the expected size when using these primers (975 bp), suggesting that the *adc* gene encoding an *acetoacetate decarboxylase* could be localised in this fragment (Figure 3).

DISCUSSION

Genome size measurement and PFGE

PFGE is capable of resolving large fragments of DNA having a practical range of 10 Kb to ~7 Mbp by using rare-cutting enzymes (Warner and Onderdonk, 2003). Many efforts have been made in recent years to improve and standardise PFGE procedures for determining genome size in different bacteria. They generally consist of preparing DNA in agarose plugs, electrophoresis, image processing and data interpretation (Fiett et al. 2004). One of the main PFGE applications consists of calculating bacterial genome size by adding the estimated size of resolved restriction fragments, preferably cut by a variety of suitable rare-cutting restriction enzymes. PFGE was used for determining the genome size of thirteen native Clostridia strains in this work and establishing the presence of megaplasmids.

Three parameters were taken into account to standardise the method: establishing a DNA concentration allowing good band resolution in electrophoresis gels, controlling agarose concentration and using at least three proteinase K treatments. Some bands were not clearly distinguished at DNA concentrations of less than 0.8 µg/ml and visualising the bands became hampered by the excessive quantity of

DNA at concentrations greater than 3.5 µg/ml. 1.2% agarose concentration for preparing DNA inserts showed better band resolution than 1.4% and 1.5% concentrations. This fact was probably related to the formation of larger pores increasing the diffusion of enzymes, reagents and detergents which led to obtaining high purity DNA. It was also observed in this study that three proteinase K treatments led to agarose inserts becoming more translucent (suggesting efficient protein degradation by enzymes and thereby leading to obtaining DNA which had no remaining proteins), whilst one or two treatments were not enough to obtain suitable DNA (Wilkinson and Young, 1993; Lin and Johnson, 1995; Hielm et al. 1998).

It is difficult to resolve all fragments resulting from DNA digestion when using just one set of PFGE conditions. Three different electrophoretic ramps were therefore tested for determining genome size as described by Hielm et al. (1998). This allowed comparing conditions aimed at selecting just the one where most bands were located in the highest gel resolution area. Size should be most accurately estimated when fragments migrate within this region of the gel (Birren and Lai, 1993). This was especially observed in electrophoresis gels having DNA digested with *Eco*52I in condition one where a compression area was observed showing several bands; this became resolved in gels run in conditions two and three.

PFGE is one of the most direct and accurate ways of determining genome size compared to other methods (Lin and Johnson, 1995). Studies using either PFGE or whole-genome sequencing have revealed a diversity of bacterial genome sizes, ranging from as low as 0.5 Mbp to as high as 10 Mbp (Bansal and Meyer, 2002). However, clostridial genome sizes determined by PFGE have ranged from 3.5 Mbp to 6.5 Mbp (Wilkinson and Young, 1993; Keis et al. 2001; Lin and Johnson, 1995; Hielm et al. 1998) although sequenced genomes range from 2.79 to 3.94 Mbp (Nölling et al. 2001; Shimizu et al. 2002; Brüggemann et al. 2003). Mean genome size determined for the *C. acetobutylicum* ATCC 824 reference strain was 4,136 ± 169 Kbp in our work, which is comparable to the size estimated for the same strain by Wilkinson and Young (1993) and Cornillot et al. (1997b). Genome sizes for clostridia native strains ranged from 4.0 to 4.2 Mbp (Table 2).

ATCC 824 genome size calculated with the *Apa*I restriction profile was initially 5,087 Kbp; however, differences between this value and data obtained by sequencing were associated with comigrating fragments and partial digests producing larger fragments. Wilkinson and Young (1993) have shown that the first band from *C. acetobutylicum* ATCC 824 digested with *Apa*I migrates at a distance similar to that of the Lambda concatemers marker 485 Kbp fragment. Two additional bands migrating above the same marker band were observed in this work which reinforces the hypothesis of possible partial digestion increasing genome size measurement. These first two bands were therefore not taken into account when determining final genome size. A band having greater intensity (212 Kbp) was detected in *Apa*I profiles and summed twice for determining size, bearing in mind that it might have corresponded to two comigrating fragments, this being in agreement with considerations reported by Wilkinson and Young (1993), Bergthorsson and Ochman (1995), Lima and Correia (2000) and Warner and Onderdonk (2003). The estimated genome size was therefore close to that obtained by sequencing.

