J Mol Evol (1989) 28:322–326

Journal of Molecular Evolution

Histone Genes in *Physarum polycephalum*: Transcription and Analysis of the Flanking Regions of the Two H4 Genes

Marcelle L. Wilhelm and F.-Xavier Wilhelm

Institut de Biologie Moléculaire et Cellulaire du C.N.R.S., 15 rue René Descartes, 67084 Strasbourg, France

Summary. The histone H4 multigene family of *Physarum polycephalum* consists of two genes, H41 and H42. Both genes have an unusual structure in that they are interrupted by a small intron. The structure of the *P. polycephalum* H4 genes is discussed and compared to the structure of histone genes of other organisms. S1 nuclease analysis was used to map the 5′ and 3′ ends of the histone H4 messengers. We show that the histone H4 genes have a hybrid structure; they are interrupted by an intervening sequence, as in replacement variant histone genes of higher eukaryotes, but their 5′ and 3′ noncoding regions have the properties of replication-dependent histone genes: the 5′ and 3′ leader and trailer sequences are short, possess a 3′- hyphenated dyad symmetry element, and a CAGA sequence is found 3′ to the hyphenated hairpin structure. This report also provides evidence that both genes are expressed in late G2 phase as well as in S phase and that their expression is temporally coordinated and quantitatively similar during the cell cycle.

Key words: Histone genes — S1 mapping — 5′ and 3′ flanking regions — Cell cycle — *Physarum polycephalum*

Introduction

The histone H4 multigene family of *Physarum polycephalum* consists of two genes, H41 and H42, that are present in only one or two copies per haploid genome. In this report we have used S1 nuclease analysis to map the 5′ and 3′ ends of the histone H4 messenger RNAs. The unusual structure of the *P. polycephalum* histone H4 genes is discussed and compared with the structure of histone genes of other organisms.

The sustained synchrony over several successive nuclear division cycles in the syncytial plasmodia phase of the myxomycete *P. polycephalum* has been exploited often to study cell cycle events. In our laboratory we have been studying histone gene expression during the cell cycle in this organism (Wilhelm et al. 1984). The synthesis of histone proteins and mRNA occurs periodically during the cell cycle of *P. polycephalum* as in higher eukaryotes. Translation of histone mRNA is essentially limited to S phase and is tightly coupled to DNA replication. It is now clearly established that in a certain number of cases, histone gene transcription is initiated before entry of the cells into S phase. This was first shown in yeast (Hereford et al. 1982) but has also been observed in mammalian cells (Sittman et al. 1983; Artishevsky et al. 1984). In naturally synchronous plasmodia of *P. polycephalum* this phenomenon is even more pronounced because histone genes are transcriptionally activated in mid-G2 phase several hours before the messengers are actually translated (Wilhelm et al. 1984; Carrino and Laffler 1986). The rate of transcription increases approximately 20-fold over its minimum value during the second part of G2, reaches a maximum at the end of G2, and drops sharply after entry of the cells into S phase (Wilhelm et al. 1988). The newly synthesized histone mRNAs that accumulate during the second half of G2 are not translated immediately but constitute a pool of molecules in anticipation of the next S phase. We have shown recently that the histone mRNA is stored as a small ribonucleoprotein in late G2 and constitutes a dowry of histone

