Volume 13 Number 2 1985 Nucleic Acids Research

Nucleotide sequence of a macronuclear DNA molecule coding for α-tubulin from the ciliate *Stylonychia lemnae*. Special codon usage: TAA is not a translation termination codon

E.Helftenbein

Universitaet Tuebingen, Biologie III, Abt. Zellbiologie Auf der Morgenstelle 28, 7400 Tuebingen, FRG

Received 10 October 1984; Revised 10 December 1984; Accepted 3 January 1985

ABSTRACT

The gene-sized macronuclear DNA of the hypotrichous ciliate Stylonychia lemnae contains two size classes of DNA molecules (1.85 and 1.73 kbp) coding for α-tubulin. Each macronucleus contains about 55000 copies of the 1.85 kbp molecules and about 17000 copies of the 1.73 kbp DNA molecules. Five macronuclear molecules of these sequences were cloned and sequenced, one, from the 1.85 kbp size class in its entirety. The 5 sequences fell into two classes suggesting that Stylonychia lemnae contains at least two different α-tubulin genes. All 5 clones show the codon TAA in the same nucleotide positions of the coding region. In this position the TAA codon cannot function as a translational stop codon and we suggest that this codon codes for the amino acid glutamine. The nucleotide sequence of the coding region as well as the encoded amino acid sequence is highly conserved compared to α-tubulin genes from vertebrates. The noncoding regions show several putative transcription-regulatory sequences as well as sequences presumably functioning as replication origins.

INTRODUCTION

Stylonychia lemnae shows the nuclear dualism typical for ciliated protozoa. Cells contain two types of nuclei, small diploid micronuclei resembling gametic nuclei of higher eukaryotes and the DNA-rich macronuclei (1, 2). Micronuclei contain high molecular weight DNA (chromosome-sized DNA) in form of about 150 chromosomes per haploid genome (2). Macronuclear DNA is organized in short molecules, 0.4 to 20 kbp in size (2, 3). In the course of sexual reproduction (conjugation) the new macronucleus is formed from a micronucleus in a process which involves polytene chromosome formation, degradation of these chromosomes accompanied by the loss of over 90 % of the DNA sequences, and the specific fragmentation of the DNA into gene-sized DNA molecules (2, 4, 5). Macronuclear DNA contains 10000 - 20000 different DNA molecules each present in average 15000 times per macronucleus (2, 6, 7, 8). Hybridization experiments of total macronuclear DNA with cloned macronuclear DNA fragments showed that the amplification

differs from fragment to fragment. Copy numbers vary from 3000 to more than 100000 per nucleus (unpublished data). In contrast to micronuclei, macronuclei do not divide by a mitotic process but the copy number of each fragment are maintained constant through several subsequent cell divisions (measured up to 200 cell division, unpublished). Thus each macronuclear DNA fragment seems to be a replication unit with one or two replication origins together with all essential control sequences. Preliminary electron microscopic observations suggest that replication origins are at or very close to one or both ends of each DNA molecule (9, Steinbrück, unpublished).
Only macronuclei are transcriptionally active during vegetative cell growth (1). Hybridization of specific gene probes (e. g. rDNA, histone-, actin-, tubulin-genes) to macronuclear DNA revealed hybridization to defined macronuclear DNA molecules (3, 6, 10). These results together with the ability to obtain stable transformants of Stylonychia, when macronuclear DNA sequences were used as a vector (11) lead to the assumption, that each macronuclear DNA molecule is a complete transcription and replication unit. Thus, all signals and regulatory sequences involved in inititation of transcription such as the "Hogness" box (12, 13) and the transcriptional termination signals AATAAA (14,15) and TTTT (16) as well as the replication origins and other essential regulatory sequences should be present on each macronuclear DNA molecule.
Sequence analysis of total macronuclear DNA from several hypotrichous ciliates revealed the existence of terminal inverted repeats present at both ends of each DNA molecule (17,18,19). The sequence of these terminal inverted repeats in Stylonychia is 5'-$C_4A_4C_4A_4C_4$ with an additional single strand extension of 3'-$G_4T_4G_4T_4$ (19,unpubl. data). Similar tandemly repeated sequences have been found in other lower eukaryotes where they function presumably as telomeric structures in the DNA replication process (20).
Since ciliates show a variety of microtubuli-containing structures, e. g. ciliary axonemes, subpellicular microtubular systems and the spindle apparatus of the dividing micronuclei, (21,22) we have chosen the macronuclear DNA molecules containing the α-tubulin genes for detailed analysis. Of particular interests are gene multiplicity, DNA replication and its regulation on the nucleotide sequence level as well as regulatory sequences involved in tubulin synthesis. Therefore macronuclear DNA molecules of Stylonychia lemnae coding for α-tubulin were cloned and sequenced.

MATERIAL AND METHODS

Cultivation of cells and isolation of macronuclear DNA, nuclear and polysomal RNA

Cells of Stylonychia lemnae were cultivated in Pringsheim solution with the alga Chlorogonium elongatum as food (23), and after isolation of macronuclei, DNA was purified as previously described (2, 7). Nuclear RNA was isolated from vegetative cells as described by Nock (8), and purified by cesium chloride gradient centrifugation (24). Polysomal RNA was isolated from purified polysomes (25) following the procedure of Buell et al. (26).

Isolation of plasmid- and recombinant DNA

Plasmid- and recombinant DNA was extracted from bacterial cultures as described previously (27), followed by CsCl/ethidium bromide density gradient centrifugation.

Agarose gel electrophoresis and hybridization reactions

Native and restriction endonuclease digested macronuclear DNA as well as recombinant DNA was subjected to electrophoresis on 1 % to 1.5 % vertical agarose slab gels. Gels containing 1.5 % agarose with 50 % formaldehyde were used for electrophoresis of small restriction fragments of recombinant DNA (28). For denaturing gels RNA and DNA were treated with glyoxal (29), and electrophoresed on 1.5 % agarose gels. After electrophoresis the gels containing DNA were either dried (30), or the DNA was transferred to nitrocellulose membrane filters (Schleicher & Schüll, BA 85) as described by Southern (31). From denaturing gels RNA and DNA were transferred to nitrocellulose membrane filters by the method of Thomas (32). The DNA in the dried gels was denatured in 0.5 M NaOH and 150 mM NaCl and then neutralized in 1.5 M TRIS-HCl pH 8 and 150 mM NaCl for 2 x 15 min. Prehybridization and hybridization was carried out at 62 ° C in 5 x SSC, 10 x Denhard's solution (33) and 0.1 % SDS. For hybridization the labeled probes were denatured in TE-buffer and mixed with prewarmed hybridization solution. After hybridization for 24 - 48 hours the gels and filters were rinsed down to 0.1 x SSC at 58 ° C and exposed to Agfa Curix RP1 X-ray films with Dupont Quanta III intensifying screens at - 70 ° C.

Preparative agarose gel electrophoresis was carried out either on 1.5 % low melting agarose gels (Seakam) or on 1 % and 1.5 % normal agarose gels. DNA was eluted from low melting agarose gels as described previously (34), and from high melting agarose gels by the electrophoretical method of Smith (35).

