Eur Biophys J (2000) 29: 48–56

BIOPHYSICS LETTER

A.M. Riveron · L. Lopez-Canovas · M. Baez-Camargo
E. Flores · G. Perez-Perez · J.P. Luna-Arias · E. Orozco

Circular and linear DNA molecules in the *Entamoeba histolytica* complex molecular karyotype

Received: 19 January 1999 / Revised version: 3 November 1999 / Accepted: 22 November 1999

Abstract *Entamoeba histolytica* genome was analysed by pulsed field gel electrophoresis under conditions to separate linear chromosomes in the 170–1400 kb range. We identified linear DNA molecules of 227, 366, 631, 850, 1112 and 1361 kb (mean sizes obtained by three different methods) and we estimated their reorientation times and migration velocities at various experimental conditions. DNA shift mobility assays, using ethidium bromide, suggested that bands migrating at 227 and 631 kb contain linear and circular DNA, whereas a band at 436 kb has only circular DNA. We obtained a regression equation relating sizes of supercoiled DNA molecules with their migration velocities during a pulse at constant electric field and temperature. We also developed a computer program (EHPATTERNS) that predicts the migration per pulse and the resolution order of circular and linear *E. histolytica* DNA at different pulse times and constant driving and frictional forces. The simulation showed that linear DNA molecules frequently co-migrate with circular molecules, but circular molecules change when the pulse time varies. This molecular mixture generates broad bands and difficulties in the interpretation of the molecular karyotype of *E. histolytica*.

A.M. Riveron · L. Lopez-Canovas · G. Perez-Perez
Department of Molecular Biology, Neuroscience Branch,
National Center for Scientific Research,
P.O. Box 6990, Havana, Cuba

A.M. Riveron · J.P. Luna-Arias · E. Orozco
Molecular Biomedicine Program, CICATA-IPN,
Legaria 694, Col. Irrigación,
Mexico 11500 D.F.

M. Baez-Camargo · E. Flores · E. Orozco (✉)
Department of Experimental Pathology,
CINVESTAV-IPN, A.P. 14-740,
Mexico 07300 D.F.
e-mail: esther@mail.cinvestav.mx

J.P. Luna-Arias
Multidisciplinary Program in Molecular Biomedicine,
CINVESTAV-IPN, A.P. 14-740,
Mexico 07300 D.F.

Key words *Entamoeba histolytica* · Linear and circular DNA · Migration velocities · Reorientation time · Pulsed field

Abbreviations *ARS* autonomous replication sequence · *CHEF* contour clamped homogeneous electric field equipment · *cirDNA* circular DNA · *EtBr* ethidium bromide · *linDNA* linear DNA · *miniCHEF* miniequipment of contour clamped homogeneous electric field · *scDNA* supercoiled DNA · *TAFE* transverse alternating field electrophoresis equipment · *TelTh Tetrahymena* telomeric probe

Symbols

D	DNA migration distance
d	DNA migration per pulse
d_m	DNA migration per pulse after reorientation
d_r	DNA migration per pulse during reorientation
$d_{scDNAEh}$	migrations per pulses of *E. histolytica* scDNA molecules
E	electric field
L	DNA contour length
N_p	number of pulses
Q	formal DNA net charge
t_e	electrophoresis time
t_p	pulse time
t_r	DNA reorientation time
v_m	migration velocity after DNA reorientation
v_r	migration velocity during DNA reorientation
V_{scDNA}	migration velocity of *E. histolytica* scDNA molecules
η	buffer viscosity

Introduction

Entamoeba histolytica, the causative protozoan of human amoebiasis, has a complex genome composed of:

(1) at least 6–8 chromosomes of undetermined sizes (Argüello et al. 1992); (2) circular DNA (cirDNA) molecules ranging from 5 to 50 kb (Dhar et al. 1995; Lioutas et al. 1995; Baez-Camargo et al. 1996a); (3) extranuclear DNA in the EhkO organelle (Orozco et al. 1997). Pulsed field gel electrophoresis (PFGE) karyotypes present 8–16 bands ranging from 170 to 3000 kb (Valdes et al. 1990; Orozco et al. 1993). Bands of 170, 250, 400, 800, 1200 and 1600 kb hybridise with a telomeric probe, and a huge amount of ribosomal concatamers are resolved at pulse times larger than 120 s, interfering with the visualisation of other molecules (Baez-Camargo et al. 1996b).

Migration per pulse of linear DNA (linDNA) is a function of the pulse time (t_p), DNA reorientation time (t_r) and migration velocities during (v_r) and after (v_m) DNA reorientation (Riveron et al. 1989, 1994). In addition, t_r, v_r and v_m depend on the experimental conditions. In contrast, cirDNA migration depends on topology and size. However, running conditions affect cirDNA and linDNA migration in a different way (Hightower et al. 1989). Previously, we described migrations per pulses of linDNA molecules at constant driving and frictional forces in contour clamped homogeneous electric field equipment (CHEF), and proposed the equation that describes linDNA migration at different experimental conditions (Riveron et al. 1989; Lopez-Canovas et al. 1998a). Here, using these equations, we identified *E. histolytica* linDNA molecules and calculated their sizes, t_r, v_r and v_m. We also proposed an equation to describe cirDNA migration per pulse and developed a computer program (EHPATTERNS[1]) that predicts the order of resolution of cirDNA and linDNA molecules.

Experimental

TAFE, CHEF and miniCHEF experiments

E. histolytica DNA molecules were separated in CHEF and miniCHEF (Riveron et al. 1995) and TAFE (transverse alternating electric field) Geneline I (Beckman) experiments under conditions that resolve 170–1400 kb linDNA, but retain the ribosomal concatamers in the gel compression zone. We used 1.5% agarose gels (Lachema or GTG Gold Seakem agarose FMC) and 0.5X TBE buffer (1X TBE: 89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.3). An automatic controlled switching unit set the pulse time (t_p) and the number of pulses (N_p). A MultiTemp LKB heat exchanger buffer maintained constant temperature. Plugs for CHEF, TAFE (0.2 cm) and miniCHEF (0.1 cm) experiments were prepared as reported (Baez-Camargo et al. 1996a). *Saccharomyces cerevisiae* (strain 196-2, *Matα his*$^-$) was a kind gift of M. Luzzati, France. Gels were stained with ethidium bromide (EtBr) (0.5 μg/ml), blotted to nylon membanes and hybridised with [α-^{32}P]dATP random primed labelled *Tetrahymena* telomeric (*TelTh*)) and *E. histolytica* ribosomal [16S, 25S and autonomous replication sequence (ARS)] probes (Burke et al. 1987; Grodberg et al. 1990; Que and Reed 1991; Michel et al. 1995). The *TelTh* probe was hybridised at 37 °C in 2X SSC and homologous probes at 42 °C in 0.1X SSC (2X SSC: 0.3 M NaCl, 0.03 M sodium citrate, pH 7.5). Washings were carried out at 45 °C in 1X SSC and 0.5% SDS.

