INTERNATL MICROBIOL (1998) 1: 35-43

RESEARCH ARTICLES

Hervé Moreau
Marie-Line Géraud
Yvonne Bhaud
Marie-Odile Soyer-Gobillard

Observatoire Océanologique de Banyuls,
Laboratoire Arago, Banyuls-sur-Mer, France

Cloning, characterization and chromosomal localization of a repeated sequence in *Crypthecodinium cohnii*, a marine dinoflagellate

Received 19 October 1997
Accepted 15 December 1997

Correspondence to:
Hervé Moreau. Observatoire Océanologique de Banyuls. Laboratoire Arago.
UMR CNRS 7628. BP44.
F-66651 Banyuls-sur-Mer. France.
Fax: +33-468887398.
E-mail: h.moreau@arago.obs-banyuls.fr

Summary Genomic DNA of *Crypthecodinium cohnii* has been extracted in the presence of cetylmethylammonium bromide and hydrolysed by 13 restriction enzymes. No typical ladder-like pattern or isolated band of satellite sequences were found with any of these enzymes. A "mini" genomic DNA library had been made and screened by reverse hybridization to isolate highly repeated sequences. Seven such DNA fragments were sequenced. The copy number of one of them (Cc18), 226 bp long, was estimated at around 25,000, representing 0.06% of the total genome. Cc18 was found to be included in a higher fragment of 3.0 kb by Southern blot analysis after cleavage by *Pst*I. This higher molecular weight fragment could be composed either of tandemly repeated Cc18 sequences, or by only one or a very low copy number of Cc18. In this latter case, these fragments, also repeated 25,000 times would represent 1 to 2% of the total genome. Genomic localization of Cc18 by in situ hybridization on squashed *C. cohnii* cells showed that it was widely distributed on the different chromosomes. All the chromosomes observed displayed Cc18 labeling, which appeared homogeneously distributed. The ability of Cc18 to be a specific molecular marker to distinguish sibling *C. cohnii* species is discussed.

Key words *Crypthecodinium cohnii* · Repeated sequence · Dinoflagellates · Chromosomal localization · Genomic organization

Introduction

Dinoflagellates, which can be heterotrophic or autotrophic, free living or parasitic, are widely distributed in the phytoplankton. Both toxic and non-toxic dinoflagellates can proliferate in sea water, and they can cause important economic and health problems. Dinoflagellates are eukaryotes that have kept several prokaryotic features, and have some original characteristics such as a permanent nuclear envelope, chromosomes condensed throughout the cell cycle and the lack of histones and nucleosomes. The mitotic microtubule spindle is extranuclear, and passes through the nucleus via cytoplasmic channels (for review, see [27]).

All species of dinoflagellates show a high content of DNA in the nucleus, which ranges from 7.0 pg/cell in *Crypthecodinium cohnii*, whose nucleus is haploid [16, 18], to 200 pg/cell in *Gonyaulax polyedra* [3] compared to a range of 0.046–3 pg/nucleus in other unicellular eukaryotes [16]. DNA filaments are packaged in a various number (from 4 to 200, depending of the species [28]) of morphologically identical chromosomes without longitudinal differentiation at either the Q, G or C banding [27]. An unusual base, the hydroxymethyluracil has been found to replace thymine in dinoflagellate genomes at different rates from 37% in *C. cohnii* [14] to 63% in *Prorocentrum micans* [11]. Furthermore, in *C. cohnii*, hydroxymethyluracil is not found uniformly interspersed with thymine, and about 10% of its genomic DNA displays a low hydroxymethyluracil content [14].

Renaturation kinetic studies demonstrated the presence of 55 to 60% of repeated, interspersed DNA in *C. cohnii* genome [1]. This proportion of repeated sequences was confirmed later [12], and their organization in the genome was determined. Half of the genome is composed of unique sequences interspersed with repeated sequence elements with a period of around 600 bp, representing around 95% of the total number of interspersed unique elements [12].

Dinoflagellate chromosomes kept permanently condensed throughout the cell cycle, and in ultrathin sections, nucleofilaments appear in a series of arches [25]. Such a condensation of DNA raises the problem of gene transcription, and the accessibility of the transcription machinery to the coding

sequences. Sigee [24] showed that transcription occurs on extra chromosomal filaments and not on DNA in the main body of the chromosome. He suggested that a large part of genetically inactive DNA should locate inside the chromosome and should play an major role by stabilizing chromosome structure in association with protein matrix. Furthermore, a special conformation of DNA (zDNA) has been described at the periphery of the chromosomes of *C. cohnii*, in accordance with the loops protruding from the chromosome masses [26].

More direct evidence of a different chromosomal localization of coding and non coding sequences in *C. cohnii* has been provided by Anderson et al. [2]. These authors showed that coding sequences in intact nuclei were preferentially digested by restriction enzymes, and that bulk chromosomal DNA, which was inaccessible to these enzymes, contained few, if any, coding sequences. This work confirmed two previous observations which had shown that only 10 to 13% of the DNA in isolated dinoflagellate nuclei was accessible to micrococcal nuclease [5, 10].

In prokaryotes, the great majority of the DNA present in nucleoids codes for proteins. The presence of large part of non coding DNA sequences in dinoflagellates is a major distinction from prokaryotes. In eukaryotes, non coding DNA is an important part of the genome, and it is usually organized as repeated sequences interspersed or tandemly repeated (satellite sequences), in more or less special regions of the chromosomes, as telomeres and centromeres [6, 20]. This organization is rather similar to the dinoflagellate genomic organization described from renaturation kinetic experiments [12]. However, in eukaryotic cells, the non coding DNA does not appear as structural DNA, located into the body of the chromosome [2]. Our present investigations focuse on the idea that this peculiar organization of the dinoflagellate genome could be associated with the permanently condensed state of chromosomes.