The molecular weight marker of all gels was λ-DNA (BRL) digested with Hind III and Eco R I.

Restriction endonuclease digestions

Restriction endonucleases were purchased from Boehringer, Mannheim and Bethesda Research Laboratories. Digestions were carried out as suggested by the manufacturers.

Labeling of DNA for hybridization- and sequencing experiments

DNA was labeled by the nick-translation procedure (36) with α-^{32}P dCTP (spec. act. 2000 - 3000 Ci/mMol. Amersham), and used for hybridization experiments. For sequencing the DNA fragments were isolated from agarose gels by electrophoretic elution, phosphatase treated and labeled at the 5' ends with γ-^{32}P ATP (5000 Ci/mMol, Amersham) using T_4 polynucleotide kinase or at the 3' ends with α-^{32}P ddATP (3000 Ci/mMol, Amersham) and terminal deoxynucleotide transferase. All enzymes were purchased from Boehringer, Mannheim and reactions were carried out as recommended by the manufacturers.

Construction of recombinant plasmids containing macronuclear DNA

Hybrid plasmids were constructed with pBR 322 and full length or restriction enzyme fragmented macronuclear DNA. Plasmid DNA was linearized with the restriction endonuclease Pst I and about 20 nucleotides of dG were added to the 3'-ends by terminal deoxynucleotide transferase (Boehringer) following the method of Roychoudhury (37). About the same number of dC residues was added to the macronuclear DNA molecules. Equimolar amounts of tailed pBR 322 and macronuclear DNA were heated to 60 ° for 5 min, reannealed by cooling to 42 ° C for 2 h and maintained at room temperature overnight. Restriction fragments of Sal I-Pvu II digested macronuclear DNA were cloned in the following way: pBR 322-DNA was digested with the restriction enzymes Sal I and Pvu II. The larger fragment was ligated (T_4-Ligase, Boehringer, Mannheim) to the macronuclear DNA-fragments in a molar relation 1 : 5, following the procedures suggested by the manufacturers. Cells of *E. coli* HB 101 and SK 375 were used for transformation as described by Mandel and Higa (38).

Colonies containing recombinant DNA with full length macronuclear DNA were tetracycline resistant and ampicillin sensitive. Colonies consisting of pBR 322 and Sal I-Pvu II restriction fragments of macronuclear DNA showed resistance to ampicillin and sensitivity to tetracyclin. Both kinds of colonies were screened for the desired sequences.

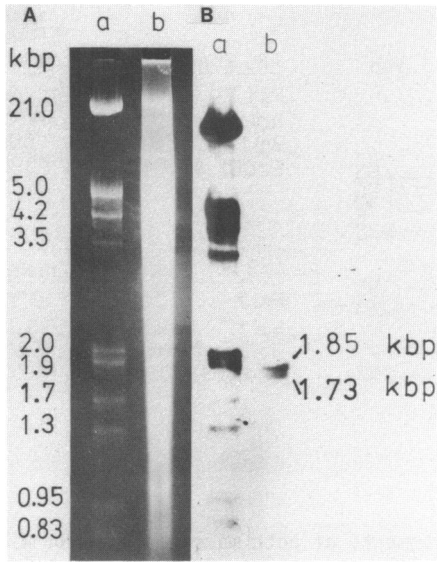


Fig. 1: Detection of macronuclear DNA molecules coding for α-tubulin. A: Total macronuclear DNA was electrophoresed on a 1% agarose gel (12 µg DNA per lane) and stained with ethidium bromide. B: The gel was dried and hybridized with a ^{32}P-labeled heterologous cloned α-tubulin cDNA probe from chicken as described in the "Materials and Methods". Lane a: DNA molecular weight marker (Eco RI, Hind III digested λ-DNA) lane b: macronuclear DNA.

Screening for α-tubulin-gene sequences

Colonies containing recombinant plasmids with full length macronuclear DNA (i. e. clones received by GC-tailing) were screened for the α-tubulin genes with a ^{32}P-labeled DNA probe of cloned cDNA of the α-tubulin mRNA from chicken. The recombinant DNA containing the α-tubulin cDNA was kindly given to us by D. Cleveland (39). Colony-hybridization (40) was used to screen for the α-tubulin genes.

Colonies containing recombinant plasmids with restriction fragments of macronuclear DNA, were screened for α-tubulin gene containing sequences with the ^{32}P-labeled cloned α-tubulin gene form Stylonychia lemnae by the same method.

DNA-sequencing

Sequencing was performed as described by Maxam and Gilbert (41) For the majority of the fragments, sequences were determined on both strands.

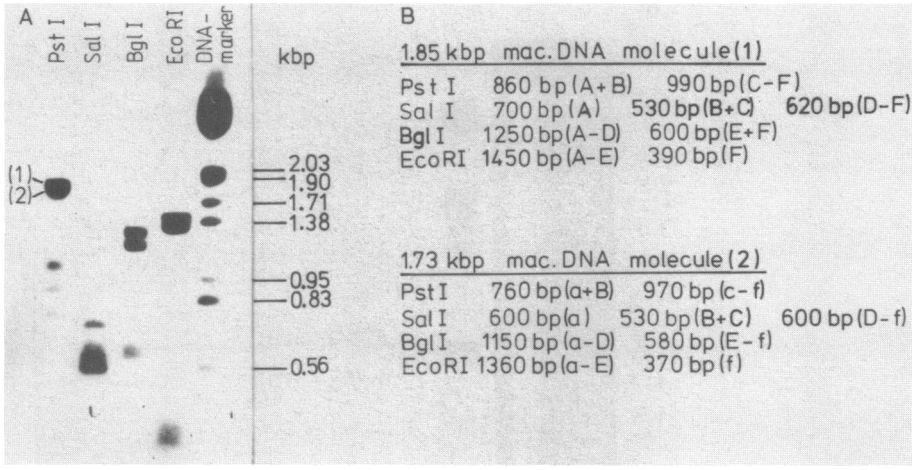


Fig. 2: Restriction fragments of both macronuclear DNA molecules containing the α-tubulin genes. Total macronuclear DNA was digested with Pst I, Sal I, Bgl I and Eco RI and electrophoresed on a 1.5% agarose gel (10 µg DNA per lane). After DNA-transfer to nitrocellulose filters, the filters were hybridized with a ^{32}P-labeled heterologous chicken α-tubulin gene probe.
A: autoradiograph; B: size of restriction fragments of both macronuclear DNA molecules containing the α-tubulin genes. Fragments A - F are shown in Fig. 4.

RESULTS

Identification and cloning of macronuclear DNA molecules coding for α-tubulin

To identify the macronuclear DNA molecules coding for α-tubulin, total macronuclear DNA was separated according to size by gel electrophoresis and hybridized with a heterologous α-tubulin gene probe. For this purpose a cloned cDNA probe from chicken α-tubulin mRNA was used (39). Fig. 1 shows that two different size classes of macronuclear DNA molecules hybridize to the chicken α-tubulin probe, one with a size of 1.85 kbp and one with 1.73 kbp.