DNA migration distance and migration per pulse

Gels and autoradiographies were analysed by the Gel Doc 1000 system (BioRad) and by software of public domain. The DNA migration distance D (cm) was measured from the migration origin to the midpoint of each band and the DNA migration per pulse d was calculated as D/N_p [$N_p = t_e/2t_p$; t_e (s) is the electrophoresis time]. Each measurement was carried out at least three times with a mean error of 0.001 cm. Migrations per pulses of supercoiled DNA molecules (scDNA) were calculated using mobilities (μ in cm^2/V s) reported by Sobral and Atherley (1989). D was obtained multiplying μ by the electric field (E) and the t_e value.

Determination of *E. histolytica* DNA sizes by co-migration with yeast chromosomes (method I)

E. histolytica DNA molecules and *S. cerevisiae* chromosomes were separated in TAFE experiments at 5.8 V/cm, 20 °C, 120 s of t_p and 360 pulses. Sizes of yeast chromosomes (Goffeau et al. 1996) and migrated distances (D_{yeast}) were related by $D_{yeast} = c_0 + c_1 kb_{yeast}$ where c_0 and c_1 are the regression coefficients. To fit the equation we obtained by simulation two replicas of each D_{yeast}. These values were normally distributed with a standard error of $0.07E$, where $E \approx N(0,1)$. *E. histolytica* DNA sizes were determined by inverse interpolation of D_{Eh} in this equation. By applying Fieller's theorem (Fleiss 1986), we constructed the 95% confidence interval of sizes estimates.

Sizes, t_r and v_m of *E. histolytica* DNA, studying relationships between d and t_p (method II)

Migrations per pulses of DNA molecules were determined in CHEF experiments at 10, 25, 50, 80, 100, 120, 160 and 180 s, 1.5% agarose gels, 0.5X TBE, 5.8 V/cm and 20 °C. Migrations per pulses (d_m) obtained at $t_p > t_r$ were analysed by:

$$d_m = v_m(t_p - t_r) = -v_m t_r + v_m t_p \quad (1)$$

CHEF-separated DNA bands were indexed by their order of appearance (from the bottom) as Eh_1, Eh_2, ...,

[1] EHPATTERNS software is available to those interested in using it.

Eh_p. Subscripts were identified by the variable k and the number of molecules was p. The t_p were t_{p_1}, t_{p_2}, ..., t_{p_m}. Variable i was used for these subscripts and m for the number of assays. Migration distances (D_{ki}) of Eh_k molecules were measured at $t_{p_{ki}}$ and migrations per pulse (d_{ki}) were calculated as $D_{ki}/N_{p_{kj}}$. For each Eh_k, $d_{m_{ki}}$ values were fitted to a linear function of $t_{p_{ki}}$:

$$d_{m_{ki}} = (b_0)_k + (b_1)_k t_{p_{ki}} \quad (2)$$

for a given k, E and T (°C), with i ranging from 1 to m, where $(b_0)_k$ and $(b_1)_k$ are the least square regression coefficients. If $d_{m_{ki}}$ and $t_{p_{ki}}$ obey Eq. (1) we will obtain regression coefficients differing significantly from zero and $(b_0)_k < 0$. By analogy between Eqs. (1) and (2), $t_{r_k} = -(b_0)_k/(b_1)_k$ and $v_{m_k} = (b_1)_k$. *E. histolytica* DNA sizes were estimated by replacing t_{r_k}, E and η [buffer viscosity that depends on temperature (T)] in Eq. (7) below, that relates t_r and DNA size.

Sizes, t_r, v_r and v_m using the equation describing migration per pulse of linDNA at different conditions (method III)

Migrations per pulses of linDNA molecules were described as a function of running variables (E, η, t_p) and DNA size (Lopez-Canovas et al. 1998a). The $d_{m_{kji}}$ of an Eh_k molecule after the reorientation (at $t_{p_{ji}} > t_r$) is

$$d_{m_{kji}} = 0.665 Q_k E_j^{1.76} [t_{p_{ji}} - 0.134 (L_k^{1.14}/v_{r_{kj}})^{0.926}]/[8\pi\eta_j L_k^{1.08}] \quad (3)$$

where L_k (cm) is the DNA contour length [0.34 nm × base pairs number (bp)], Q_k the formal DNA net charge ($2 \times bp \times 1.0 \times 4.806 \times 10^{-10}$ in statcoulomb), E_j is the electric field (statvolt/cm) and η_j is buffer viscosity. Subscript j indexes the running conditions (E_j, T_j) and i the pulse times (t_{p_i}) that resolve a particular DNA molecule (Eh_k). η was calculated in poise by replacing the temperature T (°C) in the fourth degree polynomial function that relates both variables (Lopez-Canovas et al. 1998a):

$$(\eta)_{H_2O} = 1.7844 \times 10^{-2} - 5.9388 \times 10^{-4} T + 1.3494 \times 10^{-5} T^2 - 1.9278 \times 10^{-7} T^3 + 1.2455 \times 10^{-9} T^4 \quad (4)$$

The $v_{r_{kj}}$ of an Eh_k molecule is

$$v_{r_{kj}} = 0.0207 Q_k E_j^{1.45}/(8\pi\eta_j L_k^{1.35}) \quad (5)$$

LinDNA migrations per pulses $d_{m_{kj}}$ were calculated from experiments at $t_p > t_r$ (not used in method II) performed in CHEF for various t_p and N_p at 5.8 V/cm, 15 and 20 °C, or in miniCHEF at 10.71 V/cm and 10 °C. The L_k corresponding to each $d_{m_{kji}}$ at E_j, η_j and $t_{p_{ji}}$ was calculated by Eq. (3) (Lopez-Canovas et al. 1998a). We used an iterative procedure that starts with a small initial L and stops when L reaches a value of a theoretical migration per pulse $d_{mt_{kji}}$ equal to the experimental $d_{m_{kji}}$. The $v_{r_{kj}}$, $v_{m_{kj}}$ (cm/s) and $t_{r_{kj}}$ (s) were obtained replacing L_k in Eqs. (5)–(7):

$$v_{m_{kj}} = 0.665 Q_k E_j^{1.76}/(8\pi\eta_j L_k^{1.08}) \quad (6)$$

$$t_{r_{kj}} = 0.134 [L_k^{1.14}/v_{r_{kj}}]^{0.926} \quad (7)$$

Comparisons of estimates were done using t-statistic (Lopez-Canovas et al. 1998b).

