

Extreme heterogeneous composition of the *Paramecium caudatum* macronuclear genomic DNA between hemoglobin and nucleosome assembly protein-1 genes

Norihito Nishiyama^{1,3†}, Kazuyuki Mikami², Arika Matsuoka³,
Takehiko Ochiai³ and Kiyoshi Yamauchi^{1*}

¹*Department of Biological Science, Faculty of Science, Shizuoka University,
Shizuoka 422-8529, Japan*

²*Environmental Education Center, Miyagi University of Education,
Aoba-ku, Sendai 980-0845, Japan*

³*Department of Biology, School of Medicine, Fukushima
Medical University, Fukushima 960-1295, Japan*

(Received 31 January 2010, accepted 4 March 2010)

The intergenic region between the hemoglobin (*hb*) and nucleosome assembly protein-1 (*nap-1*) genes in the *Paramecium caudatum* macronuclear genome was previously found to be heterogeneously composed. Cloning of this intergenic region from the macronuclear genomic DNA identified four unique DNA fragments of different sizes. Sequencing of the cloned fragments revealed extreme heterogeneity and characteristics of both internal eliminated sequence (IES) and imprecise internal deletion sequences (IIDSs) in the intergenic region. Missing sequences were an AT-rich and direct repeats existed in their boundaries. Southern blotting of the total genomic DNA and polymerase chain reaction (PCR) of the total genomic DNAs indicated that there exist a dozen DNA fragments of different sizes in this intergenic region. It is likely that the heterogeneity found in the *P. caudatum* macronuclear genome results from the variable removal of an intergenic region.

Key words: macronucleus, intergenic heterogeneity, internal eliminated sequences, *Paramecium caudatum*

INTRODUCTION

Ciliated protists have a characteristic nuclear dimorphism. Each vegetative cell contains two types of nuclei: the micronuclei and the macronuclei. The micronuclei are small diploid nuclei that are almost inactive transcriptionally, whereas the macronuclei are large polygenomic nuclei that are vegetative and transcriptionally active (Prescott, 1994). The micronuclei and macronuclei are, therefore, thought to correspond to the germline and somatic nuclei of multicellular eukaryotes, respectively. Macronuclei divide amitotically, are responsible

for vegetative proliferation, and are destined to be destroyed during the subsequent sexual processes: conjugation and autogamy. After fertilization, the mitotic products of the diploid zygotic nucleus differentiate into new macronuclei and micronuclei, depending on their position in the cell (Mikami, 1980; Grandchamp and Beisson, 1981).

During the sexual processes, new macronuclear chromosomes are processed by extensive programmed rearrangements of DNA. These include chromosome fragmentation and precise deletion of germline-specific elements known as internal eliminated sequences (IESs), and then amplification of the fragmented chromosomes (Prescott, 1994). Fragmentation occurs at a few hundreds of breakage sites and gives rise to macronuclear chromosomes in *Paramecium*. The *de novo* addition of telomeric repeats usually does not occur at a precise nucleotide position (Le Mouél et al., 2003). In some cases, developmentally regulated chromosome fragmentation linked to imprecise elimination in macronuclear genome (Caron, 1992;

Edited by Yasuyoshi Nishida

* Corresponding author. E-mail: sbkyama@ipc.shizuoka.ac.jp

† Present address: Department of Biology, School of Medicine, Fukushima Medical University, Fukushima 960-1295, Japan.

Abbreviations: Cbs, chromosome breakage sequence; *hb*, hemoglobin; *nap-1*, nucleosome assembly protein-1; IES, internal eliminated sequence; PCR, polymerase chain reaction; RNAi, RNA interference; IIDS, imprecise internal deletion sequence

Klobutcher and Herrick, 1997; Forney, 1997), thus generating microheterogeneity among the multiple copies of the genome present in each developing macronucleus. We named this type of the deleted sequences here imprecise internal deletion sequences (IIDSs). In *Paramecium*, characteristics of IIDSs remain largely unknown, but recently, some of them were uncovered; short TA-containing direct repeats in their boundaries (2–6 bp), colocalization of telomere addition regions and imprecise internal deletion boundaries, clustering of their heterogeneous boundaries in TA-rich hot spots located outside the deleted sequences, deletion sequences containing an AT-rich minisatellite and truncated transposase gene (Le Mouël et al., 2003). In *Tetrahymena* and *Euplotes*, chromosome breakage also requires *cis*-acting sequence as chromosome breakage sequence (Cbs). Cbs has been identified so far in the vicinity of chromosome fragmentation sites. Classic Cbs have the sequence TAAACCAAC-CTCTTT in *Tetrahymena* (Yao et al., 1987), and a 10 bp conserved sequence element (5'-HATTGAAaHH-3', H = A, C or T) in *Euplotes* near chromosome fragmentation sites (Klobutcher et al., 1998). These *cis*-acting sequences are necessary and sufficient for site-specific chromosome fragmentation. But no conserved sequence has been identified so far in vicinity of *Paramecium* chromosome fragmentation sites (Le Mouël et al., 2003).

In contrast, characteristics of IES excision has been well investigated. In *Tetrahymena*, IESs are precisely excised, by a RNA interference (RNAi) like mechanism (Mochizuki and Gorovsky, 2004; Kowalczyk et al., 2006), from the mitotic copies of the diploid zygotic nucleus. Direct repeats of 4–8 bp, which do not inevitably contain the 5'-TA-3' sequence, are present at the boundaries of the IESs (Prescott, 1994; Klobutcher and Herrick, 1997). IESs in the *Paramecium aurelia* complex genome are short sequences (28–882 bp) located between 5'-TA-3' direct repeats, which were called TA-IESs. Upon excision of the TA-IESs, one of the 5'-TA-3' direct repeats remains in the macronucleus-destined sequences (Steele et al., 1994). Sequencing of many IESs also revealed that the inverted repeat (IR) consensus sequences from *Euplotes crassus* (5'-TATAGAGG-3', Jaraczewski and Jahn, 1993) and *P. aurelia* complex (5'-TAYAGYNR-3', Klobutcher and Herrick, 1995) were similar to those of the metazoan Tc1/*mariner* transposons (Klobutcher and Herrick, 1995). This suggested that the origin of ciliate IESs might have been some type of transposons (Klobutcher and Jahn, 1991; Doak et al., 1994).