In order to clone these fragments macronuclear DNA was electrophoresed on a preparative low melting agarose gel. The region containing the DNA of the corresponding molecular weight was then cut out and the DNA eluted. These DNA molecules were cloned as described in the "Materials and Methods". Bacterial colonies were screened by colony hybridization on nitrocellulose filters with the radioactive labeled heterologous cDNA probe. One positive clone was found, pmac C9.The positive clone was identified by compari-

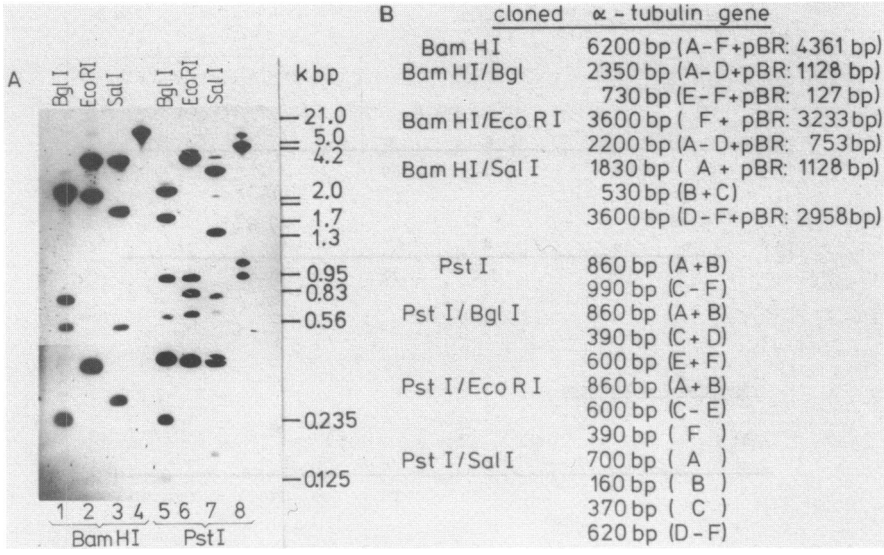


Fig. 3: Restriction fragments of the cloned macronuclear DNA molecule containing a α-tubulin gene. The recombinant DNA was linearized with Bam H I, lane 4, and digested with Bgl I, Eco RI and Sal I, lane 1 - 3. In lane 8 the insert is cut off from the plasmid pBR 322 with a Pst I digestion and further digested with Bgl I, Eco RI and Sal I, lane 5 - 7. All fragments were separated on a 1.5% agarose, 50% formaldehyde gel. Each lane contains 2 µg of recombinant DNA. After blotting to nitrocellulose filters, the filters were hybridized to a ^{32}P- labeled heterologous chicken α-tubulin gene probe. In all lanes the restriction fragments of the plasmid pBR 322 are also seen. A: autoradiograph; B: size of restriction fragments. Fragments consisting of plasmid DNA only are not listed and the fragments A - F are shown in Fig. 4.

son of its restriction map with those of the uncloned genes in macronuclear DNA. Fig. 2 and 3 show the autoradiographs of digested macronuclear DNA and of digested DNA from clone pmac C9. Fig. 4 demonstrates the restriction maps of both size classes containing the α-tubulin genes in the macronuclear DNA and of the cloned DNA molecule. All 3 restriction maps do not differ with respect to the tested restriction sites. The 1.73 kbp macronuclear DNA molecules differ from the 1.85 kbp molecules by approximately 100 bp at one end and about 20 bp at the other.

Copy numbers per macronucleus of the 1.85 and 1.73 kbp α-tubulin genes

Hybridization experiments with macronuclear DNA and pmac C9 indicate that about 1/100000 of total nuclear DNA codes for α-tubulin (Fig. 5). From the measured radioactivity in the gel of Fig. 5 we can estimate the copy number for the 1.85 kbp DNA molecule to be about 55000 and about

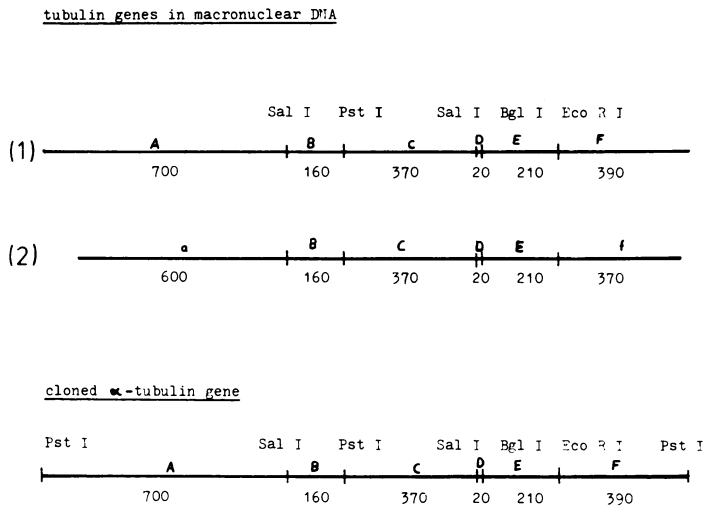


Fig. 4: Restriction maps of both macronuclear α-tubulin genes and of the cloned α-tubulin gene pmac C9. The restriction fragments of Fig. 2 and Fig. 3 were used for constructing the map. The length of the fragments A - F are given in base pairs.

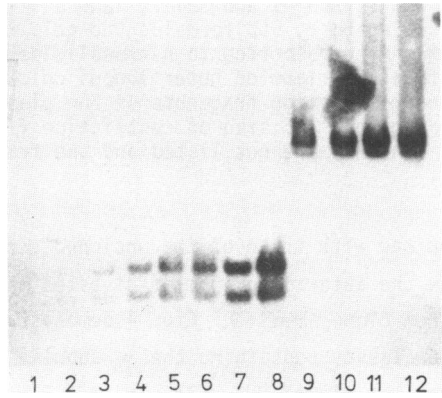


Fig. 5: Hybridization of macronuclear DNA with the cloned homologous α-tubulin gene for copy number calculations. In lane 1 - 8 the insert of clone pmac C9 is electrophoresed as reference on a 1.5% agarose gel (0.05; 0.1; 0.2; 0.4; 0.7; 1.0; 1.5; 2.5 ng DNA per lane). Pst I, which was used to remove the insert from the plasmid, creates two fragments with 860 bp and 990 bp. In lane 9 - 12 total macronuclear DNA is separated with respect to size in the same gel (5; 10; 15; 20 µg DNA per lane). After drying the gel, it was hybridized with the ^{32}P-labeled cloned α-tubulin gene (insert of clone pmac C9). For copy number calculations of both α-tubulin genes, the radioactivity of each band was meassured by liquid scintillation counting.