C. cohnii had been studied for a long time, and its cell cycle course well described [3, 4]. This species presents many interesting characteristics such as the availability of pure axenic cultures of high level of synchronization rate potentiality, high cell rate growth, relative small genome among dinoflagellates, typical dinoflagellate behavior and worldwide distribution. To further characterize the genomic organization of *C. cohnii,* we cloned and characterized a repeated 226 bp length sequence, which appeared interspersed in the genome, and was found included in a larger DNA fragment. Its genomic localization by in situ hybridization was widely and homogeneously distributed on the whole chromosomes.

Material and methods

Biological Models *C. cohnii* Biecheler 1938, strain Whd (Woods Hole), a heterotrophic dinoflagellate, was grown in MLH medium [29] in the dark at 27°C. Under these conditions, the cell cycle lasts 8 h [3].

Genomic DNA extraction The isolation of genomic DNA, was performed following the method described by Rogers and Bendich [19], and slightly modified by Lee et al. [13], utilizing cethylmethylammonium bromide (CTAB) to remove polysaccharides. In the end, the nucleic acid was precipitated in 60% ethanol, and recovered by centrifugation (500 g). The pellet was air dried, and resuspended in Tris-HCl 10 mM (pH 8.0) and EDTA 1 mM buffer at a concentration around 1 mg/ml.

Genomic Southern blots Five µg of genomic DNA of *C. cohnii* were restricted in a volume of 20 µl, and run on an 1% agarose gel, and blotted onto a Hybond N^+ nylon membrane (Amersham) according to manufacturer's recommendations. The membrane was probed with the Cc18 DNA fragment randomly labeled with biotin, using the high-prime kit (Boehringer). Prehybridization and hybridization steps were performed under high stringency conditions: $NaPO_4$ buffer (pH 7.2) 0.2M, EDTA 1mM, bovine serum albumin (BSA) 1%, SDS 7% and formamide 15%. Prehybridization was done at 65°C for one hour, and hybridization at the same temperature overnight. Washes were performed three times at 65°C in $NaPO_4$ (pH 7.2) 40 mM, 1 mM EDTA and 1% SDS.

Biotinylated probe was revealed according to Boehringer's protocol, by using streptavidin conjugated with peroxidase (dilution 1/20,000), and the luminol as chemiluminescent substrate. After development, membranes were Saran wrapped into saran, and exposed to Kodak X-OMAT films for 5 to 20 min.

Genomic DNA library and screening 10 µg of CTAB purified genomic DNA of *C. cohnii* were incubated 120 min at 37°C with 20 units of *Pst*I. DNA was then extracted with phenol, and phenol/chloroform/isoamylalcohol (25/24/1), and finally precipitated with ethanol. 0.4 µg of this DNA were ligated for 120 min at room temperature with 0.2 µg of Bluescript vector previously cleaved by *Pst*I and dephosphorylated. Efficiency of ligation was measured after transformation and plating of *E. coli* bacteria on LB agar medium containing X-gal and IPTG. White colonies, considered as positive clones, were numbered.

One hundred white colonies were then seeded on each Hybond N^+ (Amersham) membrane, and were hybridized overnight with total genomic DNA of *C. cohnii* previously restricted by *Pst*I, and random labeled with biotin (high-prime kit, Boehringer). Prehybridization, hybridization and washing conditions were as described above for genomic southern blots.

Amplification of clones All clones were obtained in the Bluescript vector, which is flanked by primers T7 and T3. Polymerase chain reactions (PCR) were performed in a volume of 50 µl, with denaturation 1 min at 94°C, annealing 1 min at 55°C and extension 1 min at 72°C (30 cycles of

amplification). DNA was added under a suspension (5 µl) of exponentially growing bacteria containing the vector. Primers T3 and T7 were present at a concentration of 5 pM, and 2 units of *Taq* polymerase (EuroTaq, Eurogentec) were added. Amplification products were analyzed on 1 or 2% agarose gels, and size of DNA determined by comparison with standard MW markers.

Copy number of Cc18 sequence Increasing amounts of total genomic DNA of *C. cohnii* and of purified Cc18 band were immobilized on Hybond N^+ nylon membrane (Amersham) with a manifold device. DNA samples were previously denatured by boiling 5 min, followed by rapid cooling on ice 10 min. Then, after loading onto the membrane, fixation of the DNA to the membrane was performed under alkaline conditions (NaOH 0.4 M, 15 min), and prehybridization, hybridization and revelation steps were identical to those described above for Southern blots.

In situ hybridization Cell squashes were prepared according to Soyer-Gobillard et al. [26]. Slides were then treated with 100 µg/ml RNase A (Boehringer) in 2 × SSC (0.15 M sodium chloride and 0.015 M sodium citrate in distilled water, pH 7.0) for 1 h at 37°C and washed twice in 2 × SSC. A third wash was in Tris-$CaCl_2$ buffer (20 mM Tris-HCL, 2 mM $CaCl_2$, pH 7.4). Samples were treated with 0.5 µg/ml proteinase K (Sigma, preincubated for 3h at 37°C to eliminate possible DNase activity) in Tris-$CaCl_2$ buffer for 10 min at 37°C. Slides were washed in PBS containing 50 mM $MgCl_2$. *C. cohnii* slides were then incubated 30 min at room temperature with 4 N HCl, and washed three times in PBS-$MgCl_2$ buffer. Cells were then fixed for 10 min in 4% paraformaldehyde, washed again in PBS-$MgCl_2$ and finally washed three times in bidistilled water. Slides were washed in PBS-$MgCl_2$, fixed for 10 min in 4% paraformaldehyde, washed again in PBS-$MgCl_2$ and finally washed three times in bidistilled water and air-dried. The hybridization mixture (2 × SSC, 50% deionized formamide, 10% dextran sulfate) contained 3 ng/µl of the biotinylated probe. 25 µl of the mixture were placed on the slide under a coverslip. Denaturation of the probe and the cell samples was done simultaneously by putting the slides into an incubator at 90°C for 10 min. The slides were then transferred overnight in moist chambers to an incubator at 37°C, and then washed three times for 10 min in 50% formamide in 2 × SSC at 42°C followed by two washes in 2 × SSC and one wash in 4 × SSC containing 0.05% Tween 20 (SSC-Tween) at room temperature.