A lesser value for the *C. acetobutylicum* ATCC 824 reference strain was obtained (2.5 Mbp) during genome size determination with *Eco52*I; such reduction in size could also be related to comigrating fragments. Cornillot et al. (1997b) previously determined *C. acetobutylicum* ATCC 824 genome size with the *Eco*52I and revealed the presence of 10 bands which had the same quantity of fragments as obtained in our work. It was thus decided to estimate genome size taking the pattern strain as molecular size marker. ATCC 824 genome size was around 4.1 Mbp when using this strategy and native strain genome size values ranged from 3.8 to 4.3 Mbp (Table 1 and Table 2).

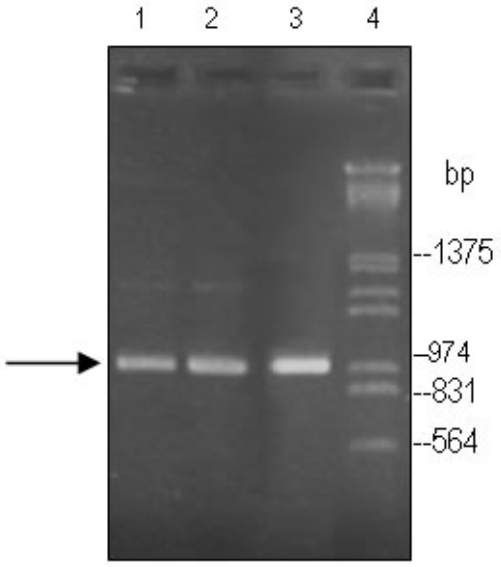

Figure 3. pSOL1 975 bp amplified fragment. Lane 1, *C. acetobutylicum* ATCC 824–25 ng; 2, *C. acetobutylicum* ATCC 824–90 ng; 3, IBUN 18A; 4, Lambda with *Hind*III and *Eco*RI. PCR amplification for 30 cycles, denaturing at 94ºC for 1 min, aligning at 40ºC for 1 min and extension at 72ºC for 1 min using PADC-UP and PADC-DN primers.

Detecting megaplasmids by PFGE and evaluating the pSOL1 plasmid

Another of this study's aim was using PFGE-CHEF to detect megaplasmids in native strains, since conventional methods for extracting large plasmids are tedious and time-consuming (Buchrieser et al. 1994). Megaplasmids are likely to be much more common amongst bacteria than is currently believed. This prediction is based upon their

sometimes poor recovery in plasmid DNA extracts obtained by many popular methods and the difficulties of observing them in standard electrophoretic gels (Barton et al. 1995). Amils et al. (1998) have reported that some extra-chromosomal elements can be resolved in PFGE conditions, although shorter pulse time runs are required for better characterisation of these elements. Only one extra-chromosomal element was detected in our work as migrating between 145.5 kb and 194 kb marker bands in all native strains Figure 2. Shorter pulse time runs and total DNA purified with CTAB (Jaimes et al. 2005) were therefore used for visualising megaplasmids; it should be noted that poor intensity bands were observed when DNA was directly obtained from agarose inserts (Cornillot and Soucaille, 1996).

The *adc* gene region was successfully amplified by PCR using plasmidic DNA from *C. acetobutylicum* ATCC 824 and native strain IBUN 18 A as template. The *adc* gene encodes acetoacetate decarboxylase which plays a fundamental role in solventogenic Clostridia's metabolic route. This fact suggests that genes related to solventogenesis in native strain IBUN18A may be localised in the extra-chromosomal element detected in this work, as observed by Cornillot et al. (1997a) in *C. acetobutylicum* ATCC 824. It is presumed that genes encoding enzymes implied in the solventogenic metabolic route could be localised in other native strains' megaplasmids, unlike *C. beijerinckii* which contains widespread solventogenic genes in its genome (Wilkinson and Young, 1998).

CONCLUDING REMARKS

The PFGE method developed in this study was shown to be useful for obtaining estimated genome sizes for thirteen *Clostridium* strains isolated from Colombian soils. Further analysis is required for detecting solventogenic genes in native strains. Determining genome size and the physical structure of bacterial genomes as revealed by endonuclease mapping could provide an alternative parameter to be employed in taxonomic studies (Liu et al. 1999). Results obtained in this work are useful for increasing genetic information about native Colombian strains.

ACKNOWLEDGMENTS

This work was carried out within the framework of the project entitled, “Anaerobic solvent producing bacteria: biomass fermentation and isolation of cellulolytic strains” financed by COLCIENCIAS (the Instituto Colombiano Francisco José de Caldas), Banco de la República and Universidad Nacional de Colombia. The authors would like to express their thanks to Jason Garry for patiently reading the manuscript.

REFERENCES

AMILS, Ricardo; IRAZABAL, Nagore; MOREIRA, David; ABAD, Jose and MARÍN, Irma. Genomic organization analysis of acidophilic chemolithotrophic bacteria using pulsed field gel electrophoretic techniques. *Biochimie*, November 1998, vol. 80, no. 11, p. 911-921.