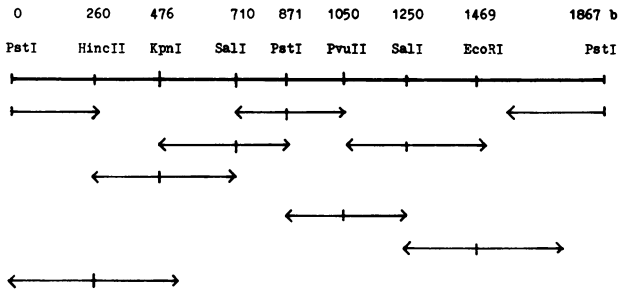


Fig. 6: Sequencing strategy of the cloned macronuclear DNA molecule containing a α-tubulin gene. Nine different restriction sites, separated by some 200 bp were used for the chemical sequencing method of Maxam and Gilbert (41). Restriction fragments of Sal I and Eco RI were labeled on the 5'-end by kinase reactions and restriction fragments of Hinc II, Kpn I, Pst I and Pvu II were labeled on the 3'-end with ddATP and terminal deoxynucleotide transferase.

17000 for the smaller one in each macronucleus, assuming a DNA content of 788 pg/macronucleus (2).

Nucleotide sequence of coding and noncoding regions

The cloned α-tubulin gene was sequenced by the chemical method of Maxam and Gilbert (41) following the strategy shown in Fig. 6. Nine different restriction sites separated by some 200 bp were employed. In most parts of the gene both strands were sequenced. Fig. 7 shows the nucleotide sequence of the total cloned macronuclear DNA molecule with the predicted amino acid sequence of the protein. The coding region starts 192 bp behind the terminal structure (repeats) on the 5'-end and is terminated 242 bp before the repeats on the 3'-end. It includes 1338 bp (incl. TGA = stop codon) and contains no intervening sequences.

Nucleotide sequences of restriction fragments from the 1.85 and the 1.73 kbp macronuclear α-tubulin genes

For a more exact comparison of both size classes of macronuclear α-tubulin genes 4 further independently obtained restriction fragments (341 bp) from macronuclear DNA in the molecular weight range of 1.7 to 1.9 kbp were cloned and sequenced. For this purpose macronuclear DNA molecules containing both size-classes of α-tubulin genes were digested with Sal I and Pvu II (Fig. 6) and the fragments containing Sal I and Pvu II restriction sites were cloned in pBR 322 as described in the "Materials and Methods". By screening the colonies with the cloned α-tubulin gene pmac C9, 48 positive clones were obtained. Four of them (pE3, pC2, pC10, pC12) were

sequenced as before (41). The sequences of the 4 α-tubulin gene fragments (Fig. 8) and the sequence of the total α-tubulin gene, pmac C9 (Fig. 7a), show two types of nucleotide sequences. The clones pC2, pC10 and pC12 are identical in respect to their nucleotide sequences but differ from those of the clones pE3 and pmac C9 in nucleotide position 516 and 522. This implies the existence of at least two different α-tubulin genes in the macronucleus of Stylonychia.

Codon usage

The most surprising feature in the codon usage is the existence of the codon TAA, beginning at nucleotide pos. 526 in all 5 clones of the α-tubulin gene sequences (Fig. 7 , 8). This implies that TAA, which is known to be a stop codon in all organisms investigated so far does not terminate the translation in the α-tubulin genes of Stylonychia lemnae. Since the cloned α-tubulin gene sequence, pmac C9, beginning at the putative initiation point + 1 to the termination point + 1338 is highly homologous with other known α-tubulin gene sequences (42, 43) it is highly improbable that termination occurs at position 526. Table 1 shows that the nucleotide sequence of the cloned α-tubulin gene, pmac C9, contains a remarkably low number of codons with G in the 3rd position. Out of 15 such codons only 6 are used for the amino acids Leu, Met, Gln, Lys and Trp. For nearly all other amino acids codons with C in the 3rd position are preferred. The codon TTT for Phe, which has never been found in lower eukaryotes (44, 45)is not present in the Stylonychia lemnae α-tubulin gene,too.

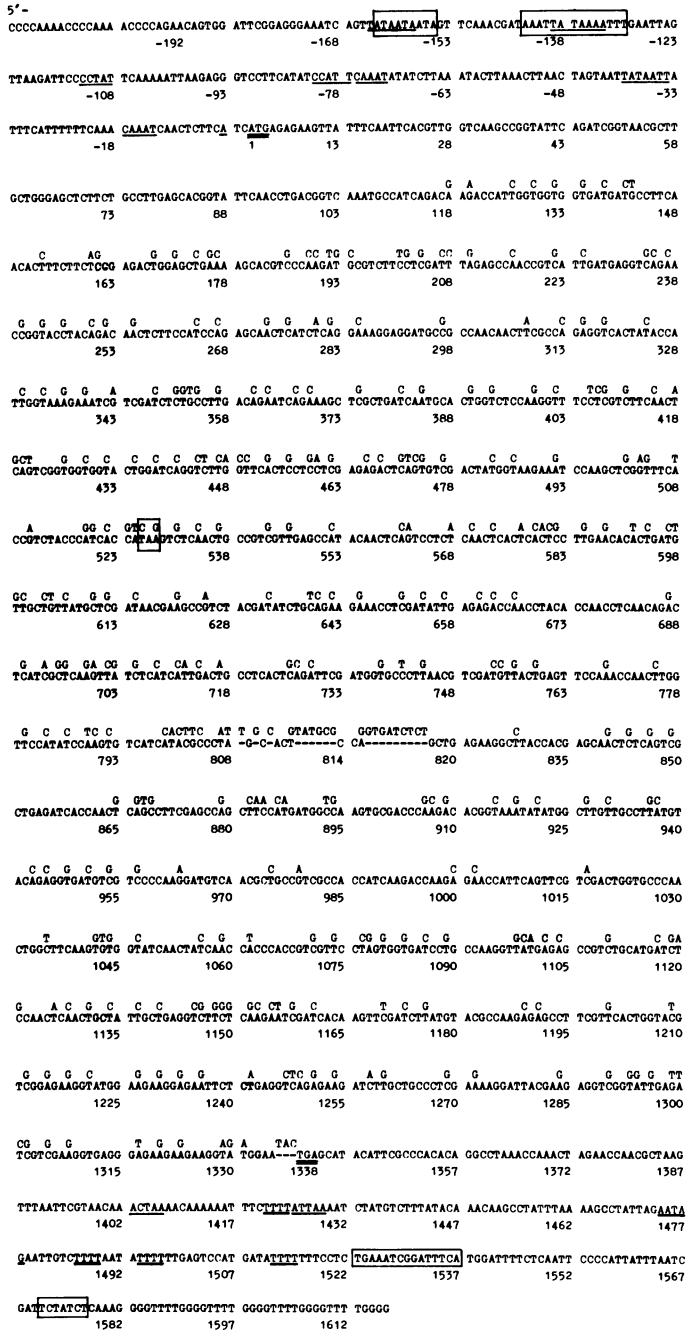
G + C content of coding and noncoding region