DNA mobility at different EtBr concentrations

E. histolytica DNA was run in agarose gel frames (1.5%) that were cast using a multilane comb (Serwer 1980), in which each lane was supplemented with 0, 1, 2.5 or 5 ng/ml EtBr (Baez-Camargo et al. 1996a). Samples were predialysed against each EtBr concentration and co-electrophoresed in miniCHEF for 4 h at 20 °C, 25 s t_p and 0.5X TBE buffer.

Migration per pulse of cirDNA as a function of t_p

Sobral and Atherley (1989) studied in CHEF the migration of 5.01, 6.03, 7.05, 8.07, 10.10, 12.14, 14.17 and 16.14 kb scDNA molecules. They used t_p from 1 to 25 s, 6.4 V/cm, 13 °C, 1% agarose gel and 0.5X TBE buffer. Using their data we studied if Eq. (1) explained migrations per pulses of scDNA molecules at different t_p. The scDNA were named $scDNA_1$, $scDNA_2$, ..., $scDNA_n$ and indexed by subscript w (ranging between 1 and the number of scDNA molecules n). The t_p were t_{p_1}, t_{p_2}, ..., t_{p_m} and indexed by subscript a as t_{p_a} (a ranging between 1 and number of assays m). Linear dependence between d (named d_{scDNA}) and t_p was studied by

$$d_{scDNA_{wa}} = g_{0scDNA_w} + V_{scDNA_w}(t_{p_a}) \quad (8)$$

where g_{0scDNA_w} and V_{scDNA_w} are the least square regression coefficients (intercept and slope, respectively). V_{scDNA_w} is the migration velocity during a single pulse for each $scDNA_w$. Equation (8) was fitted for each $scDNA_w$ using $d_{scDNA_{wa}}$ and t_{p_a}. The fit was done for all $scDNA_w$, and we examined the relation between V_{scDNA_w} and sizes (kb_w) by the function

$$V_{scDNA_w} = f(kb_w^x) \quad (9)$$

Theoretical approach to analyse *E. histolytica* complex karyotype formed by cirDNA and linDNA molecules; prediction of band patterns

Using Eqs. (1) and (3)–(7) we developed the computer program EHPATTERNS to predict $t_{r_{kj}}$, $v_{r_{kj}}$ and $v_{m_{kj}}$ at 6.4 V/cm and 13 °C for the Eh_k linDNA molecules identified here. For various t_{p_i} assayed in CHEF, the program estimated migrations per pulses of linDNA during ($d_{r_{kji}}$) and after their reorientation ($d_{m_{kji}}$) using the

equations: $d_{r_{kji}} = v_{r_{kj}}(t_{p_i})$ if $t_{p_i} \leq t_{r_{kj}}$, and $d_{m_{kji}} = v_{m_{kj}}(t_{p_i} - t_{r_{kj}})$ if $t_{p_i} > t_{r_{kj}}$ (Lopez-Canovas et al. 1998a). The migrations per pulses ($d_{scDNAEh_{wi}}$) of *E. histolytica* scDNA molecules of 5, 13, 26, 38 and 50 kb ($scDNAEh_1$ to $scDNAEh_5$) were predicted at the pulse times (t_{p_i}) used, considering that $d_{scDNAEh_{wi}} = v_{scDNAEh_{wi}}(t_{p_i})$ For each t_{p_i}, the program arrayed the migrations per pulses ($d_{scDNAEh_{wi}}$, $d_{r_{kji}}$ or $d_{m_{kji}}$) calculated for all $scDNAEh_{wi}$ and Eh_k linDNA from the smallest value. The results predicted the order of appearance of *E. histolytica* molecules in CHEF patterns. Data were processed using algorithms and programs previously described (Lopez-Canovas 1998a). EHPATTERNS was written in TURBO-PASCAL v. 7.0 (MS-DOS operating system).

Results and discussion

CHEF and TAFE *E. histolytica* DNA patterns and determination of sizes by co-migration with yeast chromosomes (method I)

E. histolytica DNA and *S. cerevisiae* chromosomes were co-electrophoresed at different CHEF, miniCHEF and TAFE conditions. At 5.8 V/cm, 20 °C and 80 s t_p, CHEF resolved a fast migrating band, which contains cirDNA, two bands close to 270 kb marker, one to 440 kb, one below and another above 577 kb and the compression zone (Fig. 1A). At 120 s t_p appeared the fast migrating band, diffuse bands near to 270, 440 and 813 kb markers, one above and another below 1091 kb and the compression zone (Fig. 1B). MiniCHEF (15 °C, 200 s t_p), gave wide bands (Fig. 1C): the fast migrating band, another at 813, one above 1091 kb, the compression zone and near to 440 kb a diffuse area containing stacked molecules (Fig. 1C). CHEF and TAFE at 120 s t_p, 5.8 V/cm and 20 °C gave 6–7 bands (Fig. 1B, D). By TAFE (Fig. 1D), D_{Eh} of *E. histolytica* DNA bands were 0.75, 1.60, 2.40, 2.75, 3.25 and 3.50 cm. We inversely interpolated D_{Eh} in $D_{yeast} = c_0 + c_1 kb$ ($c_0 = 4.3$, $S^2c_0 = 2.23 \times 10^{-3}$ and $c_1 = -3.0928 \times 10^{-3}$, $S^2c_1 = 1.28 \times 10^{-8}$) and obtained 1149 ± 71, 874 ± 53, 615 ± 39, 502 ± 36, 341 ± 36 and 259 ± 37 kb (estimated size ±95% confidence interval).

In GTG Gold Seakem agarose gels, patterns were better defined. These gels were blotted and hybridised with the *TelTh* and ribosomal probes (Fig. 2). *TelTh* revealed two wide bands, one around 160 kb and another between 1400 and 1600 kb. It also detected bands at 211, 376, 630, 843 and 1080 kb (Fig. 2C). Except for the band at 211 kb, the others fall into the confidence interval of the above estimates. The ribosomal probes hybridised at 200, 1200 and 1400 kb (Fig. 2B), suggesting that these bands contain cirDNA. Co-migration of cirDNA and linDNA causes the *E. hystolytica* DNA bands to appear diffuse, whereas *S. cerevisiae* chromosomes gave sharp bands (Fig. 1).