Southern blotting of the *P. caudatum* hemoglobin gene (*hb*) revealed at least three major different sized fragments in the macronuclear genome (Yamauchi et al., 1992b). This result raised several possibilities: (1) a *hb* gene family exists whose members are structurally and functionally related (Usuki et al., 1989), and are distributed on distinct chromosomes, like the mammalian α and

β *hb* genes (Hardison, 1998); (2) several copies of the same or highly related *hb* gene, produced recently by gene duplication generating a tandem repeated gene cluster similar to mammalian and invertebrate *hb* gene clusters (Efstratiadis et al., 1980; Antoine and Niessing, 1984), exist; (3) IESs at the *hb* gene or its flanking regions are excised alternatively or differentially to generate heterogeneity during macronuclear development (Austerberry et al., 1984; Herrick et al., 1987); (4) IIDSs are imprecisely excised, leading either to chromosome fragmentation followed by the addition of new telomeres or to the rejoining of flanking sequences (Caron, 1992; Le Mouël et al., 2003); (5) homologous or non-homologous recombinations occur among the hundreds of copies of macronuclear DNA during or after macronuclear development (Deak and Doerder, 1998).

To understand why several DNA fragments containing the *hb* gene are generated in the *P. caudatum* macronuclear genomic DNA, four DNA fragments were cloned and sequenced. Their organization in the macronuclear genome was analyzed by Southern blotting and polymerase chain reaction (PCR). The *hb* genes in the DNA fragments different in size had the same sequence. A variety of missing sequences were found in the intergenic region between the *hb* and nucleosome assembly protein-1 (*nap-1*) genes. These missing sequences and their boundaries showed some features characteristic of IESs and IIDSs. It is likely that some segments of the intergenic region may be imprecisely deleted, during macronuclear development, producing at least three major and several minor closely related macronuclear DNA fragments.

MATERIALS AND METHODS

Cell lines and cultivation The *P. caudatum* stock (syngen 3, stock YC) previously analyzed (Yamauchi et al., 1992b) was used here. *P. caudatum* cells were cultured at 20–25°C in fresh lettuce juice diluted with Dryl's physiological salt solution (Dryl, 1959), which was modified by the substitution of KH₂PO₄ for NaH₂PO₄, pH 7.0, and inoculated with *Klebsiella pneumoniae* the day before use (Mikami, 1987). To produce offspring, cells from stock YC were crossed with those from stock 27aG3. Under starvation conditions, sexually matured cells of complementary mating types were mixed together, forming tight conjugation pairs. After conjugation, the single cell that was separated was transferred into fresh culture medium. Three exconjugant F1 clones, G3YC-1, G3YC-2 and G3YC-4, that have the same zygotic genome but differ in maternal inheritance, were isolated from each conjugant pair (Mikami and Hiwatashi, 1975). Karyonide clones that received an independently differentiated new macronucleus and a mitotic copy of the functional micronucleus, were isolated from an exconjugant clone.

Isolation and characterization of genomic clones

Total genomic DNA was extracted as described previously (Yamauchi et al., 1992b). Partial genomic libraries of *P. caudatum* total genomic DNA were constructed using λ EMBL3 and λ gt11. These libraries were screened by plaque hybridization (Sambrook et al., 1989) using 32 P-labeled *P. caudatum* *hb* cDNA (Yamauchi et al., 1992a) under highly stringent conditions. The insert DNA of the longest clone from λ EMBL3 library was digested with *Eco*RI or *Bam*HI. The fragment containing the *hb* gene was identified by Southern blot hybridization as described below. Genomic DNA fragments identified by plaque hybridization or PCR amplification were recloned into either pUC19 or pBluescript II SK(-). DNA sequences were determined for both strands by the method of Sanger et al. (1982) with a BcaBEST dideoxy sequencing kit (Takara, Kyoto, Japan).

Southern blotting DNA was fractionated by electrophoresis on 1% agarose gels. The separated DNA was

transferred to nylon membranes after treatment with 50 mM HCl for 15 min, 0.5 N NaOH and 1.5 M NaCl for 15 min and then 30 min, and 0.5 M Tris, pH 7.5, and 3.0 M NaCl for 15 min and then 30 min. The nylon membranes were then hybridized with the *hb* cDNA or genomic DNA corresponding to the *hb* and *nap-1* intergenic region in 50% formamide, 10 \times Denhardt's solution, 4 \times SSC (1 \times SSC: 150 mM NaCl, 15 mM sodium citrate, pH 7.0), 100 μ g/ml salmon sperm DNA, and 50 mM Hepes,

Table 1. Oligonucleotides for PCR primers

Name	Oligonucleotide
L1	5'-ACATCAAACAGAAGAAAATAGCAGG-3'
L2	5'-CTCAGAGTATTCATCAGAATAGG-3'
L3	5'-GCAAGCTTGGGATTCATTTCTAGGATAAGG-3'
R1	5'-ATGATCCTCTACTTTTGGTTGCTG-3'
R2	5'-CAGGAGATTTATCAGTAGGAGAGGCAATTC-3'
R3	5'-CCGAATTCTGACGATATTGCACCAGCC-3'