Slides were incubated first in 4 × SSC, 0.05% Tween 20 and 5% non-fat dry milk (SSC-milk) for 10 min at room temperature and then in the above mixture containing 5 µg/l avidin-FITC (Sigma) for 30 min at room temperature and washed three times (5 min each) in SSC-Tween. After 30 min incubation in SSC-milk containing 10 µg/ml biotynylated goat anti-avidin (Sigma), slides were washed three times in SSC-Tween. Finally slides were incubated again in SSC-milk containing 10 µg/ml avidin-FITC for 30 min at room temperature and washed three times in SSC-Tween. Slides were then washed in 4 × SSC alone, mounted in 4× SSC/glycerol (1:1) containing 5% antifading compound N-propylgallate and observed under fluorescent light with a Reichert Polyvar.

Results

Extraction of genomic DNA Genomic DNA isolated from *C. cohnii* by conventional methods such as cell lysis followed by phenol/chloroform extraction was refractory to cleavage by restriction enzymes. This has also been reported in other systems and notably for another dinoflagellate: *Gonyaulax polyedra* [13]. This has been attributed to the presence of polysaccharides and/or pigments co-purified with DNA [13]. The protocol used in this work, using CTAB as selective precipitating reagent [19], yielded more cleavable DNA. The extracted DNA appeared on electrophoresis in 1% agarose gel as a unique band between 20 and 25 kb, which is usually described as the limit of electrophoretic mobility of long DNA fragments in such gels. A very low diffuse background or any was visible below this band after ethidium bromide staining, indicating that minimal shearing had taken place.

13 different restriction enzymes have been tested on this DNA (Fig. 1), and of which, 3 enzymes were found to be the most effective to cleave the DNA: *Nco*I, *Pst*I and *Pvu*II. Seven others (*Apa*I, *Bam*HI, *Bgl*II, *Eco*RI, *Hind*III, *Sal*I and *Xba*I)

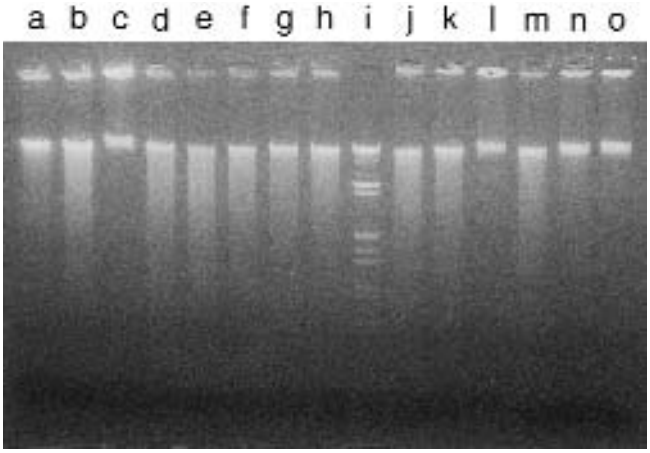

Fig. 1 Agarose gel electrophoresis of *C. cohnii* genomic DNA extracted by the CTAB protocol and cleaved with various restriction enzymes: *Apa*I (a), *Nco*I (b), *Not*I (c), *Pst*I (d), *Pvu*II (e), *Bam*HI (f), *Bgl*II (g), *Eco*RI (h), molecular weight markers (i), *Hind*III (j), *Sal*I (k), *Sma*I (l), *Xba*I (m), *Xho*I (n) and not digested (o). 4 µg of genomic DNA were incubated 60 min at 37°C (except for SmaI: 25°C) in the presence of 10 units of restriction enzyme. DNA bands were revealed by ethidium bromide staining, and molecular weight markers were from lambda DNA phage digested by *Eco*RI and *HindIII*. Sizes were: 21,226 bp; 5,148 bp; 4,973 bp; 4,268 bp; 2,027 bp; 1,904 bp; 1,584 bp; 1,375 bp; 947 bp, 831 bp and 564 bp

Table 1 Reverse hybridization on *Crypthecodinium cohnii* DNA genomic library

Plate	I	II	III
Total number of clones	115	121	123
Number of clones presenting a strong hybridization signal	7 (6.0%)	9 (7.4%)	4 (3.2%)
Number of clones presenting an intermediate hybridization signal	11 (9.6%)	10 (8.3%)	15 (12.2%)
Number of clones presenting a low hybridization signal	62 (53.9%)	54 (44.6%)	72 (58.5%)
Number of clones presenting no hybridization signal	5 (30.5%)	48 (39.7%)	32 (26.1%)

On each plate, 2 clones containing only bluescript without any insert as negative control were seeded, and did not display any hybridization signal.