ANDREESEN, Jan; BAHL, Hubert and GOTTSCHALK, Gerhard. Introduction to the Physiology and Biochemistry of the Genus *Clostridium*. In: MINTON, Nigel P. and CLARKE, David J. eds. *Clostridia (Biotechnology Handbooks)*. New York, Plenum Publishing Corporation, 1989, p. 27-53.

ARÉVALO, Catalina; AGUILERA, Greicy; ARRIETA, Adriana; ARISITIZÁBAL, Fabio and MONTOYA, Dolly. Caracterización de cepas nativas colombianas de clostridios solventogénicos por perfiles de plásmidos. *Revista Colombiana de Ciencias Químico-Farmaceútica,* 2002, p. 51-60.

BARTON, Bret M.; HARDING, Gordon P. and ZUCCARELLI, Anthony J. A general method for detecting and sizing large plasmids. *Analytical Biochemistry*, April 1995, vol. 226, no. 2, p. 235-240.

BANSAL, Arvind K. and MEYER, Terrance E. Evolutionary analysis by whole-genome comparisons. *Journal of Bacteriology,* April 2002, vol. 184, no. 8, p. 2260–2272.

BERGTHORSSON, Ulfar and OCHMAN, Howard. Heterogeneity of genome sizes among natural isolates of *Escherichia coli. Journal of Bacteriology,* October 1995, vol. 177, no. 20, p. 5784-5789.

BIRREN, Bruce and LAI, Eric. *Pulsed field gel electrophoresis*: *A practical guide*. London, Academic Press Limited, 1993. 272 p. ISBN 0-12-101290-5.

BRÜGGEMANN, Holger; BÄUMER, Sebastian; FRICKE, Wolfgang Florian; WIEZER, Arnim; LIESEGANG, Heiko; DECKER, Iwona; HERZBERG, Christina; MARTINEZ-ARIAS, Rosa; MERKL, Rainer; HENNE, Anke and GOTTSCHALK, Gerhard. The genome sequence of *Clostridium tetani*, the causative agent of tetanus disease. *Proceedings of the National Academy of Sciences of the United States of America,* February 2003, vol. 100, no. 3, p. 1316-1321.

BUCHRIESER, C.; WEAGANT, S. and KASPAR, C.W. Molecular characterization of *Yersinia enterocolitica* by pulsed-field gel electrophoresis and hybridization of DNA fragments to ail and pYV probes. *Applied and Environmental Microbiology*, December 1994, vol. 60, no. 12, p. 4371-4379.

CORNILLOT, Emmanuel and SOUCAILLE, Philippe. Solvent-forming genes in Clostridia. *Nature,* April 1996, vol. 380, no. 6574, p. 489.

CORNILLOT, Emmanuel; NAIR, Ramesh V.;

PAPOUTSAKIS, Eleftherios T. and SOUCAILLE, Philippe. The genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 reside on a large plasmid whose loss leads to degeneration of the strain. *Journal of Bacteriology*, September 1997a, vol. 179, no. 17, p. 5442–5447.

CORNILLOT, Emmanuel; CROUX, Christian and SOUCAILLE, Philippe. Physical and genetic map of the *Clostridium acetobutylicum* ATCC 824 chromosome. *Journal of Bacteriology*, December 1997b, vol. 179, no. 23, p. 7426-7434.

FIETT, Janusz; KUCHARCZYK, Krzysztof and GNIADKOWSKY, Marek. A new method of DNA preparation for pulsed-field gel electrophoresis analysis. *Journal of Microbiological Methods,* December 2004, vol. 59, no. 3, p. 433-436.

FONSTEIN, Michael and HASELKORN, Robert. Physical mapping of bacterial genomes. *Journal of Bacteriology*, June 1995, vol. 177, no. 12, p. 3361-3369.

HIELM, Sebastian; BJÖRKROTH, Johanna; HYYTIÄ, Eija and KORKEALA, Hannu. Genomic analysis of *Clostridium botulinum* group II by pulsed-field gel electrophoresis. *Applied and Environmental Microbiology,* February 1998, vol. 64, no. 2, p. 703-708.

JAIMES, Claudia P.; SUÁREZ, Zulma R.; ARISTIZÁBAL, Fabio A.; BERNAL, J. Mauricio and MONTOYA, Dolly. AFLP fingerprinting of Colombian *Clostridium* spp. strains, multivariate data analysis and its taxonomical implications. In: *157th Meeting of the Society for General Microbiology* [online]. 12-14 September 2005, Keele, UK. Abstracts, p. 37. Available from Internet: http://www.sgm.org.uk/meetings/pdfabstracts/keele2005abs.pdf.