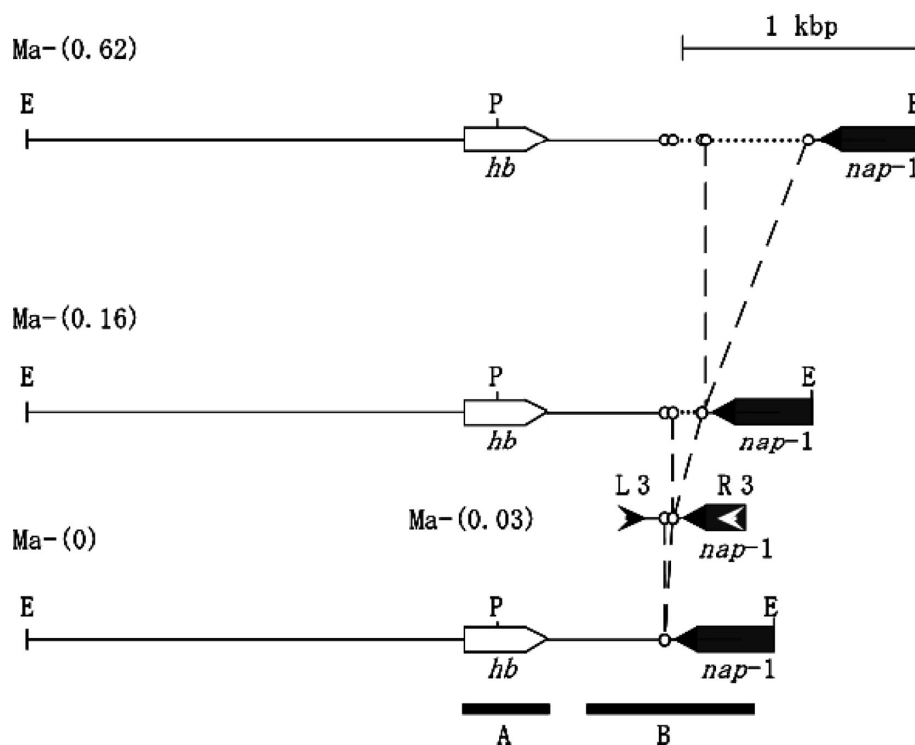


Fig. 1. *P. caudatum* macronuclear genomic DNA fragments cloned around the *hb* and *nap-1* genes. Cloned DNA fragments, Ma-(0.62), Ma-(0.16) and Ma-(0), are indicated by thin horizontal lines with the *hb* and *nap-1* genes (boxed arrows) opposing to each other. The DNA fragment amplified by PCR using the primer set L3/R3, Ma-(0.03), is indicated by a thin horizontal line, a boxed arrow, and two arrowheads which indicate the L3 and R3 primer positions. The Southern hybridization probes are indicated by bold horizontal bars (probes A and B). Missing sequence segments and their boundaries are indicated by dotted lines and open circles, respectively. The nucleotide sequences of the intergenic region between *hb* and *nap-1* were submitted to the DDBJ/EMBL/GenBank under the accession numbers AB107639 for Ma-(0.62), AB107640 for Ma-(0.16) AB107643 for Ma-(0.03) and AB107641 for Ma-(0). Vertical broken lines show the correspondence among the clones, at the putative excised and rejoining sites in the shorter clones. P and E indicate *Pvu*II and *Eco*RI restriction sites, respectively.

pH 7.0 at 42°C (Sambrook et al., 1989). The membranes were washed three times with 0.1% SDS and $2 \times$ SSC for 5 min, and then once with 0.1% SDS and $0.1 \times$ SSC for 5 min. The membranes were air-dried and exposed to XAR5 film (Kodak, Rochester, N.Y., USA) at -85°C with Dupont Cronex Lighting-Plus intensifying screens for 1 to 7 days.

PCR amplification PCR reactions (20–50 μl) contained 100–200 ng of total genomic DNA, 0.4–2.0 μM of each primer, 0.2 mM dNTPs, 1.5 mM MgCl_2 , 10 mM Tris, pH 8.2, 50 mM KCl, 0.001% gelatin, and 1–2 U Ampli Taq Gold DNA polymerase (Perkin-Elmer Applied Biosystems, Foster, CA, USA). Amplification was for 35–55 cycles: denaturation at 95°C for 30 s, annealing at 52°C for 1 min, and extension at 72°C for 1 min, followed by a final

extension at 72°C for 12 min. PCR amplification was performed in thermocyclers (GeneAmp PCR System 2400 and 9700, Perkin-Elmer Applied Biosystems). The primers used are shown in Table 1. Their positions within the *P. caudatum* macronuclear DNA are shown in Figs. 1, 4B and 5B.

Single cells, from which DNA was extracted for PCR, were isolated from stock G3YC-4 in logarithmic phase of vegetative growth and were washed three times with distilled water. A single cell was mixed with 5 μl of buffer containing 20 mM DTT, 1.75 μM SDS, 0.05 mg/ml proteinase K, 1.5 mM MgCl_2 , 10 mM Tris-HCl, pH 8.2, 50 mM KCl, and 0.001% gelatin at 37°C for 1 hr. To inactivate proteinase K, the mixture was boiled for 5 min and then chilled on ice for 5 min. This mixture was used immediately in PCR (Li et al., 1988).