Migration of *E. histolytica* DNA molecules at different pulse times in CHEF; t_r, v_m and DNA sizes (method II)

In CHEF, at constant *E*, *T*, agarose and buffer concentrations, migration per pulse (d) of linDNA depends on the pulse time. Two straight lines describe d as function of t_p, one during DNA reorientation ($t_p \leq t_r$) and other after reorientation ($t_p > t_r$). The two slopes of these plots correspond to the velocities during (v_r) and after (v_m) reorientation. Reorientation time (t_r) is the t_p making $d_m = 0$ in Eqs. (1) and (2) (Riveron et al. 1989,

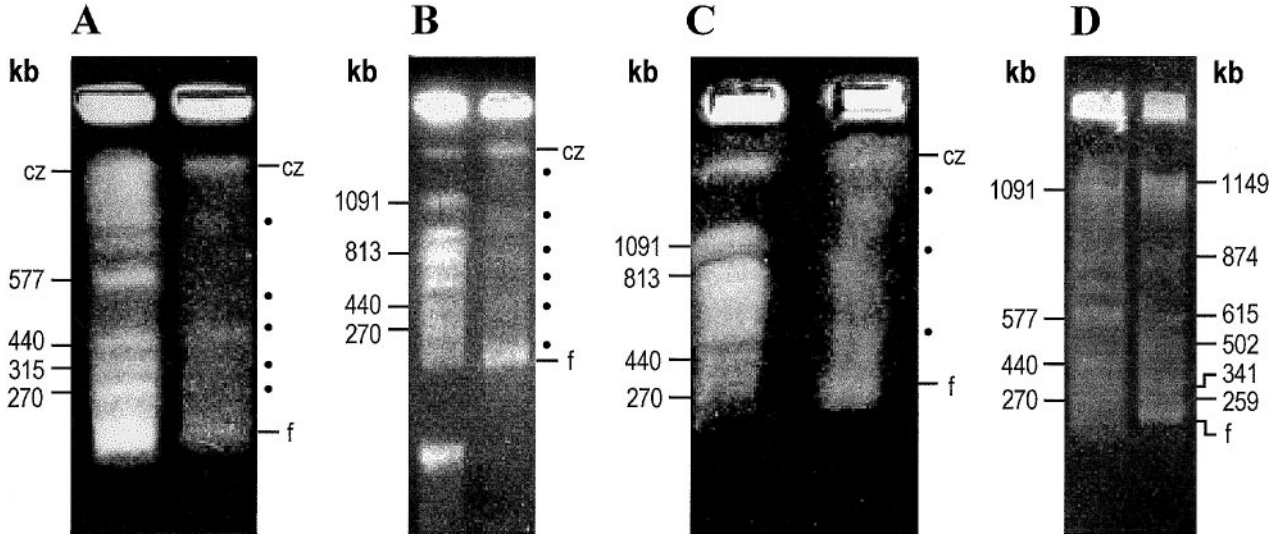

Fig. 1A–D *S. cerevisiae* (*left*) and *E. histolytica* (*right*) DNA electrophoretic patterns obtained in CHEF, miniCHEF and TAFE experiments at 5.8 V/cm, 0.5X TBE, 1.5% agarose gel (Lachema). **A** 24 h, 20 °C, 80 s, CHEF; **B** 24 h, 20 °C, 120 s, CHEF; **C** 15 h 30 min, 15 °C, 200 s, miniCHEF; **D** 24 h, 20 °C, 120 s, TAFE. *Dots* identify bands in the *E. histolytica* karyotype; *f*, fast migrating band; *cz*, compression zone

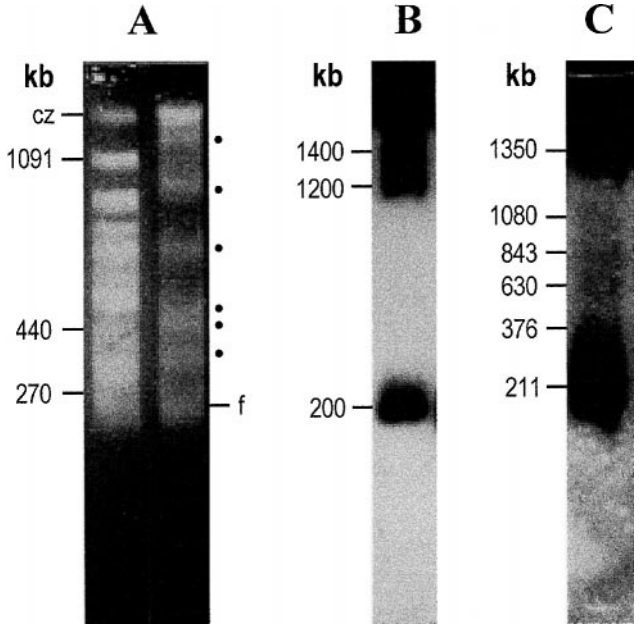

Fig. 2A–C Hybridisation of TAFE-separated *E. histolytica* DNA bands with the ^{32}P radiolabelled *TelTh* and ribosomal probes (16S and 25S). The electrophoresis was performed at 5.8 V/cm, 20 °C, at 120 s of pulse time for 24 h in 0.5X TBE and 1.5% agarose gel (GTG Gold Seakem). **A** Ethidium bromide stained gel showing *S. cerevisiae* chromosomes (*left*) and *E. histolytica* DNA (*right*). **B** Hybridisation with the ribosomal probe, and **C** with the *TelTh* probe. *Dots* identify bands in *E. histolytica* karyotype; *f*, fast migrating band; *cz*, compression zone

1994). As t_r depends on DNA contour lengths, these plots give information on DNA sizes. We identified Eh_1, $Eh_2 \ldots$, Eh_p DNA bands (Eh_k) according to their order of appearance in every CHEF pattern and to their relative position to linear markers. Migrations per pulses after reorientation ($d_{m_{ki}}$) of Eh_k molecules were calculated using data obtained at $t_{p_{ki}}$ larger than the t_r expected at the experimental conditions used (Table 1). Arrays of $d_{m_{ki}}$ and $t_{p_{ki}}$ of each Eh_k were fitted to Eq. (2). Migrations of Eh_1, Eh_4, Eh_5 and the smallest yeast chromosomes behaved similarly (Fig. 3), because $(b_0)_k$ and $(b_1)_k$ differed significantly from zero (Table 2) and $(b_0)_k$ was lower than zero. Eh_2 and Eh_3 molecules gave regression equations with non-significant intercepts $(b_0)_k$ (Table 2). According to Eqs. (1) and (2), the t_{r_k} and v_{m_k} were calculated for Eh_1 to Eh_5 as $t_{r_k} = -(b_0)_k/(b_1)_k$ and $v_{m_k} = (b_1)_k$. To obtain Eh_k sizes we replaced, in Eqs. (5) and (7), E for 5.8 V/cm, η for buffer viscosity at 20 °C and the t_{r_k} previously calculated. For t_{r_k} of 8.1, 17.4, 38.2, 66.4 and 106.3 s the sizes were 211, 368, 650, 969 and 1363 kb, respectively (Table 2). The 368 and 650 kb molecules may correspond to the bands identified by TAFE at 341 and 615 kb (in method I).