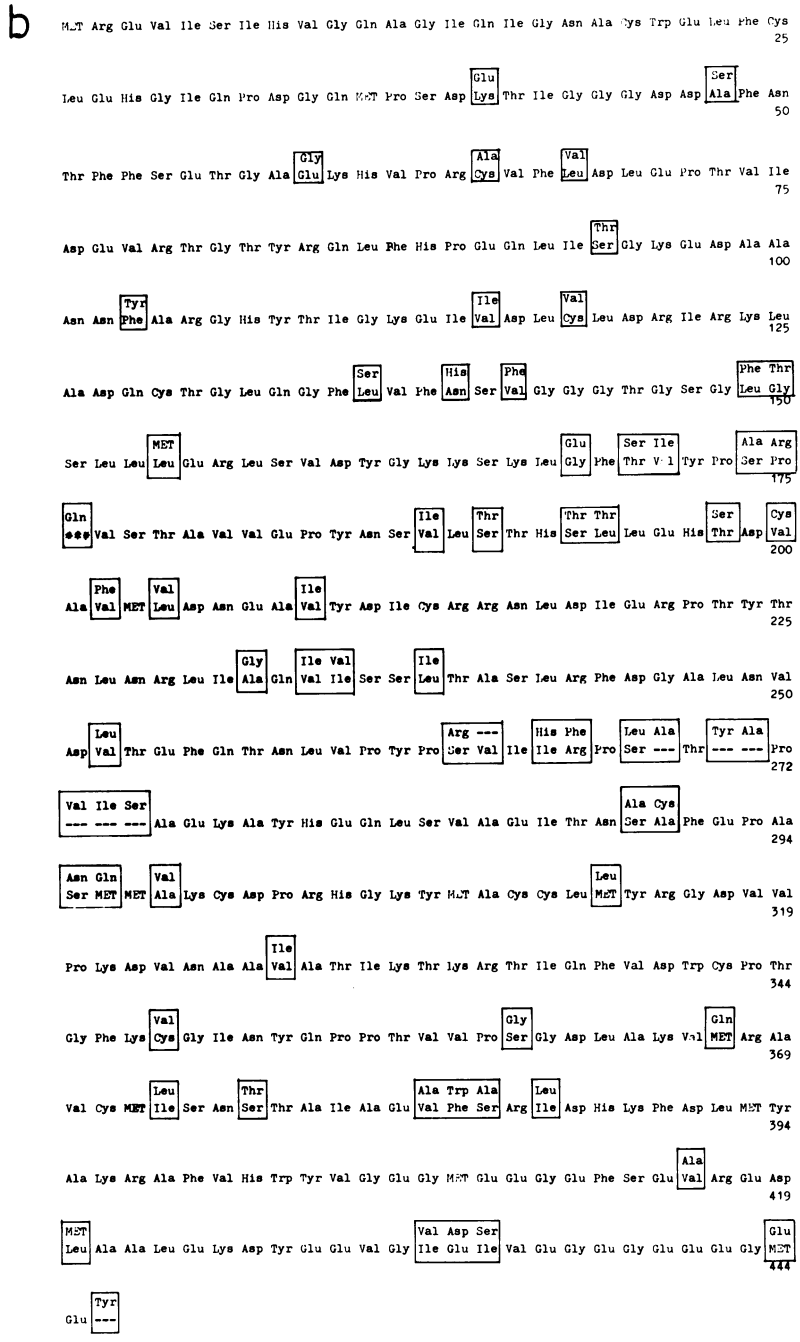
A considerable difference is found in the G + C content of coding and noncoding regions in the cloned macronuclear DNA molecule. The coding region consist of 49 % G + C residues. However the noncoding regions have only G + C contents of 25 % in the 5'-region and 28 % on the 3'-side.

Putative control signals for transcription and translation

At nucleotide positions -164, -139 and -40 (Fig. 7a) sequences are located which are identical or very similar to the "Hogness"-box TATAAAA, which is generally believed to be a transcriptional promotor (12, 13). Four other sequences at positions -112, -81, -76 and -17 (Fig. 7a) resemble the pentanucleotide CCAAA, a sequence presumably involved in the control of transcription (46). As well as the 5'-end, the 3'-noncoding region contain several possible control signals. Three sequences similar to the hexanucleotide AATAAA, which may function in eukaryotes as a polyadenylation signal (14, 15), have been found at nucleotide positions 1403, 1425, 1474. Down-

a




b
MET Arg Glu Val Ile Ser Ile His Val Gly Gln Ala Gly Ile Gln Ile Gly Asn Ala Cys Trp Glu Leu Phe Cys
25
Leu Glu His Gly Ile Gln Pro Asp Gly Gln MET Pro Ser Asp Glu Lys Thr Ile Gly Gly Gly Asp Asp Ser Ala Phe Asn
50
Thr Phe Phe Ser Glu Thr Gly Ala Gly Glu Lys His Val Pro Arg Ala Cys Val Phe Val Leu Asp Leu Glu Pro Thr Val Ile
75
Asp Glu Val Arg Thr Gly Thr Tyr Arg Gln Leu Phe His Pro Glu Gln Leu Ile Thr Ser Gly Lys Glu Asp Ala Ala
100
Asn Asn Tyr Phe Ala Arg Gly His Tyr Thr Ile Gly Lys Glu Ile Ile Val Asp Leu Val Cys Leu Asp Arg Ile Arg Lys Leu
125
Ala Asp Gln Cys Thr Gly Leu Gln Gly Phe Ser Leu Val Phe His Asn Ser Phe Val Gly Gly Gly Thr Gly Ser Gly Phe Thr Leu Gly
150
Ser Leu Leu MET Leu Glu Arg Leu Ser Val Asp Tyr Gly Lys Lys Ser Lys Leu Glu Gly Phe Ser Ile Thr Val Tyr Pro Ala Arg Ser Pro
175
Gln Val Ser Thr Ala Val Val Glu Pro Tyr Asn Ser Ile Val Leu Thr Ser Thr His Thr Thr Ser Leu Leu Glu His Ser Thr Asp Cys Val
200
Ala Phe Val MET Val Leu Asp Asn Glu Ala Ile Val Tyr Asp Ile Cys Arg Arg Asn Leu Asp Ile Glu Arg Pro Thr Tyr Thr
225
Asn Leu Asn Arg Leu Ile Gly Ala Gln Ile Val Val Ile Ser Ser Ile Leu Thr Ala Ser Leu Arg Phe Asp Gly Ala Leu Asn Val
250
Asp Leu Val Thr Glu Phe Gln Thr Asn Leu Val Pro Tyr Pro Arg --- Ser Val Ile His Phe Ile Arg Pro Leu Ala Ser --- Thr Tyr Ala --- --- Pro
272
Val Ile Ser --- --- --- Ala Glu Lys Ala Tyr His Glu Gln Leu Ser Val Ala Glu Ile Thr Asn Ala Cys Ser Ala Phe Glu Pro Ala
294
Asn Gln Ser MET MET Val Ala Lys Cys Asp Pro Arg His Gly Lys Tyr MET Ala Cys Cys Leu Leu MET Tyr Arg Gly Asp Val Val
319
Pro Lys Asp Val Asn Ala Ala Ile Val Ala Thr Ile Lys Thr Lys Arg Thr Ile Gln Phe Val Asp Trp Cys Pro Thr
344
Gly Phe Lys Val Cys Gly Ile Asn Tyr Gln Pro Pro Thr Val Val Pro Gly Ser Gly Asp Leu Ala Lys Val Gln MET Arg Ala
369
Val Cys MET Leu Ile Ser Asn Thr Ser Thr Ala Ile Ala Glu Ala Trp Ala Val Phe Ser Arg Leu Ile Asp His Lys Phe Asp Leu MET Tyr
394
Ala Lys Arg Ala Phe Val His Trp Tyr Val Gly Glu Gly MET Glu Glu Gly Glu Phe Ser Glu Ala Val Arg Glu Asp
419
MET Leu Ala Ala Leu Glu Lys Asp Tyr Glu Glu Val Gly Val Asp Ser Ile Glu Ile Val Glu Gly Glu Gly Glu Glu Glu Gly Glu MET
444
Glu Tyr ---

stream of these putative signals a sequence, TTTT, known to be a transcriptional termination signal (16), is localized at positions 1421, 1486, 1494, 1512 (Fig. 7a).
The first five nucleotides just upstream of the coding region show pyrimidines in position -5, -4, -2, -1 and a A in pos. -3. Sequences similar to this seem to act together with the codon ATG as a eukaryotic initiation site for translation (47).