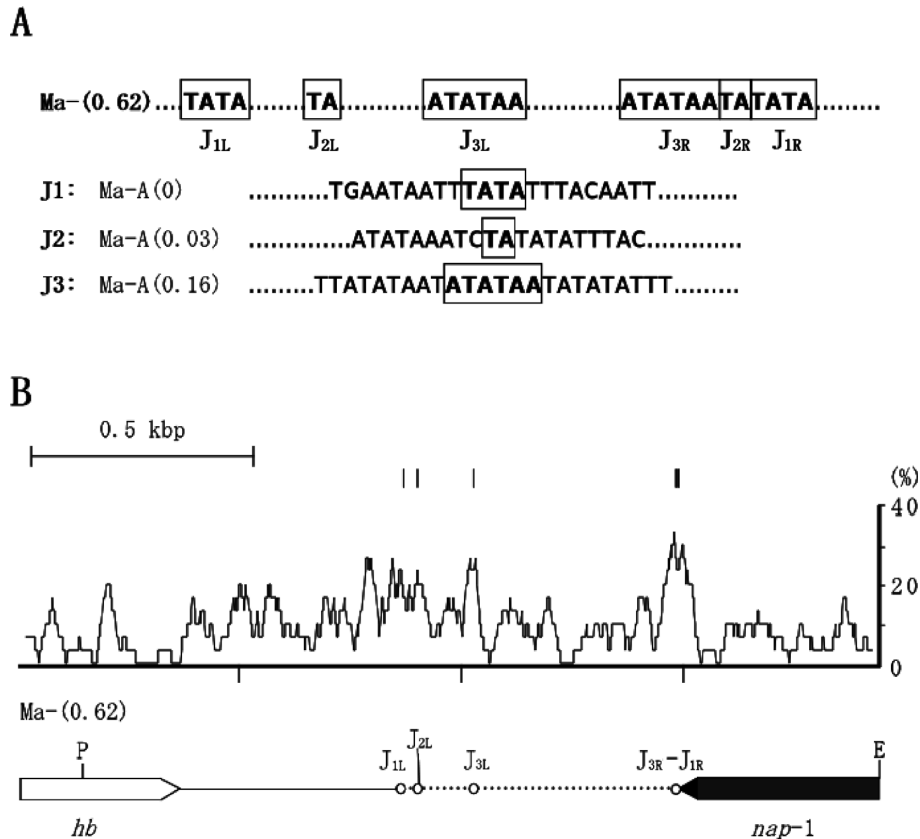


Fig. 2. Direct repeats at boundaries of missing sequence segments and local density of TA dinucleotides in Ma-(0.62). (A) Direct repeats present at boundaries of the missing sequence segment. When compared with the sequences of the shorter clones, Ma-(0), Ma-(0.03) and Ma-(0.16), missing sequence segments are present in Ma-(0.62). Putative deletion junctions (J_{1L} - J_{3L} and J_{3R} - J_{1R}) and direct repeats at their boundaries (boxed) are aligned. One of direct repeats that is maintained in the shorter macronuclear sequences is in boxed bold letters. (B) Local density of TA dinucleotides. Local density of TA dinucleotides was plotted from the *hb* gene (left) through the intergenic region to the *nap-1* gene (right) (percentage of total dinucleotides computed in a 30 bp window). The short vertical lines above the graph mark the positions of the boundaries of the missing sequence segments. Under the plot, Ma-(0.62) are indicated by thin horizontal lines with the *hb* and *nap-1* genes (boxed arrows) opposing to each other and the junctions (open circles) of the missing sequence segments (dotted line).

RESULTS

Sequence analysis of the heterogeneous macronuclear DNA fragments of the intergenic region between the *hb* and *nap-1* genes

The *hb* gene is located in a 15 kbp *Bam*HI fragment or 3–5 kbp *Eco*RI fragments obtained from stock YC genomic DNA (Yamauchi et al., 1992b). Partial genomic libraries in λ EMBL3 and λ gt11 were prepared and screened with a 0.4 kbp cDNA encoding the major *hb* gene (Yamauchi et al., 1992a). The longest 15 kbp fragment was obtained from the λ EMBL3 library. *Eco*RI sub-fragments of this DNA fragment were recloned into pUC19. A 3.2 kbp *Eco*RI fragment containing the *hb* gene was selected by Southern blotting (Yamauchi et al., 1992b). The *hb* cDNA did not hybridize to the other *Eco*RI sub-fragments. This positive 3.2 kbp sub-fragment was named Ma-(0). Clones Ma-(0.16) and Ma-(0.62), 3.4 kbp and 3.8 kbp *Eco*RI fragments, respectively, were isolated from the λ gt11 library. An additional clone, Ma-(0.03) was isolated by PCR amplification using the primer set L3/R3. Differences in the restriction maps among these clones exist in the intergenic region only (Fig. 1).

Nucleotide sequencing found no difference in the *hb* gene, which included one short intron, of the three clones, Ma-(0), Ma-(0.16), and Ma-(0.62). The terminus of the three genomic clones that encodes the C-terminal region of Nap-1 (Nishiyama et al., 2001) also had the same sequence (Fig. 1). Differences in the nucleotide sequences among the four cloned fragments were found in the intergenic region between the *hb* and *nap-1* genes. Ma-(0.03), Ma-(0.16), and Ma-(0.62) contained 0.03 kbp, 0.16 kbp, and 0.62 kbp of additional sequence, respectively, in this region; however, Ma-(0) did not. In the nucleotide sequences, there was only a 32 bp difference in size between Ma-(0) and Ma-(0.03).

When compared with the Ma-(0.62) sequence, the sequence segments missing in Ma-(0), Ma-(0.03) and Ma-(0.16) had high A + T content (73.9–75.1% vs. 52% in the *hb* coding region and vs. 69% in the *nap-1* coding region) and direct 2–6 bp repeats at their boundaries with only one copy of these repeats preserved in the shorter clones (Fig. 2A, in boxed bold letters). The sequence completely conserved in these direct repeats was 5'-TA-3' alone. These features were discovered on both IIDSs and IESs (Prescott, 1994; Le Mouël et al., 2003). The sizes of the missing sequences, estimated from the isolated clones, ranged from 32 bp to 616 bp.

A plot of the local density of TA dinucleotides, computed in a 30-bp window sliding along the sequence, of Ma-(0.62) demonstrated that the boundaries of the sequences missing in the shorter clones coincided with peaks in TA density (Fig. 2B). The right boundaries (J_{1R} to J_{3R}) were clustered within a 10-bp region with the highest TA density (~30%), whereas the left boundaries were (J_{1L} to J_{3L})

interspersed in the intergenic region of 158 bp with secondary higher TA density (17.8–22.3%). Coincidence of missing segment boundaries and peaks of local TA density was one of the hallmarks of IIDSs (Le Mouël et al., 2003).

As the other hallmark of IIDSs, Cbs existed in the vicinity of the missing sequence (Prescott, 1994). To ascertain the existence of Cbs, we surveyed whole sequence of Ma-(0.62). Any consensus Cbs, which was identified in *Euplotes* and *Tetrahymena* (Yao et al., 1987; Klobutcher et al., 1998), was not detected as well as previous studies in *Paramecium* (Baroin et al., 1987; Coyne et al., 1996; Le Mouël et al., 2003).