In method II, migration data came from CHEF while the method I data were from TAFE. As a molecule moves forward in TAFE gel the reorientation angle increases and the migration per pulse decreases, giving sharper bands than in CHEF even when linDNA and cirDNA are co-migrating in a single band. Thus, in the analysed CHEF assays we were only able to detect a single migration boundary in the region limited by 874–1149 kb, identifying a putative molecule with a size of 969 kb instead of two molecules. The same explanation may be also valid for the size estimation of the 211 kb band.

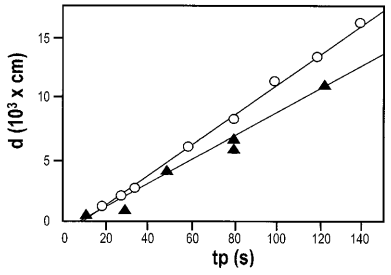

Fig. 3 Migration per pulse (d) as a function of pulse time (t_p) of the (○) 270 kb *S. cerevisiae* chromosome and (▲) Eh_1 211 kb *E. histolytica* band. CHEF electrophoresis conditions: 5.8 V/cm, 20 °C, 0.5X TBE, 1.5% agarose gel

Table 1 Migrations per pulse of *E. histolytica* DNA in CHEF experiments at different pulse times[a]

Migration data							
Electrophoresis conditions			Migrations per pulse (d_{ki}) (10^3 cm) of resolved Eh_k DNA				
i	t_{p_i}(s)	N_{p_i}	(Eh_1) $d_{m_{1i}}$	(Eh_2) $d_{m_{2i}}$	(Eh_3) $d_{m_{3i}}$	(Eh_4) $d_{m_{4i}}$	(Eh_5) $d_{m_{5i}}$
1	10	3600	0.33	–	–	–	–
2	25	1730	1.21	–	–	–	–
3	50	864	4.56	–	–	–	–
4	80	540	6.88	5.30	4.32	0.98	–
5	80	500	6.38	5.31	4.32	1.11	–
6	100	646	8.48	6.65	5.97	2.12	–
7	120	500	10.90	8.72	7.72	3.98	2.37
8	160	400	–	–	–	–	9.99
9	180	270	–	–	–	–	13.05

[a] Experiments were carried out at 5.80 V/cm and 20 °C. The value of i in each row corresponds to the subscripts of t_{p_i} (pulse time), N_{p_i} (pulse number) and d_{ki} (migration per pulse). The migration data in the fourth row were obtained from the electrophoretic pattern shown in Fig. 1A

Table 2 Estimated t_r, v_m and sizes of *E. histolytica* DNA molecules (Eh_k) obtained by fitting to Eq. (2) the migration per pulse data of Table 1

Parameters[a]	Bands				
	Eh_1	Eh_2	Eh_3	Eh_4	Eh_5
$b_0 \times 10^3$	−0.7712	−1.4591*	−3.2422*	−4.7552	−19.1529
$b_1 \times 10^3$: v_m (cm/s)	0.09560	0.08373	0.08493	0.07161	0.18021
$(S^2b_0) \times 10^6$	8.61×10^{-2}	4.04×10^{-1}	1.25	5.05×10^{-1}	1.91
$(S^2b_1) \times 10^6$	1.49×10^{-5}	5.35×10^{-5}	1.35×10^{-4}	5.43×10^{-5}	7.93×10^{-5}
t_r (s)	8.1	17.4	38.2	66.4	106.3
Size (kb)	211	368	650	969	1363

[a] Parameters refer to the regression coefficients b_0 and b_1, their variances S^2b_0 and S^2b_1 and t_r (reorientation time) and v_m (velocity after the reorientation). The $t_{r_k} = (-b_0)_k/(b_1)_k$ and $v_{m_k} = (b_1)_k$. The sizes (in kb) were determined using Eqs. (5) and (7). The symbol * tags the regression coefficients that did not differ significantly from zero

DNA sizes, t_r, v_r and v_m of *E. histolytica* DNA molecules (method III)

CirDNA co-migrates with different linDNA molecules as the pulse time changes. Therefore, no reproducible estimates of size, t_r, v_r or v_m can be obtained for cirDNA. By CHEF and miniCHEF we identified four times bands at 843 ± 9, three times at 1474 ± 13, twice at 1360 ± 4, 1094 ± 27, 630 ± 19 and 376 ± 2 kb (mean size ± mean standard error) and once at 211 and 1781 kb (Table 3). By method III, the 211 kb band has an estimated t_r of 8 s, the 378 kb of 18 s, whereas the 611 kb has a 33 s t_r (Table 3). These $t_{r_{kj}}$ were statistically compared with the $t_{r_{kj}}$ obtained by method II (Table 2). We used the variances (S^2b_0, S^2b_1, Table 2) of the regression coefficients [b_0 and b_1 in Eq. (2), method II] to estimate the variance of t_r ($S^2t_{r_{kj}}$):

$$S^2t_{r_{kj}} = \{(b_0)_{kj}/(b_1)_{kj}\} \times \{((S^2b_0)_{kj}/(b_0)^2_{kj}) + ((S^2b_1)_{kj}/(b_1)^2_{kj})\} \quad (10)$$

Using $S^2(b_1)_{kj}$, $v_{m_{kj}}$ estimates were also compared. According to t-statistic [$p(t) = 0.05$], methods II and III gave similar estimates for the $t_{r_{kj}}$ or $v_{m_{kj}}$ of 211, 378 and 611 kb molecules at 5.8 V/cm and 20 °C, except for the v_m of the 211 kb band. These results suggested that these bands contain linDNA. It remains to be clarified why the v_m of the 211 kb band significantly differed, when measured by both methods. At 5.8 V/cm, 15 °C, 160 and 240 s t_p, molecules of 1364 and 1356 kb appeared (Table 3); they were statistically equal (significant level of 0.05) to the 1363 kb molecule estimated by method II (Table 2). Different t_r (119, 120 and 106 s, respectively) relate to the distinct buffer temperatures used. In the 823–868 kb interval a band appeared at four experimental conditions (Table 3). The mean size (843 kb, Table 4) was statistically equal to the 874 kb band identified by TAFE (Fig. 1D), where a band of 1149 kb was also identified. It was statistically equal to the 1121 and 1067 kb bands (mean size = 1094 kb, Table 4) identified by method III, at two different conditions (Table 3). These bands

Table 3 Kinetic parameters and sizes of *E. histolytica* DNA separated in CHEF and miniCHEF experiments at different conditions[a]