Table 2 Sequences of clones selected for their strong signal after reverse hybridization

Cc18 TGCAGTGGCA TCAAAGATCC AGTTCCTGAA GTTATTCGGC
TCCCCTACTTCGCTGGCGTC TGCATGACCT TGAAGCTCCA CCACGGCACC TGCAGGAACA
GGTACTTCAA GGCTGGAGGG GCGAACCCGA GGTGGCTGCT GGAGTGGTTC
CCAGAGGACC CCAGCGACAT CCCAGAGCAC TTTGGCACCC CCCATGTCTC
CCTGGTGTCC TACAGGGAGT TCCTGC
Length: 226 bp - 72C; 63G; 47T; 44A. GC ratio = 60%

Cc19 TGCAGCCTCT TGGTCAGACT TCTTCTCGGG GACCTTCCGC
TCTCGGGCTG GTTC
Length: 52 bp - 18C; 15G; 17T; 4A. GC ratio = 63%

Cc20 TGCAGACTGA ATCCTTTGGC TATATGGGTC CCGACTCCTG TAGTAGTCTC
TCCAAAAGC
Length: 59 bp - 16C; 13G; 17T;13A. GC ratio = 49%

Cc21 TGCAGTGCGC GCCAAGCGCT TTGGCATGCC AAAGCAGGGC
ACAGTGTGTC
Length: 50 bp - 14C, 17G, 9T; 10A. GC ratio = 62%

Cc22 TGCAGTGCGC GCCAAGCGCT TTGGCATGCG AAAGCAGGGC
ATAGTGTGTC
Length: 50 bp - 12C; 18G; 10T; 10A. GC ratio = 60%

Cc23 TGCAGGCACT GATGCGCCC
Length: 19 bp - 7C; 6G; 3T; 3A. GC ratio = 68%

Cc24 TGCAGGCACT GATGCGCCC
Length: 19 bp - 7C; 6G; 3T; 3A. GC ratio = 68%

were less efficient, and three of the enzymes displayed no activity or a very weak one (*Not*I, *Sma*I and *Xho*I). No obvious relation could be found between the composition of their recognition sites, their efficiency to cleave the genomic DNA of *C. cohnii* and the replacement of around 37% of thymine residues by the modified nucleotide 5'-hydroxymethyluracil in this organism [14].

Cloning of repeated sequences Bands representing ladders of oligomer did not repeat, nor were isolated bands observed after digestion of the DNA with any of the enzymes used. This indicated that no satellite/repeated sequence could be directly isolated from agarose gels, as is usually done to isolate satellite sequences in eukaryotes [20].

To overcome this difficulty, the reverse hybridization system was used [21]. Genomic DNA fragments generated by *Pst*I digestion were ligated into bluescript vector, and after transformation in bacteria and plating, a total of 10,000 independent positive (white) colonies were obtained. Twelve clones randomly picked gave an average size between 100 and 1000 bp per clone (data not shown). Assuming a quantity of DNA per *C. cohnii* cell of around 7.0 pg [16], this "mini" genomic library would represent between 0.01 and 0.1% of the total haploid genome. Although it is a low rate, it appeared representative of the total genome to isolate the most repeated sequences present in the genome of *C. cohnii*. Around 300 white clones were plated on three different membrane lifts and were analyzed by reverse hybridization with biotin-labelled total genomic *C. cohnii* DNA previously cleaved by *Pst*I. Intensity of hybridization signal was deemed to be representative of the repetitiveness of the insert in the genome. Hybridization and subsequent washes were performed under high stringency conditions. Film exposure showed that clones could be classified into three groups, as a function of the intensity level of the hybridization signal (Table 1). Group I presented the higher level of hybridization, and was found in 6% of total clones, representing 20 independent clones on the three membranes.

Inserts of the seven clones of the first membrane reacting strongly with the probe were amplified by PCR and visualized on agarose gel. Their size was low, below 50 bp for Cc23 and Cc24, around 60 bp for Cc19, Cc20, Cc21 and Cc22, and 250 bp for Cc18. 16 other clones were taken randomly from the library, amplified by PCR and their size was determined in an agarose gel. The average size of these inserts of the same range as those obtained with clones selected after reverse hybridization.

The seven clones from group I of the first membrane were further analyzed and totally sequenced (Table 2). Their molecular weights, previously determined electrophoretically, were confirmed by the determination of the sequences. These seven clones had been isolated independently, but sequences of four of them appeared highly related two by two. Cc23 and Cc24 were very short (i.e., 19 bp), and totally identical. Cc21 and Cc22 displayed 96% of homology on their sequences of 50 bp length for both inserts. Furthermore, for these two sequences, a palindromic sequence, composed of 18 nucleotides, was seen in Cc21 DNA, and with one different nucleotide in Cc22. Fragments Cc11, Cc17 and Cc18, Cc19 and Cc20 were unique, and no obvious repeat, inversion or other structural particularity could be observed, except in an inverse repeat of 8 bp in Cc19. A *Pst*I site had been sequenced inside the Cc18 fragment, indicating that this site was not accessible to *Pst*I enzyme in the genomic DNA of *C. cohnii*. The GC content (60%) found in all the clones, except Cc20, was high if compared with the average previously reported for genomic DNA of *C. cohnii* (40%) [14]. No clear homology could be found in genbank and EMBL data bank, using the blast homology research system for the seven clones. Only very partial homologies were found on limited portions of these sequences, with various coding or non coding DNA sequences, without any clear significance.

Characterization of Cc18 Cc18 was the longer isolated sequence (226 bp), allowing a high rate of biotin incorporation by random priming method, and was further characterized.

To estimate the relative number of copies of Cc18 sequences in *C. cohnii* genome, the intensity of the hybridization signal was compared between increasing amounts of both total *C. cohnii* DNA, and Cc18 DNA. DNA samples were spot-blotted onto nylon membranes and subsequently hybridized to Cc18 insert restricted and purified from its vector. 2,500 ng of unrestricted genomic DNA gave a hybridization signal equivalent to 1.5 ng for Cc18 (Fig. 2). This result has been confirmed three times. Since the *C. cohnii* cell contains around 7.0 pg of DNA (around 8.0×10^9 bp) (16), this corresponded to a value of about 20,000 copies of Cc18 per cell, and represented 0.06% of the total genome of *C. cohnii.*

On each dot blot performed to determine the copy number of Cc18 in *C. cohnii* genome, the same increasing amounts of genomic DNA from another dinoflagellate species *P. micans* were spotted onto the membranes (Fig. 2), to determine the specificity of Cc18 sequence among the dinoflagellate group. No hybridization signal with labelled Cc18 could be detected at any quantity of *P. micans* DNA tested (up to 5 µg).