Heterogeneous composition of the intergenic region between the *hb* and *nap-1* genes in macronuclear genomic DNA

Southern blotting of total genomic DNA obtained from stock YC with the 0.4 kbp *hb* cDNA (probe A, Fig. 1) revealed three strong positive bands of 3.2 kbp, 3.7 kbp and 4.8 kbp (lane 1, Fig. 3), as previously shown (Yamauchi et al., 1992b). The 3.2 kbp band corresponded to Ma-(0)/Ma-(0.03); however, the 3.7 kbp band did not correspond to Ma-(0.62). The fact that the 3.2 kbp band is a major fragment in the heterogeneous *hb* macronuclear genomic DNA was confirmed by PCR using the primer set L3/R3. The size of the most intense band was estimated to be 0.4–0.6 kbp (indicated by an open dot, lane 1, Fig. 4A). This agreed with the distance between

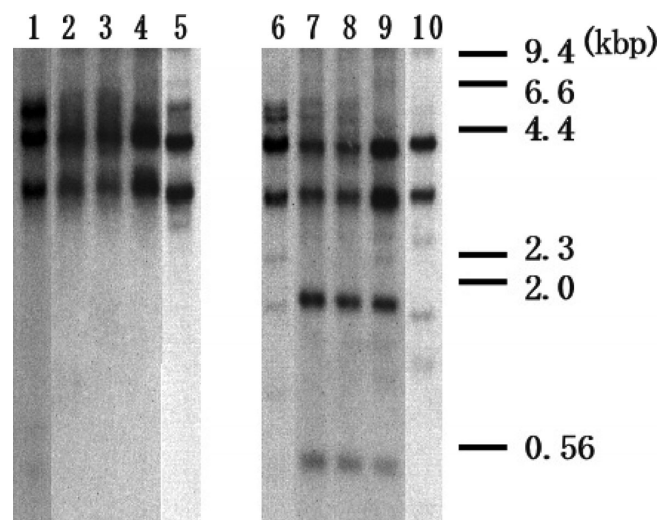


Fig. 3. Southern blotting of the *P. caudatum* genomic DNA. Genomic DNA was extracted from stocks YC (lanes 1 and 6) and 27aG3 (lanes 5 and 10), and their offspring, G3YC-1 (lanes 2 and 7), G3YC-2 (lanes 3 and 8) and G3YC-4 (lanes 4 and 9), each of which is an exconjugant clone. Each sample (6.7 μ g/lane) was digested with *Eco*RI, and was electrophoresed on a 1% agarose gel. DNA, immobilized on a nylon membrane, was hybridized with the *hb* cDNA (probe A in Fig. 1; lanes 1–5), and with both flanking regions of the missing sequence segments in the macronuclear genome (probe B in Fig. 1; lanes 6–10).

the primer sites L3 and R3 calculated for Ma-(0)/Ma-(0.03), 420/452 bp and for Ma-(0.16), 581 bp (indicated by open dot, Fig. 4B). Using the L3 and R3 primers, band corresponding to Ma-(0.62) should be 1036 bp. This band was visible and clearly distinct yet (indicated by a closed dot, in lane 1, Fig. 4A), but it was not as intense as the 0.4–0.6 kbp band.

When a probe specific for the intergenic regions (probe B, Fig. 1) was used in Southern blotting, a result similar to that of the *hb* cDNA (lane 6, Fig. 3), except for additional weak bands of approximately 1.7 kbp, 2.2 kbp and 5.2 kbp, was obtained. PCR amplifications using the primer sets L3/R1, L2/R1, and L1/R1 revealed several minor genomic fragments longer than the 740 bp, 524 bp,

and 378 bp bands (indicated by a closed dot, lanes 4, 7, and 10, respectively, in Fig. 4A), expected for Ma-(0.62). Using the primer set L1/R2, bands ranging 1–2 kbp were found in addition to the 0.6 kbp band expected for Ma-(0.62) (indicated by a closed dot, lane 13 in Fig. 4A). These results, when combined with the sequencing data, indicated that the intergenic region between the *hb* and *nap-1* genes in the macronuclear genome of *P. caudatum* (stock YC) is heterogeneously composed of three major frequently occurring segments and several minor segments.

Comparison of the *hb* and *nap-1* intergenic region between stocks, clones, and cells in the same clone To determine if this heterogeneity in the macronuclear genome was specific to the *P. caudatum* stock YC