Chamber	Experimental conditions					Kinetic parameters			
	E (V/cm)	T (°C)	t_p (s)	Number of pulses	d_m (10^3 cm)	Size (kb)	v_r (10^3 cm/s)	v_m (10^3 cm/s)	t_r (s)
miniCHEF	5.80	15	240	116	5.309	1781	0.0179	0.07920	173
miniCHEF	5.80	15	240	116	8.848	1449	0.0192	0.08052	130
miniCHEF	5.80	15	200	140	5.256	1485	0.0190	0.08036	134
CHEF	5.80	20	120	328	2.591	1490	0.0216	0.09127	120
miniCHEF	5.80	15	240	116	9.732	1364	0.0196	0.08091	120
miniCHEF	5.80	15	160	167	3.336	1356	0.0196	0.08095	119
miniCHEF	5.80	15	200	140	9.461	1067	0.0214	0.08252	83
CHEF	5.80	20	120	328	3.628	1121	0.0239	0.09337	81
miniCHEF	5.80	15	240	116	15.041	839	0.0232	0.08412	61
miniCHEF	5.80	15	200	140	11.826	823	0.0234	0.08425	60
miniCHEF	10.71	10	50	144	3.869	868	0.0487	0.21498	32
CHEF	5.80	20	120	328	6.220	844	0.0263	0.09552	55
miniCHEF	10.71	10	50	144	6.287	649	0.0539	0.22000	22
CHEF	5.80	20	120	328	8.323	611	0.0295	0.09802	33
miniCHEF	10.71	10	50	144	9.189	375	0.0653	0.22991	10
CHEF	5.80	20	120	328	10.380	378	0.0349	0.10186	18
CHEF	5.80	20	120	328	11.942	211	0.0428	0.10672	8

[a] E electric field, T temperature, t_p pulse time, d_m migration per pulse after reorientation; v_r reorientation velocity, v_m velocity after the reorientation, t_r reorientation time; t_r, v_r and v_m were obtained by solving Eq. (3); d_m of rows 3, 7, 10 and rows 4, 8, 12, 14, 16, 17 were calculated from patterns of Fig. 1C and B, respectively

Table 4 DNA molecules identified in the *E. histolytica* karyotype[a]

Eh_k	Method			Statistics of estimated sizes (methods I, II and III)			TAFE hybridised		Shift Mobility with	
	I	II	III	Mean (kb)	MSE	RSD (%)	*TelTh*	25S	*TelTh*, *ARS* 16S, 25S	
	DNA sizes (kb)								DNA sizes (kb)	
Eh_9(?)			1474					1400		
Eh_8(L)		1363	1360	1361	2.52	0.18	1350			
Eh_7(C?)								1200		
Eh_6(L)	1149	969*	1094	1112	24.06	2.16	1080			
Eh_5(L)	874		843	850	9.45	1.11	843			
Eh_4(L,C)	615	650	630	631	10.57	1.67	630		614	614
Eh_3(C)	502									436
Eh_2(L)	341	368	376	366	8.43	2.31	376			
Eh_1(L,C)	259	211	211	227	16.00	7.05	211	200	246	246

[a] The sizes of linear Eh_k were determined by co-migration with yeast chromosomes in the method I and TAFE hybridised gel, and using Eqs. (5) and (7) in the methods II, III and shift mobility assays. (L), (C) and (?) are mean linDNA, cirDNA or unknown topology, respectively. The sizes estimated by method III are the averages of the size values of Table 3: Eh_9 (rows 2–4), Eh_8 (rows 5–6), Eh_6 (rows 7–8), Eh_5 (rows 9–12), Eh_4 (rows 13–14), Eh_2 (rows 15–16).* Tag the molecules not included in the mean size estimates. MSE: mean standard error; RSD: relative standard deviation [100(MSE/mean)]

hybridised with *Telth* probe (Fig. 2 and Table 4), suggesting that they are linDNA (Table 4).

Except for the molecules Eh_1 (258 kb by method I) and Eh_6 (969 kb by method II) the size estimates for each molecule are statiscally equal by the three methods. Then, we average the sizes estimates obtained by methods I, II and III (Table 4). To calculate the Eh_1 size average we used all values presented in Table 4 because they range between 211 and 259 kb. The mean sizes have relatively small mean error and their relative standard deviations (RSD) were always less than 10%, indicating that the three methods provide size estimates with low variability. Therefore, we used the averaged sizes (1361, 1112, 850, 631, 366 and 227 kb) in the following topics (Tables 3 and 4).

Mobilities of cirDNA and linDNA molecules in the presence of EtBr

In a multilane miniCHEF gel, using different EtBr concentrations, molecules recognised by the *TelTh* probe did not show shift mobilities (Fig. 4A, B), whereas those hybridised with ribosomal probes did (Fig. 4C–E). Plots of *D* against EtBr concentrations showed the differences in migration (Fig. 4F). LinDNA sizes (calculated from the *D* of bands separated without EtBr) were 614, 436 and 246 kb. The 614 and 246 kb bands hybridised with 25S, 16S, *ARS* and *TelTh* probes, whereas the 436 kb hybridised with 25S and 16S probes (Fig. 4A–D, Table 4). Therefore, bands at 246 and 614 kb have linDNA and cirDNA, whereas band at 436 kb may contain only cirDNA. CirDNA usually leaves a smear behind in gel electrophoresis, increasing the band broadness.

CirDNA migration and t_p

Supercoiled DNA (scDNA) reorients significantly faster than linDNA of similar size, and cirDNA