To further characterize Cc18 sequence, a kinetic digestion of *C. cohnii* genomic DNA by *Pst*I was analyzed by Southern blot and hybridized with Cc18 probe (Fig. 3). After 5 minutes of digestion (Fig. 3b), Cc18 signal was confined to a major

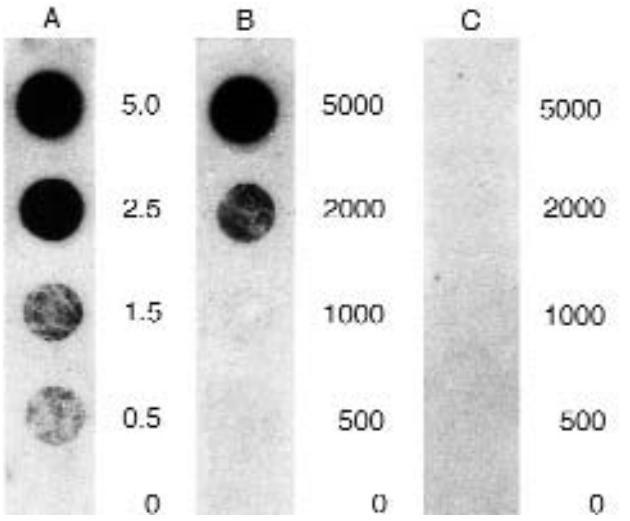

Fig. 2 Estimation of the copy number of Cc18 in the genome of *C. cohnii* and *P. micans*. Different amounts of Cc18 DNA (A), *C. cohnii* genomic DNA (B) or *P. micans* genomic DNA (C) were spotted on Hybond N^+ membrane (Amersham) and hybridised with biotinylated Cc18 probe. Values reported on each spot on the figure represent the quantity of DNA (in ng) spot-blotted on the nylon membrane

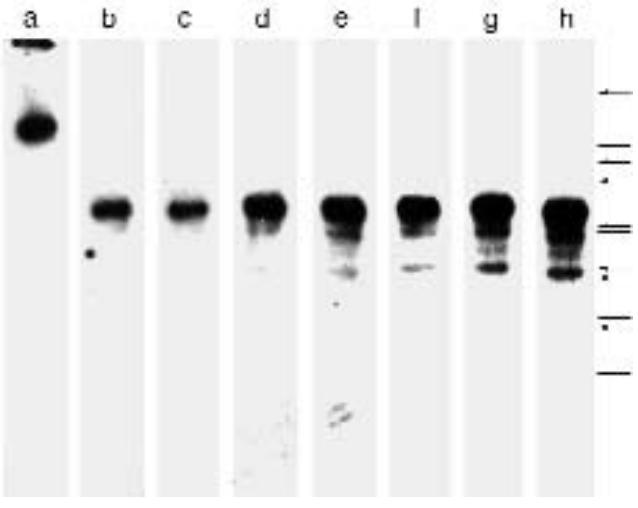

Fig. 3 Southern blot analysis of *C. cohnii* genomic DNA digested by PstI. 2.6 µg of genomic DNA were digested during various times, and hybridized with Cc18 biotinylated probe: not digested (a), 5 min (b), 10 min (c), 30 min (d), 60 min (e), 120 min (f), 240 min (g) and 360 min (h). Molecular weight markers are similar to those described in Fig. 1