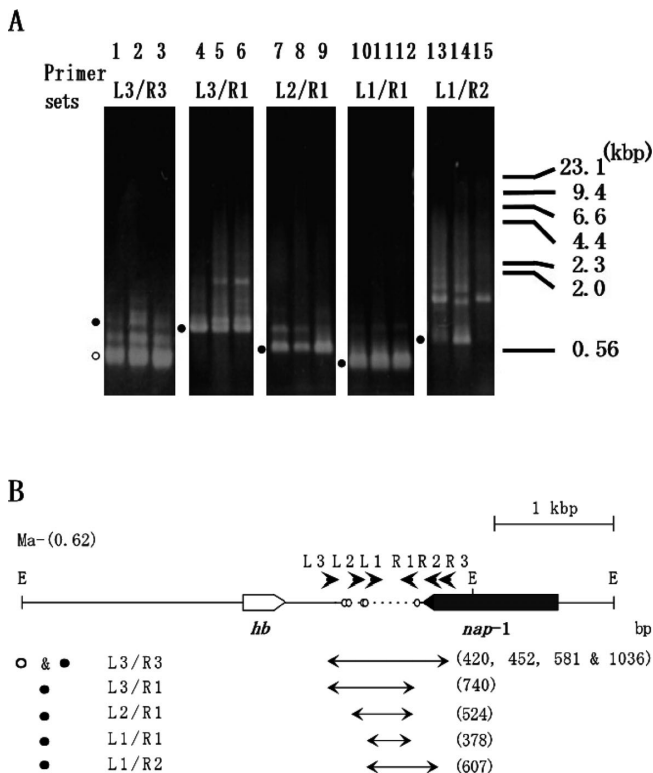


Fig. 4. Heterogeneity in the *P. caudatum* macronuclear genomic DNA. (A) PCR amplification of total genomic DNA. Genomic DNA from stock YC (lanes 1, 4, 7, 10 and 13), G3YC-4 (lanes 2, 5, 8, 11 and 14) and 27aG3 (lanes 3, 6, 9, 12 and 15) were used as templates for PCR amplification using the specific primer sets, L3/R3 (lanes 1–3), L3/R1 (lanes 4–6), L2/R1 (lanes 7–9), L1/R1 (lanes 10–12) and L1/R2 (lanes 13–15). *Hind*III-digested λ DNA (kbp) sizes are marked on the right side. The bands corresponding to the Ma-(0), Ma-(0.03) and Ma-(0.16) and that corresponding to the Ma-(0.62) PCR products are indicated by an open dot and a closed dot, respectively. (B) Cloned Ma-(0.62) with PCR primer positions. PCR primer sites are indicated by arrowheads (L1, L2, L3, R1, R2 and R3). Regions amplified by PCR using specific primer sets are indicated by horizontal arrows. The PCR product sizes assumed from the isolated clones are given. Missing sequence segments in the shorter clones and their boundaries are indicated by dotted lines and open circles, respectively.

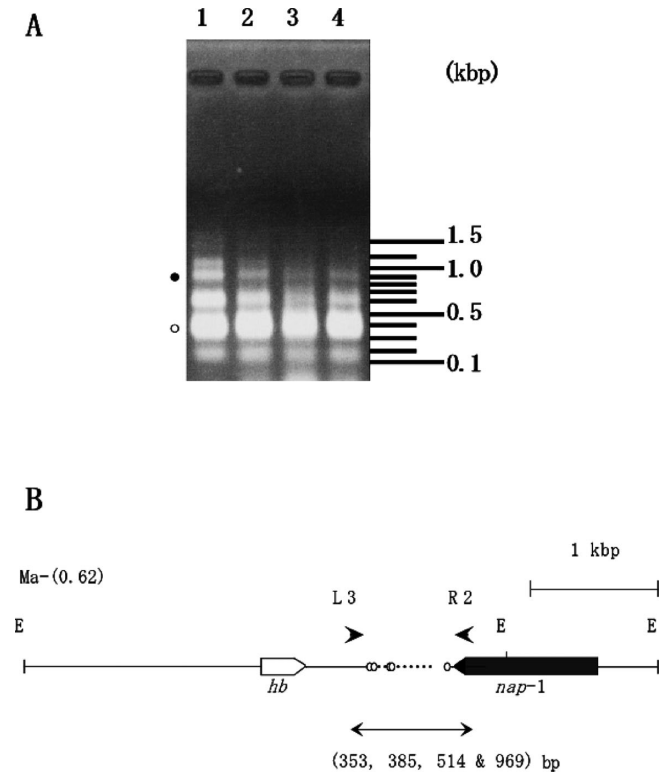


Fig. 5. PCR amplification of *P. caudatum* total genomic DNA isolated from a single cell. (A) Genomic DNA isolated from a single cell of stock G3YC-4 was used as a template for PCR amplification using the primer set L3/R2 (lanes 1–4). The expected sizes of the amplified DNA are indicated by a closed dot: 969 bp, estimated from the Ma-(0.62) sequence, and by an open dot 353–514 bp, estimated from the Ma-(0), Ma-(0.03) and Ma-(0.16) sequences. Size markers for a 100 bp ladder are indicated on the right. (B) Cloned Ma-(0.62) with PCR primer positions. The PCR primer sites are indicated by arrowheads (L3 and R2). Missing sequence segment and possible splice sites are indicated by dotted lines and open circles, respectively. Region amplified by PCR using specific primer set is indicated by a horizontal arrow. The PCR product sizes assumed for the isolated clones are indicated.

or a universal feature of most *P. caudatum* stocks, the composition of the *hb* and *nap-1* intergenic region in offspring obtained from the conjugation of stock YC with stock 27aG3 was determined. Southern blot (Fig. 3) and PCR analyses (Fig. 4) revealed few differences in the composition of the intergenic regions of stock YC and stock 27aG3, except for the bands of 4.8 kbp and 5.2 kbp on Southern blot (lanes 6 and 10, Fig. 3). Furthermore, there was no significant difference in the heterogeneity of the intergenic region in three offspring exconjugant clones (lanes 2–4 and lanes 7–9, Fig. 3). Southern blot analysis revealed significant differences in 1.7 kbp, 4.8 kbp and 5.2 kbp bands between the offspring and their parents when probe B was used (lanes 6–10, Fig. 3), however PCR analysis indicated few differences between them (Fig. 4A).

To determine if each cell derived from an exconjugant clone had the same or an individually distinct heterogeneous composition in their macronuclear DNA, single cell PCR using the primer set L3/R2 (Fig. 5B) was performed. At least five bands were detected (Fig. 5A). The size and proportion of these bands when compared among the four different cells were almost identical. The second longest band (a closed dot in Fig. 5A) corresponded to the Ma-(0.62) PCR product, 969 bp in length (Fig. 5B). The fourth longest and prominently amplified band (an open dot in Fig. 5A) spanning from 0.35 kbp to 0.5 kbp corresponded to the Ma-(0), Ma-(0.03) and Ma-(0.16) PCR products, 353–514 bp in length (Fig. 5B). Therefore, the heterogeneous composition of the *P. caudatum* macronuclear DNA is retained within a cell by a strictly controlled mechanism accomplishing a constant composition and proportion.