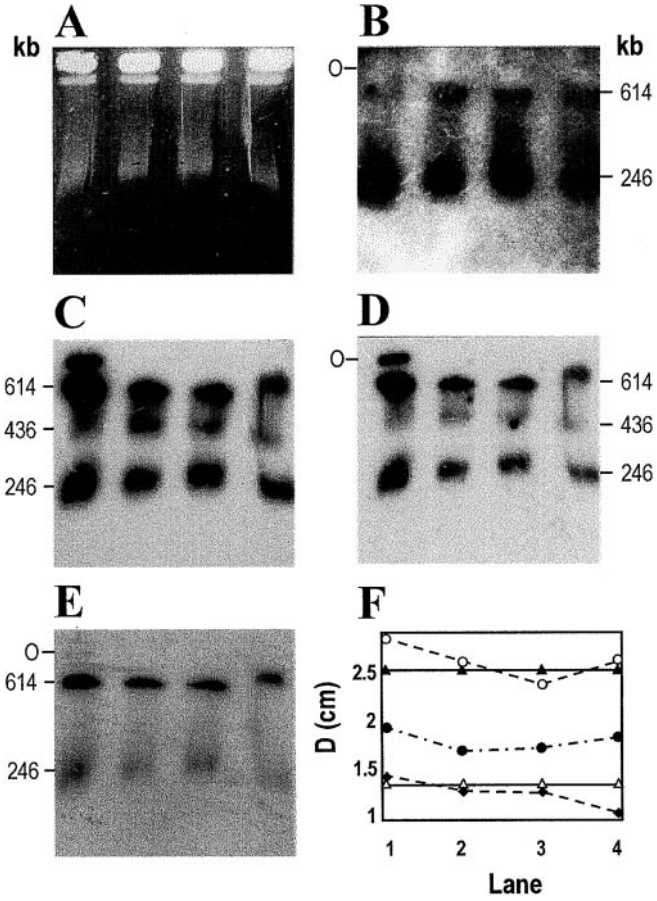

Fig. 4A–F Shift mobility assay of *E. histolytica* DNA in a multilane gel with 0, 1, 2.5 and 5 ng/ml EtBr (*left to right*), respectively. The multigel was run in miniCHEF at 9.04 V/cm, 20 °C, 25 s pulse time for 4 h. **A** Ethidium bromide stained gel. **B–E** Hybridisations with the probes: **B** *TelTh*, **C** 25S, **D** 16S, **E** *ARS*. *O*, gel origin. **F** Plot of the bands migrations (*D*) in lanes 1–4 (*from left to right*), hybridised with: *TelTh* probe, 614 (△) and 246 (▲) kb regions; 25S, 16S and *ARS* probes, respectively, 614 (◆) and 246 (○) kb regions. 25S and 16S probes, 436 (●) kb region

mobility is insensitive to changes in pulse times from 10 to 120 s (Mathew et al. 1988; Hightower et al. 1989; Simske and Sherer 1989). However, a systematic description of migration per pulse of cirDNA as function of t_p, t_r, v_r and v_m is not available. To calculate $d_{scDNA_{wa}}$ of 5.01, 6.03, 7.05, 8.07, 10.10, 12.14, 14.17 and 16.14 kb scDNA, we used mobilities determined in CHEF by Sobral and Atherley (1989). For each scDNA, an array of $d_{scDNA_{wa}}$ and t_{p_a} was fitted to Eq. (8). The coefficients g_{0scDNA} were always zero, but V_{scDNA} differed significantly from zero [$p(t) = 0.05$]. We could not calculate t_r by Eqs. (1) and (8). For 5.01, 6.03, 7.05, 8.07, 10.10, 12.14, 14.17 and 16.14 kb, V_{scDNA} were 0.1993, 0.1659, 0.1463, 0.1280, 0.1034, 0.08764, 0.079116, 0.07225 ($10^3 \times$ cm/s), respectively. Thus, from 1 to 25 s t_p, each scDNA migrated during each pulse with a single velocity V_{scDNA}. V_{scDNA_w} was inversely related to kb_{scDNA_w} by:

$$V_{scDNA} = a_0 + a_1/kb_{scDNA} \quad (11)$$

where $a_0 = 0.01279$ (variance $= 2.1964 \times 10^{-6}$) and $a_1 = 0.93064$ (variance $= 1.3704 \times 10^{-4}$). The coefficients a_0 and a_1 differed from zero [$p(t) \leq 0.05$], so:

$$d_{scDNA} = (0.01279 + 0.93064/kb_{scDNA})t_p \quad (12)$$

The inverse relation between sizes and velocities resembles the migration velocity of linDNA in conventional gel electrophoresis (Lumpkin and Zimm 1982). Equation (12) permits prediction of d_{scDNA} in CHEF at different pulse times, 6.4 V/cm and 13 °C.

Predictions of complex *E. histolytica* patterns with linear and circular DNA

Based on Eqs. (1) and (3)–(12), the EHPATTERNS computer program predicts cirDNA and linDNA migrations in CHEF at 6.4 V/cm, 13 °C and different t_p values. EHPATTERNS uses mean sizes estimated for *E. histolytica* linDNA molecules, and the reported sizes of cirDNA (Dhar et al. 1995; Lioutas et al. 1995). The program informs the array of *E. histolytica* DNA molecules as they are progressively resolved from the compression zone.

To predict migrations per pulses of Eh_k linDNA molecules ($d_{mt_{kji}}$) at the t_p assayed here, EHPATTERNS calculates $t_{r_{kj}}$, $v_{r_{kj}}$ and $v_{m_{kj}}$ by estimating L_k from mean sizes. Further, 5, 13, 26, 38 and 50 kb (sizes of *E. histolytica* supercoiled DNA) were interpolated in Eq. (11) to obtain V_{scDNA}. d_{tscDNA} (theoretical migrations per pulse of these cirDNA) was calculated at the t_p assayed. At each t_p and independently from the molecule topology, the program arranged theoretical migrations per pulses from the smallest value. The array revealed the order of appearance of circular and linear molecules in the gel and also showed the molecules with similar migrations per pulses and thus co-migrating in a band (Table 5). EHPATTERNS shows that at 10 s t_p the 38 kb scDNA comigrates with 227 kb linDNA, at 25 s it appears between 366 and 227 kb, at 50 s between 631 and 366 kb and finally at 240 s t_p it migrates in the compression zone (Table 5). Similar behaviour was found for 13, 26 and 50 kb circles (Table 5). EHPATTERNS predicted that at 100 and 120 s t_p the 38 and

Table 5 Simulations of the order of appearance (from bottom to top) and kinetic parameters of the *E. histolytica* supercoiled and linear molecules for different pulse times[a]

Pulses times (s)									
10	25	50	80	100	120	160	180	200	240
cz	cz	cz	cz	cz	cz	cz	cz	cz	cz
				1112	1361			**50**	**50**
									38
38	**50**	631	850	**50**	1112	1361	**50**	**38**	**26**
227	366				**50**	**50**		1361	1361
26	**38**	**50**	**50**	**38**	**38**	**38**	1361	**26**	1112
				850			**38**		
13	**26**	**38**	**38**	**26**	**26**	1112	**26**	1112	•850
					850	**26**	1112		631
									13
5	227	**26**	**26**	631	631	850	850	850	•227
			631						366
	13	366	366	366	**13**	631	631	•631	**5**
				13	366			**13**	
	5	**13**	**13**	227	227	•**13**	•**13**	•366	
		227	227			366	366	227	
		5	**5**	**5**	**5**	227	227	**5**	
						5	**5**		

[a] Molecules of sizes enclosed in shadowed boxes were co-migrating in a band. The symbol • indicates stacked molecules; cz, compression zone. Kinetic parameters of scDNA were obtained from the simulation at 6.4 V/cm and 13 °C for 5, 13, 26, 38 and 50 kb scDNA (bold numbers); V_{scDNA} were 0.1989, 0.0844, 0.0486, 0.0373, 0.0314 (10^3 cm), respectively. For 1361, 1112, 850, 631, 366 and 227 kb the t_r (s) were 110, 83, 58, 38, 18 and 9; the v_r were 0.0214, 0.0230, 0.0253, 0.0281, 0.0339 and 0.0401 (10^3 cm); and the v_m were 0.0912, 0.0927, 0.0947, 0.0970, 0.1013 and 0.1052 (10^3 cm), respectively

50 kb scDNA co-migrate with the 850 and 1112 kb linDNA, respectively (Table 5). Thus, in CHEF experiments (method II), we identified these bands as a single broad band around 969 kb (Table 2), explaining the results obtained above. The program also reveals that 13 kb scDNA may co-migrate with the 227, 366 or 631 kb linear molecules (Table 5).