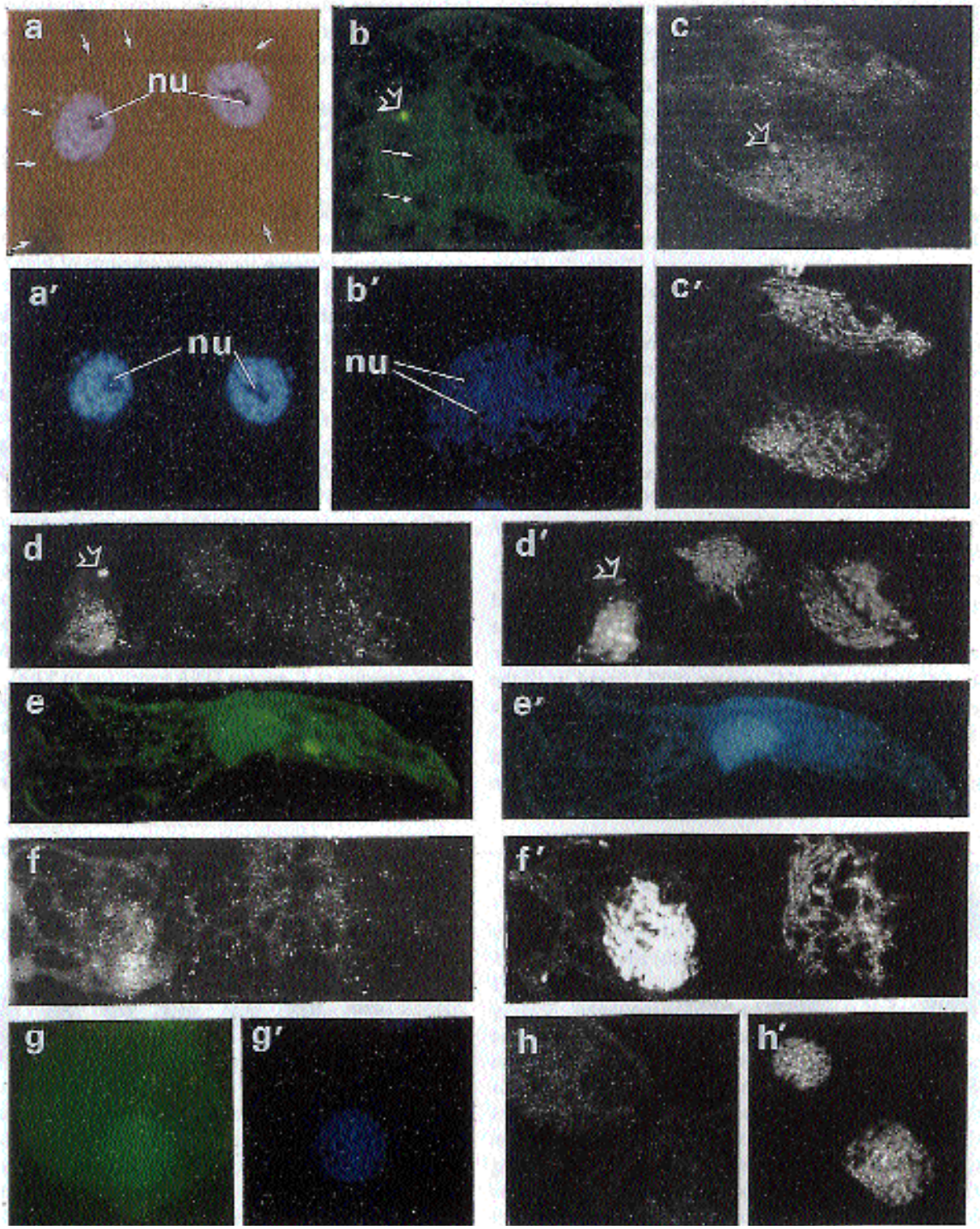

Fig. 4 Genomic localization of Cc18 sequences on *C. cohnii* squashes, by *in situ* hybridization of Cc18 biotinylated probe. a, a' Not hybridized *C. cohnii* squashed cell in cytodieresis, where nuclei were DAPI stained. Cytoplasmic limit (a, arrows) and nucleolar regions (nu) are visible. b-f The Cc18 probe was detected in the nuclei with FITC-conjugated avidin as small bright spots either in non dividing (b) or in cytodieresis (c) cells, and either on compacted (d, f, at left) or in squashed nuclei (d, f, at right). The nucleolus regions (b' arrows) were not fluorescent (b). Note the bright autofluorescent point (empty arrows), usually seen in *C. cohnii* cells. All filaments were brigtly fluorescent (e) where the chromatin was spread out from the nuclei (e'). g-h Negative controls performed either by omitting the nucleic acid probe (g), or by replacing Cc18 by an irrelevant probe (h). b'-h' are the corresponding DAPI stained nuclei. a, a', e, e', g, g', h, h': × 3,200; d, d': × 2,800; b, b', c, c', f, f': × 4,400

band of 3.0 kb, and hybridization signal totally disappeared in the undegraded DNA (21 kb band). This 3.0 kb band remained resistant to digestion, even after 24 hours of incubation. After 2 hours of incubation (Fig. 3f), a slow reinforcement of lower bands could be observed, but no typical ladder-like pattern, characteristic of satellite sequence tandemly repeated, was observed. 226 bp Cc18 band did not appear on these blots, although it should be present at least at low level after *Pst*I digestion, because Cc18 fragment had been cloned from such a digested DNA.

Chromosomal localization of Cc18 A modification of the in situ hybridization protocol previously used to localize ribosomic genes in *P. micans* [7] allowed to obtain a clear hybridization signal with Cc18 as a probe (Fig. 4). Heat incubation and/or the different RNase and protease (proteinase K and pepsin) treatments alone failed to reveal Cc18 on chromosomes, whereas incubation of squashes in the presence of 4 N HCl gave positive results. Furthermore, in contrast to squashes on *C. cohnii*, no hybridization signal could be obtained with *C. cohnii* cryosections, whatever the treatment performed (RNase, protease or HCl incubations).

Cc18 hybridization signal appeared on Cc squashes as numerous small patches widely distributed on chromosomes. The hybridization signal was uniformly dispatched along all the chromosomes (Fig. 4b and Fig. 4c), materialized after DAPI staining (Figs. 4b' and c'). The signal was detected whatever the degree of compaction of the chromatin (Figures 4d, d' and f, f'; compare right and left nuclei). Furthermore, a high intensity of labelling was also observed on long DNA spread filaments (Fig. 4e) localized outside of the chromosomes. These filaments could be easily observed after DAPI staining (Fig. 4e'), and corresponded to unwound DNA filaments dispatched outside of the nucleus during the breakdown of the nuclear membrane, when the squashes were prepared. When the nucleolus region was examined in DAPI stained cells (Fig. 4ab'), no signal was observed (Fig. 4b).

Negative controls were performed, either by omitting the nucleic acid probe (Fig. 4g), or by replacing Cc18 by an irrelevant probe (Fig. 4h): a 250 bp non repeated sequence of *C. cohnii* randomly sampled among clones that presented no hybridization signals after the reverse hybridization screening. In all these experiments, only a faint and uniform green background was observed in *C. cohnii* cells, clearly distinct from the patched staining observed with heat denatured Cc18 probe. These results also indicated that this experimental probe detection system was not sensitive enough to detect a very low copy number of a genomic DNA sequence as the irrelevant probe used.

Whereas no signal was observed after hybridization of Cc18 probe on *P. micans* squashes and cryosections, a positive signal comparable to that obtained in a previous study [23] was clearly seen on *C. cohnii*, with a specific probe to ribosomic genes (data not shown).