Transcription of the α-tubulin genes

Nuclear and polysomal RNAs were separated according to size on denaturing gels as described in the "Materials and Methods". The α-tubulin gene transcription products were localized by hybridization with the homologous cloned α-tubulin gene. In the nuclear RNA one transcript of 1.7 kb was found which is slightly smaller than both macronuclear DNA molecules (Fig. 9). Polysomal RNA contains a transcription product with a length of 1.63 kb (Fig. 9).

DISCUSSION

The existence of two size classes of macronuclear DNA molecules coding for α-tubulin (Fig. 1b) raises the question of their function within the cell. The similarity of their restriction maps(Fig. 4) together with the known sequence of the coding region of a cloned α-tubulin gene from the larger (1.85 kbp) DNA molecules (Fig. 7a) show that both genes do not contain introns and do not differ by multiple additions or deletions as reported for pseudogenes in the human (48) and rat (49) α-tubulin gene families. The differences in length of the two macronuclear genes consist of additional nucleotides only in the terminal restriction fragments of the noncoding regions of the 1.85 kbp molecules (Fig. 4, 7a). However, comparison of the nucleotide sequences of the cloned α-tubulin gene, pmac C9, and the 4 Pvu

Fig.7a: The complete nucleotide sequence of the 1.85 kbp macronuclear DNA insert from clone pmac C9. The sequence shown is that of the antisense strand. Directions are assigned with respect to the orientation of the α-tubulin coding sequence on the DNA molecule. Bases are numbered starting with the first base of the coding sequence and terminating with the last base of the noncoding region immediately before the terminal repeat. Putative regulatory signals are underlined, presumptive replication origins and the codon TAA are boxed. Bases differing in homology from the cDNA of the α-tubulin mRNA from chicken are shown in the upper lanes.
b: The amino acid sequence encoded by the cloned 1.85 kbp macronuclear α-tubulin gene. Regions differing in homolgy from the chicken α-tubulin sequence are boxed, with the chicken residues given in the upper half of each box. Positions of deleted amino acids are indicated by broken lines.

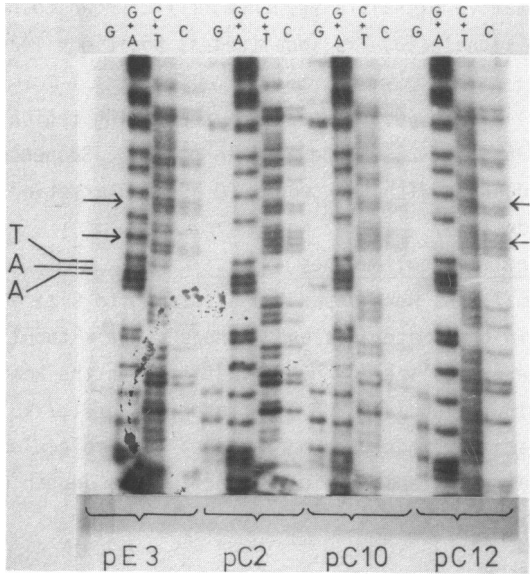


Fig. 8: Nucleotide sequences of fragments from four other macronuclear DNA-molecules coding for α-tubulin. All sequences cloned by Sal I - Pvu II restriction enzyme digestion from total macronuclear DNA contain the α-tubulin gene sequences from nucleotide 476 to 817, which include the TAA at position 526. Sequencing was done after labeling the 3'-end on the Pst I site by the method of Maxam and Gilbert (41). The arrows indicate the differences in nucleotide sequences between the 4 DNA molecules. The sequence in lane 1 is identical with that of the cloned α-tubulin gene pmac C9 in Fig. 7 a.

II - Sal I restriction fragments (Fig. 7a, 8), which were derived from both the 1.85 and the 1.73 kbp macronuclear α-tubulin genes, reveals some sequence variation. Two groups of sequences were found differing in position 516 and 522 (Fig. 7a, 8). The clones pC2, pC10 and pC12 show a T in pos. 516 and a C in pos. 522, whereas the clones pE3 and pmac C9 show a C in pos. 516 and a A in pos. 522. These variations affect the 3rd position of codons in both cases (Fig. 7a) and do not alter the encoded amino acid sequence. Therefore these data strongly suggest that both macronuclear sequences code for α-tubulin and consequently that the macronucleus of Stylonychia lemnae contains at least two types of α-tubulin genes.
The existence of codon TAA in position 526 - 528 (Fig. 7a, 8) in all α-tubulin gene sequences examined is a strong evidence that TAA is not a stop codon in Stylonychia lemnae. We suggest that it codes for the amino acid glutamine (Gln) which is encoded by CAG in the same position in rat as well

Table 1: Comparison of codon usage in α-tubulin genes from Stylonychia lemnae and chicken (42). The numbers indicate the frequency with which individual codons are used in the coding region of the cloned α-tubulin gene from Stylonychia and of the mRNA from chicken, which is shown by Valenzuela et al. (42). St. = Stylonychia, Ch = chicken.

St.	Ch.		St.	Ch.		St.	Ch.		St.	Ch.	
0	2	TTT/Phe	1	1	TCT/Ser	5	2	TAT/Tyr	2	1	TGT/Cys
19	18	TTC/Phe	5	9	TCC/Ser	10	17	TAC/Tyr	10	7	TGC/Cys
1	0	TTA/Leu	18	2	TCA/Ser	1	0	TAA	1	1	TGA
2	2	TTG/Leu	0	5	TCG/Ser	0	0	TAG	3	3	TGG/Trp
8	0	CTT/Leu	3	2	CCT/Pro	1	0	CAT/His	0	2	CGT/Arg
23	3	CTC/Leu	2	9	CCC/Pro	10	11	CAC/His	1	13	CGC/Arg
0	0	CTA/Leu	15	2	CCA/Pro	11	0	CAA/Gln	0	0	CGA/Arg
0	24	CTG/Leu	0	4	CCG/Pro	2	12	CAG/Gln	0	4	CGG/Arg
11	1	ATT/Ile	12	1	ACT/Thr	0	0	AAT/Asn	3	0	AGT/Ser
14	21	ATC/Ile	14	18	ACC/Thr	16	15	AAC/Asn	1	3	AGC/Ser
1	1	ATA/Ile	0	1	ACA/Thr	3	1	AAA/Lys	20	0	AGA/Arg
12	8	ATG/MET	0	10	ACG/Thr	16	17	AAG/Lys	0	2	AGG/Arg
11	0	GTT/Val	14	6	GCT/Ala	20	7	GAT/Asp	28	3	GGT/Gly
28	3	GTC/Val	17	22	GCC/Ala	6	18	GAC/Asp	1	13	GGC/Gly
0	1	GTA/Val	0	0	GCA/Ala	16	1	GAA/Glu	7	1	GGA/Gly
0	28	GTG/Val	0	6	GCG/Ala	21	33	GAG/Glu	0	15	GGG/Gly