EHPATTERNS reproduced two experimental findings previously noted: (1) in complex karyotypes, linDNA molecules frequently co-migrate with supercoiled circular molecules of smaller sizes, and (2) as the pulse time increases, scDNA are left behind by linear molecules of larger sizes. Pattern inversion can be understood by comparing linDNA and cirDNA migration velocities. At given experimental conditions, a scDNA moves at V_{scDNA} during a pulse, but linDNA moves with two velocities: v_r and v_m. If $V_{scDNA} = v_r$, the scDNA and linDNA will comigrate in a band when $t_p < t_r$. If $t_p > t_r$, $v_m > V_{scDNA}$, and linDNA will be left behind the cirDNA.

In conclusion, our results should help to select a priori experimental conditions to separate scDNA and linDNA and will facilitate the interpretation of complex karyotypes, such as that of the *E. histolytica* genome.

Acknowledgements We thank Dr. Lidice Galan for helping us in the statistic procedures. E.O. is an International Fellow of the Howard Hughes Medical Institute (USA). This work was also supported by CONACYT (México).

References

Argüello C, Valenzuela B, Rangel E (1992) Structural organization of chromatin during the cell cycle of *Entamoeba histolytica* trophozoites. Arch Med Res 23: 77–80

Baez-Camargo M, Lopez-Canovas L, Riveron AM, Chavez P, Orozco E (1996a) Fast procedure to distinguish circular and linDNA molecules in pulse field gel electrophoresis. Anal Lett 29: 745–753

Baez-Camargo M, Riveron AM, Delgadillo DM, Flores E, Sanchez T, Garcia-Rivera G, Orozco E (1996b) *Entamoeba histolytica*: gene linkage groups and relevant features of its karyotype. Mol Gen Genet 253: 289–296

Burke DT, Carle GF, Olson MV (1987) Cloning of a large segment of exogenous DNA into yeast by means of artificial chromosomes vectors. Science 236: 806–812

Dhar SK, Choundhury NR, Bhattacharya A (1995) A multitude of cirDNAs exist in the nucleus of *Entamoeba histolytica*. Mol Biochem Parasitol 70: 203–206

Fleiss JL (1986) The design and analysis of clinical experiments. Wiley, New York, pp 41–43

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SW (1996) Life with 6000 genes. Science 274: 546

Grodberg J, Salazar N, Oren J, Mirelman D (1990) Autonomous replication sequences in an extrachromosomal element of a pathogenic *Entamoeba histolytica*. Nucleic Acids Res 18: 5515–5519

Hightower RC, Bliska JB, Cozzarelli NR, Santi DV (1989) Analysis of amplified DNAs from drug-resistant *Leishmania* by orthogonal-field-alternation gel electrophoresis. J Biol Chem 264: 2979–2984

Lioutas C, Schamitz C, Tannich E (1995) Identification of various cirDNA molecules in *Entamoeba histolytica*. Exp Parasitol 80: 349–352

Lopez-Canovas L, Galan L, Orozco E, Riveron AM (1998a) Kinetic properties of DNA migration under clamped homogeneous electric field conditions. DNA size, velocities and reorientation time determined in a single clamped homogeneous electric field run. J Chromat A 806: 123–139

Lopez-Canovas L, Biscay R, Noa MD, Perez Perez G, Herrera JA, Orozco E, Riveron AM (1998b) Comparison of DNA migrations in two clamped homogeneous field chambers of different sizes. Relation between sample thickness and electrophoresis time. J Chromat A 806: 187–197

Lumpkin OJ, Zimm BH (1982) Mobility of DNA in gel electrophoresis. Biopolymers 22: 2315–2316

Mathew MK, Smith CL, Cantor CR (1988) High-resolution separation and accurate determination in pulsed-field gel electrophoresis of DNA. 4. Influence of DNA topology. Biochemistry 27: 9222–9226

Michel B, Lizardi PM, Alagon A, Zurita M (1995) Identification and analysis of the start siteof ribosomal RNA transcription of *Entamoeba histolytica*. Mol Biochem Parasitol 73: 19–30

Orozco E, Baez-Camargo M, Gamboa LI, Flores-Soto E, Valdes J, Hernandez F (1993) Molecular karyotype of related clones of *Entamoeba histolytica*. Mol Biochem Parasitol 59: 29–40

Orozco E, Gharaibeh R, Riveron AM, Delgadillo DM, Mercado M, Sanchez T, Gomez-Conde E, Vargas M, Lopez-Revilla R (1997) A novel cytoplasmic structure containing DNA networks in *Entamoeba histolytica* trophozoites. Mol Gen Genet 254: 250–257

Que X, Reed S (1991) Nucleotide sequence of a small subunit ribosomal RNA (16S-like) gene from *Entamoeba histolytica*: differentiation of pathogenic from isolates. Nucleic Acids Res 19: 5438

Riveron AM, Higginson D, Lopez-Canovas L, Perez G, Garcia HM, Reyes M, Manresa R (1989) Quantitative approach to the tp effect on DNA migration during pulsed field gradient gel electrophoresis: reorientation time and migration rate, Studya Biophys 133: 73–80

Riveron AM, Lopez-Canovas L, Herrera Isidron JA, Ruiz-Esquivel L, Higginson CD, Noa MD, Valdes Acosta F (1994) Molecular weight and kinetic parameters of undergoing pulsed field gel electrophoresis. Arch Med Res 25: 193–198

Riveron AM, Lopez-Canovas L, Herrera JA, Baez-Camargo M, Higginson D, Orozco E (1995) Fast pulsed field minigel electrophoresis of large DNA molecules. Anal Lett 28: 1973–1991

Serwer PP (1980) A technique for electrophoresis in multiple-concentration agarose gels. Anal Lett 101: 154–159

Simske JS, Sherer S (1989) Pulsed-field gel electrophoresis of cirDNA. Nucleic Acids Res 17: 4359–4365

Sobral BW, Atherley AG (1989) Pulsed time and agarose concentration affect the mobility of cccDNA during electrophoresis in CHEF and in FIGE. Nucleic Acids Res 17: 7359–7369

Valdes J, Hernandez F de la C, Ocadiz R, Orozco E (1990) Molecular karyotype of *Entamoeba histolytica*. Trans R Soc Trop Med Hyg 84: 537–541