DISCUSSION

Characteristics of *P. caudatum* eliminated sequences Partially overlapping and at least three different sequence segments in size were missing in the intergenic region between the *hb* and *nap-1* genes in the *P. caudatum* macronucleus. The missing sequences had characteristics common to both TA-IESs (Steele et al., 1994) and IIDSs: direct repeats at the boundaries and high A + T content. We did not detect any microsatellites and coding regions for transposase which are found in IIDSs, in our missing sequences segments. However, the missing sequences had a characteristic unique to IIDSs that are imprecisely deleted at their boundaries with high TA density. It is highly likely that extreme frequently heterogeneous composition of the *P. caudatum* macronuclear genomic DNA between the *hb* and *nap-1* genes is generated by eliminating intergenic regions of various sizes, by a mechanism that is similar but somewhat differs from those for the elimination of TA-IESs and IIDSs.

Southern blot and PCR analyses revealed that Ma-(0.62) was not a clone corresponding to micronuclear genomic DNA: the intergenic region of the macronuclear DNAs included eliminated sequences which were longer than the eliminated 0.62 kbp sequence of Ma-(0.62) (Fig. 4A). There could be a dozen eliminated sequences of various sizes in this region of the macronuclear genome. A 32 bp difference in size between Ma-(0) and Ma-(0.03) was found. This micro-heterogeneity might be the result of a variable and reciprocal recombination process after a staggered cut of the eliminated sequence (Saveliev and Cox, 1994; Yao and Yao, 1994), an alternative excision of the eliminated sequence (Dubrana et al., 1997) or an imprecise excision of the eliminated sequence, which might occur upon random recognition of the terminal repeats 5'-TA-3'. Another possibility includes generation of Ma-(0) from Ma-(0.03), by removal of a short eliminated sequence, like the 9 bp IES in *O. nova* (Dubois and Prescott, 1997). The presence of short IESs within a longer IES might be explained by reiterative integration of IES elements (Mayer et al., 1998).

Single cell PCR demonstrated that each cell within the same clone had the same heterogeneity in the intergenic region during vegetative growth (Fig. 5A). Three exconjugants of the offspring from stock YC had almost the same patterns in Southern blot and PCR analysis (Figs. 3 and 4). Southern blotting of total genomic DNAs isolated from several pairs of karyonidal clones was performed. Their Southern blotting patterns were identical (data not shown). These results imply that at least three major fragments containing the *hb* gene are reproducibly generated independent of cytoplasmic factors. These karyonide clones were immature (within 50 fissions after the latest conjugation); therefore, this heterogeneity might not result from heterogeneous senescent populations. In some instances, IES excision from the *P. tetraurelia* genome is controlled by the old macronucleus during new macronuclear development (Forney, 1997). Our data suggest that the heterogeneous composition of the macronuclear DNA is strictly controlled and is constant during vegetative growth and among exconjugant and karyonidal clones. The heterogeneity found in *P. caudatum* is different from that reported in *P. tetraurelia*. The heterogeneity found in the excision patterns of *P. tetraurelia* IESs is generated by the point mutations in the flanking 5'-TA-3' dinucleotide: one mutation prevents excision of an IES (Mayer and Forney, 1999), and another mutation prevents normal recognition of 5'-TA-3' dinucleotide by the splicing machinery, generating a few variation of elimination patterns (Dubrana et al., 1997). However, the extreme and reproducible heterogeneity found in the intergenic region of the *P. caudatum* macronuclear genomic DNA might be generated by a mechanism distinct from the point mutations found in *P. tetraurelia* IESs.

We thank H. Tada, T. Honda, K. Kurokawa, Y. Iwataki and S. Kamii for their help in sequencing the isolated DNAs and in preparing figures. We would like to thank Dr. J. Monk for a thorough and critical reading of the manuscript.