Discussion

Restriction enzymes are usually very sensitive to methylation and any other kind of chemical modification of their target sequences [22], and the replacement of thymine by hydroxymethyluracil in these target sites might render them resistant to cleavage by the corresponding restriction enzyme. Such a replacement had been described in all dinoflagellate species tested, and in *C. cohnii* genomic DNA it amounted to around 40% of thymines. In the present study, no clear relation could be established between the efficiency of the different restriction enzymes used, and the presence of thymine in their recognition sites, confirming previous observations made with another dinoflagellate species: *Gonyaulax polyedra* [13].

Hydroxymethyluracil was found to be non-uniformly interspersed with thymine in the DNA, and about 10% of *C. cohnii* DNA was contributed by a fraction with low hydroxymethyluracil content [14, 15]. The non random distribution of this unusual base in *C. cohnii* genome was indirectly confirmed by the highly reproducible cleavage pattern obtained in Southern blot. *Pst*I induced a rapid (it only took a few minutes) and total restriction of a 3.0 Kb fragment which contained Cc18 sequence. This result had been obtained with DNA from different preparations originating from different cell cultures. This clearly indicates that *Pst*I sites flanking this 3.0 Kb fragments were reproducibly sensitive to this enzyme, and consequently thymines present in their target sequence were probably never replaced by hydroxymethyluracil. At least three other *Pst*I sites have been identified in the 3.0 Kb fragment: two sites flanking the Cc18 sequence, and one inside. These sites were reproducibly cleaved after cloning of Cc18 DNA fragment in Bluescript vector, and introduced in *E. coli* for amplification, but were poorly restricted in *C. cohnii* genomic DNA. Thymines of these internal sites of the 3.0 Kb fragment could be replaced by hydroxymethyluracil, inducing a very low cleavage by *Pst*I.

Cc18 was included in a larger DNA fragment of 3.0 Kb after restriction of *C. cohnii* genomic DNA by *Pst*I. Two hypotheses can be formulated concerning the structure of this larger fragment. One possibility is that Cc18 be highly repeated in this fragment, which might consist of only Cc18 tandemly repeats. Alternatively, Cc18 might be unique in this fragment (or present in a very low copy number). In the first case, the 3.0 Kb fragment would be repeated between 20,000 and 25,000 times in *C. cohnii* genome, and such a repetition would account for 1 to 2% of the total *C. cohnii* genome. Finally, this fragment might also be included in a bigger one, representing a higher percentage of the total genome.

Satellite sequences, more or less located in peculiar chromosomal regions (centromeres and telomeres), seem to be a constant feature in eukaryotes. However, the particular behavior of dinoflagellate chromosomes, their permanent

condensed state throughout the cell cycle, non banding and the absence of morphologically discernable centromeric structures, reveal an unusual genomic organization. By cleavage of *C. cohnii* genomic DNA with any of the enzymes tested, neither bands representing ladders of oligomer repeats nor isolated bands could be seen; consequently, no satellite tandemly repeated sequences could be directly isolated. Several possibilities can explain the absence of such restriction pattern characteristic of satellite sequences in this work: (i) none of the target sequences of the 13 restriction enzymes used were present in such hypothetical satellite sequences; (ii) in many sites of these sequences, a high proportion of thymines was replaced by hydroxymethyluracil, and was resistant to cleavage, hiding such a repeated structure; and (iii) such tandemly repeated sequences do not exist in *C. cohnii* genome. Although no clear indications of the formal existence of such sequences in *C. cohnii* genome can be found in previous studies, the high percentage of repeated sequences [1] makes the last hypothesis very unlikely.

The GC content of the repeated DNA sequences isolated in this study was high for 5 sequences. This was markedly different to the average GC content of genomic DNA of *C. cohnii,* which had been found around 40% [14]. But this percentage seems variable among dinoflagellate species; 60% of GC content had been determined in *P. micans* [11]. Satellite sequences in eukaryotes display a high GC content, and this explains their different sedimentation in CsCl gradients.

In situ hybridization experiments showed that classical protocols used to denature the genomic DNA (heat treatment) were inefficient with the *C. cohnii* genome, although they had been previously successfully used for *P. micans* ribosomic genes [7]. Dinoflagellates display a particular genomic organization, and their chromosomes contain neither nucleosomes nor histones [10, 16, 23]. Several models of DNA packaging into dinoflagellate chromosomes have been proposed that were essentially based on electron microscopy studies [17, 28], although more data are needed to find a model that solves this intriguing DNA organization. However, from these studies, it was obvious that the DNA packaging in *P. micans* [8] was different than in *C. cohnii* [9]. In *C. cohnii*, whose DNA was very densely packed into small chromosomes, target sequences for in situ hybridization experiments cannot be accessible to the nucleic acid probes used. By now, the only possibility to label these sequences would be to elongate the chromosome as in the squash technique, with a signal intensity proportional to the unstacking of DNA filaments.

Cc18 labelling appeared uniformly distributed on all chromosomes, and all along each chromosome. This indicates that Cc18 is widely interspersed in the *C. cohnii* genome, and not restricted to specfic areas of the chromosomes.

This sequence was not detected in *P. micans* by Southern-blot, nor by in situ hybridization techniques, *P. micans* being a dinoflagellate species very different to *C. cohnii*. Research is in progress to test the presence of this sequence in different sibling species of *C. cohnii,* to evaluate the degree of specificity of Cc18, which could constitute a good tool to discriminate, by molecular hybridization techniques, between sibling or closely related species with no discernable morphological differences.

Acknowledgements The authors wish to thank M. Albert and P. Llaona for their excellent technical assistance, and Dr. T. M. Preston for critical reading of the manuscript. This work was supported by the CNRS (UMR 7628).