KEIS, Stefanie; SULLIVAN, John T. and JONES, David T. Physical and genetic map of the *Clostridium saccharobutylicum* (formerly *Clostridium acetobutylicum*) NCP 262 chromosome. *Microbiology*, July 2001, vol. 147, no. 7, p. 1909-1922.

LIMA, Paula T. and CORREIA, António M. Genetic fingerprinting of *Brevibacterium linens* by pulsed-field gel electrophoresis and ribotyping. *Current Microbiology*, July 2000, vol. 41, no. 1, p. 50-55.

LIN, Wei-Jen and JOHNSON, Eric A. Genome analysis of *Clostridium botulinum* type a by pulsed-field gel electrophoresis. *Applied and Environmental Microbiology,* December 1995, vol. 61, no. 12, p. 4441-4447.

LIU, Shu-Lin; SCHRYVERS, Anthony; SANDERSON, Kenneth E. and JOHNSTON, Randal N. Bacterial phylogenetic clusters revealed by genome structure. *Journal of Bacteriology*, November 1999, vol. 181, no. 21, p. 6747-6755.

MONTOYA, Dolly; PERDOMO, Liliana; ARÉVALO, Catalina; ARISTIZABAL, Fabio and SCHWARZ, Wolfgang. Caracterización de cepas nativas de *Clostridium* spp. por secuenciación parcial del gen ribosomal *16s rARN*. *Revista Colombiana de Biotecnología*, December 1999, vol. 2, no. 1, p. 35-39.

MONTOYA, Dolly; SPITIA, Sandra; SILVA, Edelberto and SCHWARZ, Wolfgang H. Isolation of mesophilic solvent-producing clostridia from Colombian sources: physiological characterisation, solvent production and polysaccharide hydrolysis. *Journal of Biotechnology*, April 2000, vol. 79, no. 2, p. 117-126.

MONTOYA, Dolly; ARÉVALO, Catalina; GONZALES, Sonny; ARISTIZABAL, Fabio and SCHWARZ, Wolfgang H. New solvent-producing *Clostridium* sp. strains, hydrolysing a wide range of polysaccharides, are closely related to *Clostridium butyricum*. *Journal of Industrial Microbiology and Biotechnology,* November 2001, vol. 27, no. 5, p. 329-335.

NAIR, Ramesh V.; BENNETT, George N. and PAPOUTSAKIS, Eleftherios T. Molecular characterization of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824. *Journal of Bacteriology*, February 1994, vol. 176, no. 3, p. 871-885.

NÖLLING, Jörk; BRETON, Gary; OMELCHENKO, Marina V.; MAKAROVA, Kira S.; ZENG, Qiandong; GIBSON, Rene; LEE, Hong Mei; DUBOIS, JoAnn; QIU, Dayong; HITTI, Joseph; G.T.C. Sequencing Center Production, Finishing, and Bioinformatics Teams; WOLF, Yuri I.; TATUSOV, Roman L.; SABATHE, Fabrice; DOUCETTE-STAMM, Lynn; SOUCAILLE, Philippe; DALY, Michael J.; BENNETT, George N.; KOONIN, Eugene V. and SMITH, Douglas R. Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *Journal of Bacteriology*, August 2001, vol. 183, no. 16, p. 4823-4838.

SHIMIZU, Tohru; OHTANI, Kaori; HIRAKAWA, Hideki; OHSHIMA, Kenshiro; YAMASHITA, Atsushi; SHIBA, Tadayoshi; OGASAWARA, Naotake; HATTORI, Masahira; KUHARA, Satoru and HAYASHI, Hideo. Complete genome sequence of *Clostridium perfringens*, an anaerobic flesh-eater. *Proceedings of the National Academy of Sciences of the United States of America*, January 2002, vol. 99, no. 2, p. 996-1001.

WARNER, John E. and ONDERDONK, Andrew B. Method for optimising pulsed-field gel electrophoresis banding pattern data. *The Journal of Molecular Diagnostics*, February 2003, vol. 5, no. 1, p. 21-27.

WILKINSON, Shane R. and YOUNG, Michael. Wide diversity of genome size among different strains of

Clostridium acetobutylicum. *Journal of General Microbiology*, 1993, vol. 139, no. 5, p. 1069-1076.

WILKINSON, Shane R. and YOUNG, Michael. *Clostridium beijerinckii* NCIMB 8052 Chromosome. In: BRUJIN, Frans J.; LUPSKI, James R. and WEINSTOCK, George M. eds. *Bacterial Genomes: Physical Structure and Analysis*. Massachusetts, Kluwer Academic Publishers, 1998, vol. 1. p. 641-644.