REFERENCES

- Antoine, M., and Niessing, J. (1984) Intron-less globin genes in the insect *Chironomus thummi thummi*. *Nature* **310**, 795–798.
- Austerberry, C. F., Allis, C. D., and Yao, M. C. (1984) Specific DNA rearrangements in synchronously developing nuclei of *Tetrahymena*. *Proc. Natl. Acad. Sci. USA* **81**, 7383–7387.
- Baroin, A., Prat, A., and Caron, F. (1987) Telomeric site position heterogeneity in macronuclear DNA of *Paramecium primaurelia*. *Nucleic Acids Res.* **15**, 1717–1728.
- Caron, F. (1992) A high degree of macronuclear chromosome polymorphism is generated by variable DNA rearrangements in *Paramecium primaurelia* during macronuclear differentiation. *J. Mol. Biol.* **225**, 661–678.
- Coyne, R. S., Chalker, D. L., and Yao, M. C. (1996) Genome downsizing during ciliate development: nuclear division of labor through chromosome restructuring. *Annu. Rev. Genet.* **30**, 557–578.
- Deak, J. C., and Doerder, F. P. (1998) High frequency intergenic recombination during macronuclear development in *Tetrahymena thermophila* restores the wild-type *SerH1* gene. *Genetics* **148**, 1109–1115.
- Doak, T. G., Doerder, F. P., Jahn, C. L., and Herrick, G. (1994) A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common “D35E” motif. *Proc. Natl. Acad. Sci. USA* **91**, 942–946.
- Dryl, S. (1959) Antigenic transformation in *Paramecium aurelia* after homologous antiserum treatment during autogamy and conjugation. *J. Protozool.* **6**, 25.
- Dubois, M. L., and Prescott, D. M. (1997) Volatility of internal eliminated segments in germ line genes of hypotrichous ciliates. *Mol. Cell. Biol.* **17**, 326–337.
- Dubrana, K., Le Mouël, A., and Amar, L. (1997) Deletion endpoint allele-specificity in the developmentally regulated elimination of an internal sequence (IES) in *Paramecium*. *Nucleic Acids Res.* **25**, 2448–2454.
- Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O’Connell, C., Spritz, R., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., et al. (1980) The structure and evolution of the human β -globin gene family. *Cell* **21**, 653–668.
- Forney, J. (1997) DNA rearrangements and mating-type determination in *Paramecium tetraurelia*. *BioEssays* **19**, 5–8.
- Grandchamp, S., and Beisson, J. (1981) Positional control of nuclear differentiation in *Paramecium*. *Dev. Biol.* **81**, 336–341.
- Hardison, R. (1998) Hemoglobins from bacteria to man: evolution of different patterns of gene expression. *J. Exp. Biol.* **201**, 1099–1117.
- Herrick, G., Hunter, D., Williams, K., and Kotter, K. (1987) Alternative processing during development of a macronuclear chromosome family in *Oxytricha fallax*. *Genes Dev.* **1**, 1047–1058.
- Jaraczewski, J. W., and Jahn, C. L. (1993) Elimination of *Tec* elements involves a novel excision process. *Genes Dev.* **7**, 95–105.
- Klobutcher, L. A., Gygas, S. E., Podoloff, J. D., Vermeesch, J. R., Price, C. M., Tebeau, C. M., and Jahn, C. L. (1998) Conserved DNA sequences adjacent to chromosome fragmentation and telomere addition sites in *Euplotes crassus*. *Nucleic Acids Res.* **26**, 4230–4240.
- Klobutcher, L. A., and Herrick, G. (1995) Consensus inverted terminal repeat sequence of *Paramecium* IESs: resemblance to termini of *Tc1*-related and *Euplotes Tec* transposons. *Nucleic Acids Res.* **23**, 2006–2013.
- Klobutcher, L. A., and Herrick, G. (1997) Developmental genome reorganization in ciliated protozoa: the transposon link. *Prog. Nucleic Acid Res. Mol. Biol.* **56**, 1–62.
- Klobutcher, L. A., and Jahn, C. L. (1991) Developmentally controlled genomic rearrangements in ciliated protozoa. *Curr. Opin. Gen. Dev.* **1**, 397–403.
- Kowalczyk, C. A., Anderson, A. M., Arce-Larreta, M., and Chalker, D. L. (2006) The germ line limited M element of *Tetrahymena* is targeted for elimination from the somatic genome by a homology-dependent mechanism. *Nucleic Acids Res.* **34**, 5778–5789.
- Le Mouël, A., Butler, A., Caron, F., and Meyer, E. (2003) Developmentally regulated chromosome fragmentation linked to imprecise elimination of repeated sequences in paramecia. *Eukaryot. Cell* **2**, 1076–1090.
- Li, H. H., Gyllensten, U. B., Cui, X. F., Saiki, R. K., Erlich, H. A., and Arnheim, N. (1988) Amplification and analysis of DNA sequences in single human sperm and diploid cells. *Nature* **335**, 414–417.
- Mayer, K. M., and Forney, J. D. (1999) A mutation in the flanking 5’-TA-3’ dinucleotide prevents excision of an internal eliminated sequence from the *Paramecium tetraurelia* genome. *Genetics* **151**, 597–604.
- Mayer, K. M., Mikami, K., and Forney, J. D. (1998) A mutation in *Paramecium tetraurelia* reveals functional and structural features of developmentally excised DNA elements. *Genetics* **148**, 139–149.
- Mikami, K. (1980) Differentiation of somatic and germinal nuclei correlated with intracellular localization in *Paramecium caudatum* exconjugants. *Dev. Biol.* **80**, 46–55.
- Mikami, K. (1987) Macronuclear development and gene expression in the exconjugants of *Paramecium caudatum*. *Dev. Biol.* **123**, 161–168.
- Mikami, K., and Hiwatashi, K. (1975) Macronuclear regeneration and cell division in *Paramecium caudatum*. *J. Protozool.* **22**, 536–540.
- Mochizuki, K., and Gorovsky, M. A. (2004) Small RNAs in genome rearrangement in *Tetrahymena*. *Curr. Opin. Genet. Dev.* **14**, 181–187.
- Nishiyama, N., Sawatsubashi, S., Ishida, M., and Yamauchi, K. (2001) Organization and expression of the *Paramecium caudatum* gene encoding nucleosome assembly protein 1. *Gene* **280**, 107–114.
- Prescott, D. M. (1994) The DNA of ciliated protozoa. *Microbiol. Rev.* **58**, 233–267.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd ed. pp. 9.31–9.62, Cold Spring Harbor Laboratory Press, New York.
- Sanger, F., Coulson, A. R., Hong, G. F., Hill, D. F., and Petersen, G. B. (1982) Nucleotide sequence of bacteriophage λ DNA. *J. Mol. Biol.* **162**, 729–773.
- Saveliev, S. V., and Cox, M. M. (1994) The fate of deleted DNA produced during programmed genomic deletion events in *Tetrahymena thermophila*. *Nucleic Acids Res.* **22**, 5695–5701.
- Steele, C. J., Barkocy-Gallagher, G. A., Preer, L. B., and Preer, J. B. (1994) Developmentally excised sequences in micronuclear DNA of *Paramecium*. *Proc. Natl. Acad. Sci. USA* **91**,

- 2255–2259.
- Usuki, I., Hino, A., and Ochiai, T. (1989) Reinvestigation of the hemoglobins from *P. caudatum*. *Comp. Biochem. Physiol.* **93B**, 555–559.
- Yamauchi, K., Mukai, M., Ochiai, T., and Usuki, I. (1992a) Molecular cloning of the cDNA for the major hemoglobin component from *Paramecium caudatum*. *Biochem. Biophys. Res. Commun.* **182**, 195–200.
- Yamauchi, K., Ochiai, T., and Usuki, I. (1992b) The unique structure of the *Paramecium caudatum* hemoglobin gene: the presence of one intron in the middle of the coding region. *Biochim. Biophys. Acta* **1171**, 81–87.
- Yao, M. C., and Yao, C. H. (1994) Detection of circular excised DNA deletion elements in *Tetrahymena thermophila* during development. *Nucleic Acids Res.* **22**, 5702–5708.
- Yao, M. C., Zheng, K., and Yao, C. H. (1987) A conserved nucleotide sequence at the sites of developmentally regulated chromosomal breakage in *Tetrahymena*. *Cell* **48**, 779–788.