References

1. Allen JR, Roberts TM, Loeblich III AR, Klotz LC (1975) Characterization of the DNA from the dinoflagellate *Crypthecodinium cohnii* and implications for nuclear organization. Cell 6:161-169
2. Anderson DM, Grabher A, Herzog M (1992) Separation of coding sequences from structural DNA in the dinoflagellate *Crypthecodinium cohnii*. Mol Mar Biotech 1:89-96
3. Bhaud Y, Salmon JM, Soyer-Gobillard MO (1991) The complex cell cycle of the dinoflagellate protoctist *Crypthecodinium cohnii* as studied in vivo and by cytofluorimetry. J Cell Sci 100:675-682
4. Bhaud Y, Barbier M, Soyer-Gobillard MO (1994) A detailed study of the complex cell cycle of the dinoflagellate *Crypthecodinium cohnii* Biecheler and evidence for variation in histone H1 kinase activity. J Euk Microbiol 41:519-526
5. Bodansky S, Mintz LB, Holmes DS (1979) The mesokaryote *Gyrodinium cohnii* lacks nucleosomes. Biochim Biophys Res Commun 88:1329-1336
6. Elder JF, Turner BJ (1995) Concerted evolution of repetitive DNA sequences in eukaryotes. Quart Rev Biol 70:297-320
7. Géraud ML, Herzog M, Soyer-Gobillard MO (1991) Nucleolar localization of rRNA coding sequences in *Prorocentrum micans* Ehr. (dinomastigote, kingdom Protoctist) by in situ hybridization. BioSystems 26:61-74
8. Haapala OK, Soyer MO (1973) Structure of dinoflagellates chromosomes. Nature New Biol 244:195-197
9. Haapala OK, Soyer MO (1974). Absence of longitudinal differentiation of dinoflagellate (*Prorocentrum micans*) chromosomes. Hereditas 78:146-150
10. Herzog M, Soyer MO (1981) Distinctive features of dinoflagellate chromatin. Absence of nucleosome in a primitive species *Prorocentrum micans* E. Eur J Cell Biol 23:295-302
11. Herzog M, de Marcillac GD, Soyer MO (1982) A high level of thymine replacement by 5-hydroxymethyluracil in nuclear DNA of the primitive dinoflagellate *Prorocentrum micans* E. Eur J Cell Biol 27:151-155
12. Hinnebusch AG, Klotz LC, Immergut E, LoeblichIII AR (1980) Deoxyribonucleic acid sequence organization in the genome of the dinoflagellate *Crypthecodinium cohnii*. Biochemistry 19:1744-1755
13. Lee DG, Mittag M, Sczekan S, Morse D, Hastings JW (1993) Molecular cloning and genomic organization of a gene for luciferin-binding protein from the dinoflagellate *Gonyaulax polyedra*. J Biol Chem 268:8842-8850
14. Rae PMM (1973) 5-Hydroxymethyluracil in the DNA of a Dinoflagellate. Proc Nat Acad Sci USA 70:1141-1145
15. Rae PMM (1976) Hydroxymethyluracil in eukaryote DNA: a natural feature of the Pyrrhophyta (dinoflagellates). Science 194:1062-1064
16. Rizzo PJ, Nooden LD (1973) Isolation and chemical composition of Dinoflagellate nuclei. J Protozool 20:666-672
17. Rizzo PJ (1987) Biochemistry of the dinoflagellate nucleus. In: Taylor FJR (ed), The Biology of Dinoflagellates. London: Blackwell Sci Pub, pp 143-173
18. Roberts TM, Tuttle RC, Allen JR, Loeblich III AR, Klotz LC (1974) New genetic and physicochemical data on structure of dinoflagellate chromosomes. Nature 248:446-447

19. Rogers SO, Bendich AJ (1988). Extraction DNA from plant tissues. In: Gelvin SB, Schilperoort RA, Verna DPS (eds) Plant Molecular Biology Manual, A6. Dordrecht: Kluwer Academic Publishers, pp 1-10
20. Ruiz-Lara S, Prats E, Sainz J, Cornudella L (1992) Cloning and characterization of a highly conserved satellite DNA from the mollusc *Mytilus edulis*. Gene 117:237-242
21. Sainz J, Prats E, Ruiz S, Cornudella L (1992) Organization of repetitive sequences in the genome of the echinoderm *Holothuria tubulosa*. Biochimie 74:1067-1074
22. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual, 2nd ed, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 5.1-5.95
23. Shupe K, Rizzo PJ (1983) Nuclease digestion of chromatin from the eukaryotic algae *Olisthodiscus luteus*, *Peridinium balticum,* and *Crypthecodinium cohnii*. J Protozool 30:599-606
24. Sigee DC (1984) Structural DNA and genetically active DNA in Dinoflagellate chromosomes. BioSystems 16:203-210
25. Soyer MO, Haapala OK (1974) Division and function of dinoflagellate chromosomes. J Microsc 19:137-146
26. Soyer-Gobillard MO, Géraud ML, Coulaud D, Barray M, Théveny B, Révet B, Delain E (1990) Location of B- and Z-DNA in the chromosomes of a primitive eukaryote Dinoflagellate. J Cell Biol 111:293-308
27. Soyer-Gobillard MO (1996) The genome of the primitive eukaryote Dinoflagellates: organization and functioning. Zoological studies 35:78-84
28. Spector DL (1984) Dinoflagellate nuclei. In: Spector DL (ed) Dinoflagellates. New York: Academic Press, pp 107-147
29. Tuttle RC, LoeblichIII AR (1975) An optimal growth medium for the dinoflagellate *Crypthecodinium cohnii*. Phycologia 14:1-8