as in chicken α-tubulin genes (Fig. 7a) (42,43). Indication that this assumption may be valid for other ciliates also, comes from other recently sequenced genes of the ciliates Tetrahymena and Paramecium. The immobilization antigen A gene of Paramecium tetraurelia possesses 4 interruptions of the open reading frame by TAA codons, which also code very likely for glutamine(J. Preer, pers. comm). In addition two histone H3 genes from Tetrahymena thermophila have both been shown to possess 2 TAA codons coding for glutamine (M. Gorovsky, pers. comm.). We can speculate that TGA which terminates the α-tubulin gene of Stylonychia and all other ciliate genes investigated so far (44, 50) is the only termination codon in ciliates. The differences in codon usage between Stylonychia and higher eukaryotes (42, 43) which are characterized by a clear reduction in the usage of codons with a G in the 3rd position and no usage of the codon TTT (table 1), result in a slight deviation in the nucleotide sequence, e. g. from that of chicken (42) and rat (43), but do not significantly affect the amino acid encoded. Between Stylonychia and rat there is a nucleotide homology of 74,4 % and a homology of the predicted amino acid sequence of 86,7 %, which is e. g. even higher than the nucleotide homologies (60 - 65 %) of the highly conserved histone H4 genes from Tetrahymena and vertebrates (44).

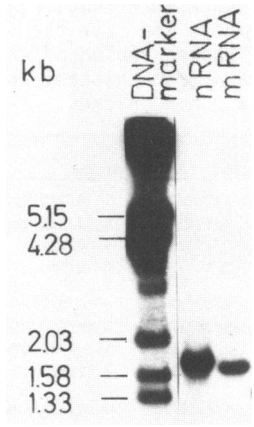


Fig. 9: Detection of transcription products of the α-tubulin genes. Total nuclear and polysomal RNA were denatured with glyoxal, electrophoresed on a 1% agarose gel and transferred to a nitrocellulose filter. Hybridization was done with the ^{32}P-labeled cloned α-tubulin gene in 5 x SSC, 10 x Denhard's solution, 0.1% SDS and 100 µg/ml tRNA at 58°C. Washings were done in 4 x SSC at 53°C.

The finding that the nuclear and polysomal transcription products of the α-tubulin genes are smaller than both size classes of the macronuclear DNA molecules (Fig. 9 , 1) supports the assumption, that both the 1.73 kbp and the 1.85 kbp molecules are transcribed. The total length of the polysomal RNA with 1630 b corresponds very well with that of the mRNA from chicken α-tubulin (42), which contains about 100 nucleotides in front of

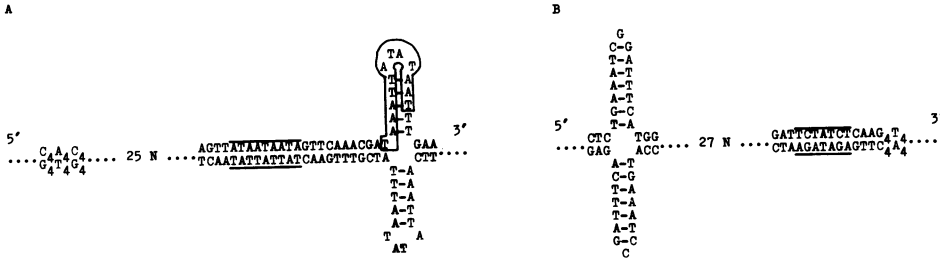


Fig. 10: Sequences possibly involved in DNA replication.
A: 5'-noncoding region of the cloned α-tubulin gene
B: 3'-noncoding region of the same cloned DNA molecule
Both noncoding regions contain A + T rich sequences capable of forming self complementary hairpin loops and also A + T rich palindromic sequences, indicated by solid lines. The boxed sequence in the 5'-hairpin loop is very similar to those of ARS known from yeast and other eukaryotes.

and about 200 nucleotides behind the translated sequence.
The cloned macronuclear DNA molecule containing the α-tubulin gene, pmac C9, shows several signals and regulatory sequences functioning in transcription initiation and termination as well as in RNA polyadenylation (Fig. 7a).
It is possible that the 8 bp containing sequence of alternating purines and pyrimidines starting immediately behind the coding region at position 1339 (Fig. 7a), which is capable of forming left handed Z-DNA (51), may be involved in regulation of gene expression. Lipps et al. have shown that Z-DNA occurs in the macronucleus of Stylonychia (52).
Both 5'- and 3'-noncoding regions contain sequences presumably involved in DNA replication. On the 5'-side an A + T rich 14 base sequence, AAATTATAAAATTT, beginning at nucleotide -143 (Fig. 7a), which is capable of forming a self-complementary hairpin loop, is found (Fig.10A). 20 bp upstream between nucleotides -163 and -155 another A + T rich palindromic sequence is localized (Fig. 7a,10A). The same combination of an A + T rich 15 nucleotides containing hairpin loop and an A + T rich palindromic sequence has also been found in the 3'-noncoding region at pos. 1523 and 1571 (Fig. 7a,10B). These structures resemble those found at the replication origin in the polyoma virus chromosome (53). The sequence beginning at nucleotide -144 in the 5'-hairpin loop (Fig. 7a,10A) is very similar to autonomously replicating sequences known from yeast and other eukaryotes (54). These similarities together with the detection of the same structures in a comparable distance of about 50 bp from the terminal C_4A_4-sequences of two other cloned and sequenced macronuclear DNA molecules (data not shown) suggest that these A + T rich sequences may function as replication origins in macronuclear DNA molecules. This assumption is supported by electron microscopic observations. They show linear molecules containing replication forks or bubbles close to one or both ends (Steinbrück, unpublished, 9). Experiments to test the hypothesis of possible replication origins are in progress.
Finally the results discussed above are further evidence that each macronuclear DNA molecule in the hypotrichous ciliate Stylonychia is a complete replication and transcription unit. Basic control and regulatory sequences for both replication and transcription processes are present on the cloned macronuclear DNA molecule containing a α-tubulin gene. But the most important finding is the abnormal codon usage of TAA in the α-tubulin gene, which seems to be a feature of other ciliates also.

ACKNOWLEDGEMENTS

I would like to thank Elke Müller for her excellent technical assistance and Dr. Axel Brennicke for his generous help in sequencing technics. I am grateful to Dr. G. Steinbrück for his valuable discussions and to Dr. H. J. Lipps for his critical comments and suggestions on the manuscript. I am also grateful to Dr. D. Cleveland for his gift of a cDNA clone of α-tubulin mRNA and to Dr. J. Preer for cooperation in the interpretation of the codon usage data. I thank R. Simon and M. Wieser for providing RNA from Stylonychia.

This work was supported by the Deutsche Forschungsgemeinschaft.

REFERENCES

1. Ammermann, D., (1971) Chromosoma 33, 209 - 238
2. Ammermann, D., Steinbrück, G., Berger, L. v. and Hennig, W., (1974) Chromosoma 45, 401 - 429
3. Lipps, H. J., Steinbrück, G. and Elsevier, S. M. (1978) Chromosoma 69, 291 - 306
4. Meyer, G. F. and Lipps, H. J. (1980) Chromosoma 77, 285 - 297
5. Meyer, G. F. and Lipps, H. J., (1982) Chromosoma 82, 309 - 314
6. Lipps, H. J., Nock, A., Riewe, M., Steinbrück, G., (1978) Nucl. Acids Res., 5, 4699 - 4709
7. Steinbrück, G., Haas, I., Hellmer, K. H. and Ammermann, D., (1981) Chromosoma 83, 199 - 208
8. Nock, A., (1981) Chromosoma 83, 209 - 220
9. Murti, K. G. and Prescott, D. M. (1983) Mol. Cell Biol. 3, 1562 - 1566
10. Lipps, H. J., Steinbrück, G., (1978) Chromosoma 69, 21 - 26
11. Wünning, I. U., Lipps, H. J. (1983) EMBO J. 2, 1753 - 1757
12. Baker, C. C., Herisse, J., Coutois, G., Galibert, E., Ziff, E., (1979) Cell 18, 569 - 580
13. Lewin, B., (1980) Gene Expression, Vol. 2, 2nd edn. (Wiley, New York)
14. Proudfoot, N. J., Brownlee, G. G., (1976) Nature 263, 211 - 214
15. Konkel, D. A., Tilghman, S. M., Leder, P., (1978) Cell 15, 1125 - 1132
16. Korn, L. J., Brown, D. D., (1978) Cell 15, 1145 - 1156
17. Oka, Y., Shiota, S., Nakai, S., Nishida, Y., Okulbo, S., (1980) Gene, 10, 301 - 306
18. Klobutcher, L. A., Swanton, M. T., Porini, P., Prescott, D. M., (1980) Proc. Natl. Acad. Sci. USA, 78, 3015 - 3019
19. Lipps, H. J., Erhardt, P., (1981) FEBS Lett. 126, 219 - 222
20. Blackburn, E. H., (1984) Cell 37, 7 - 8
21. Cohen, J., Adoutte, A., Grandchamp, S., Houdebine, L. M., Beisson, J., (1982) Biol. Cell 44, 35 - 44
22. Tucker, J. B., (1979), (ed. K. R. Roberts and J. S. Hyams) pp. 315 - 357, New York and London: Academic Press
23. Ammermann, D., (1965) Arch. Protistenk. 108, 109 - 152
24. Glisin, V., Crkvanjakov, R. and Byus, C., (1974) Biochemistry 13, 2633 - 2637
25. Palacios, R., Palmiter, R. and Schimke, R., (1972) J. Biol. Chem. 247, 2316 - 2321
26. Buell, T. and Wickens, M. P., (1978) J. Biol. Chem. 253, 2471 - 2482
27. Birnboim, H. C. and Doly, J., (1979) Nucl. Acid. Res. 7, 1513 - 1523
28. Everett, R. S., Bary, D. and Chambon, P., (1983) Nucl. Acid. Res. 8, 2447 - 2464

29. McMaster, G. K. and Carmichael, G. G., (1977) Proc. Natl. Acad. Sci. USA 74, 4835 - 4838
30. Tsao, S. G. S., Brunk, C. F. and Pearlman, R. E., (1982) Anal. Biochem. 131, 365 - 372
31. Southern, E. M., (1975) J. Mol. Biol. 98, 503 - 517
32. Thomas, P. S., (1980) Proc. Natl. Acad. Sci. USA 77, 5201 - 5205
33. Denhard, D. T., (1966) Biochem. Biophys. Res. Commun. 23, 641 - 646
34. Weislander, L., (1979) Anal. Biochem. 98, 305 - 309
35. Smith, H. O., (1980) Methods Enzymol. 65, 371 - 380
36. Rigby, P. W. J., Dieckmann, M., Rhodes. C. and Berg, P., (1977) J. Mol. Biol. 113, 237 - 251
37. Roychoudhury, R., Jay, E. and Wu, R., (1976) Nucl. Acid.Res. 3, 363 - 377
38. Mandel, M. and Higa, A., (1970) J. Mol. Biol. 53, 159 - 162
39. Cleveland, D. W., Lopata, M. A., Mac Donald. R. J., Cowan, N. J., Rutter, W. J. and Kirschner, M. W. (1980) Cell 20, 95 - 105
40. Grunstein, M. and Hogness D., (1975) Proc. Natl. Acad. Sci. 72, 3961 - 3965
41. Maxam, A. M. and Gilbert, W., (1977) Proc. Natl. Acad. Sci. USA 74, 560 - 564
42. Valenzuela, P., Quiroga, M., Zaldivar, J., Rutter, W. J., Kirschner, M. W., Cleveland, D. W. (1981) Nature 289, 650 - 655
43. Lemischka, J. R., Farmer, S., Racaniello, V. R., Sharp, P. A., (1981) J. Mol. Biol. 151. 101 - 120
44. Bannon, G. A., Bowen, J. K., Yao, M. C., Gorovsky, M. A., (1984) Nucl. Acid. Res. 12, 1961 - 1975
45. Efstradiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connel, D., Spritz, R. A., De Tiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Blechl, A. G., Smithies, O., Baralle, R. E., Shoulder, C. C., Proundfoot, N. J. (1980) Cell 21, 653 - 668
46. Koszak, M., (1984) Nucl. Acid. Res. 12, 857 - 872
47. Wilde, C. D., Crowther, C. E., Cripe, T. P., Lee, M. G. S., Cowan, N. J. (1982) Nature 297, 83 - 84
48. Lemischka, I., Sharp, P. A., (1982) Nature 300, 330 - 335
49. Ginzburg, J., Behar, L., Divol, D., Littauer, U. Z., (1981) Nucl. Acid. Res. 9, 2691 - 2697
50. Kaine, B. P., Spear, B. B., (1982) Nature 295, 430 - 432
51. Sigleton, C. K., Klysik, J., Stirdivant, S. M., Wells, R. D., (1982) Nature 299, 312 - 316
52. Lipps, H. J., Nordheim, A., Lafer, E. M., Ammermann, D., Stollar, B. D., Rich, A., (1983) Cell 32, 435 - 441
53. Soeda, E., Miura, K., Nakaso, A., Kimura, G., (1977) FEBS Lett. 79, 383 - 389
54. Montiel, J. F., Norbury, C. J., Tuite, M. F., Dobson, M. J., Mills, J. S., Kingsman, A. J., Kingsman, S. M. (1984) Nucl. Acid. Res. 12, 1049 - 1068