Offprint requests to: M.L. Wilhelm

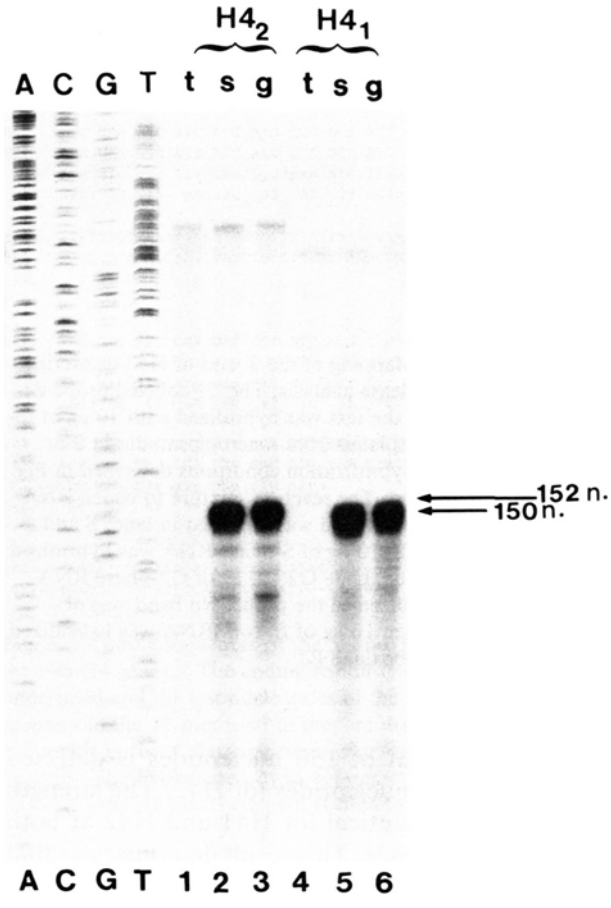

Fig. 1. Mapping of the 5′ end of H4 transcripts by S1 nuclease analysis. The ^{32}P-labeled probes described in the text were hybridized with 10 μg of total RNA isolated from macroplasmodia in S or G2. The hybridization was for 18 h at 50°C in 75% deionized formamide, 1% SDS, 20 mM Tris (pH 7.4), 0.4 M NaCl, and 1 mM EDTA. The reaction mixture was then treated with S1 nuclease, and the resistant fragment was loaded on a sequencing gel, the gel was dried, and an autoradiograph was obtained. The size of the fragments protected from hydrolysis by S1 nuclease was estimated by comparison with radioactive molecular size markers derived from an M13 clone of known sequence. Reaction mixtures to which RNA species were added were analyzed in lanes 1–6. In lanes 1–3 the H42 probe was hybridized with 10 μg *E. coli* tRNA (lane 1), 10 μg S-phase RNA (lane 2), or 10 μg G2-phase RNA (lane 3). In lanes 4–6 the H41 probe was hybridized with 10 μg *E. coli* tRNA (lane 4), 10 μg S-phase RNA (lane 5), or 10 μg G2-phase RNA (lane 6).

mRNA that may be necessary at the beginning of S because the mitotic cycle of *P. polycephalum* lacks a G1 phase. The histone messengers accumulated during the G2 phase are therefore available immediately after mitosis at the start of S phase when DNA begins to be replicated at a high rate and histone proteins are required to package newly synthesized DNA into chromatin. The low copy number of histone H4 genes may not be able to provide necessary amounts of gene products at the beginning of S phase and could explain the need for accumulating stored messengers in late G2. This report presents data showing that the two histone H4 genes (H41 and H42) are expressed during the mitotic cycle of *P. polycephalum* and that the transcripts of both genes are present in G2 and S phase of the cell cycle.

Materials and Methods

RNA Preparation. RNA was prepared as previously described (Wilhelm et al. 1984): to one synchronous macroplasmodium (1–2 g) 10 ml of lysis buffer [5.0 M guanidine isothiocyanate, 50 mM Tris-HCl (pH 7.6), 10 mM EDTA, and 2.5% 2-mercaptoethanol] were added and immediately homogenized in a Dounce homogenizer. The suspension was centrifuged for 20 min at 8000 *g* in a JS 7.6 rotor in a Beckman J21 centrifuge. The supernatant was made 4% (w/v) with respect to N-laurylsarcosine and solid CsCl was added to 0.15 g/ml. The solution was layered over a 2.0-ml cushion of 5.7 M CsCl and 0.1 M EDTA (n = 1.41) and centrifuged in a Beckman SW41 rotor for 24 h at 20°C. After centrifugation the supernatant was removed by aspiration and the tube was inverted to drain the remaining liquid. The RNA pellet was dissolved in 10 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 1% SDS and extracted with a 4:1 mixture of chloroform and 1-butanol. Then, 0.1 vol of 3 M sodium acetate (pH 5.2) and 2.2 vol of ethanol were added to the aqueous phase. After precipitation at −20°C, the RNA was recovered by centrifugation, dissolved in 1 ml of water, and the optical density was determined.

S1 Mapping. The S1 mapping of isolated RNA was carried out as described by Davis et al. (1986). An M13 clone with the inserted DNA of interest was prepared, and a radiolabeled complementary copy was synthesized by primer extension. To recover the radioactive probe, the primer-extended DNA was cut with a suitable restriction enzyme, denatured, and the products separated on a sequencing gel. The gel piece containing the radioactive probe was soaked in 0.5 M ammonium acetate, 0.01 M magnesium acetate, 0.1% SDS, and 0.1 mM EDTA. After an overnight incubation, the probe was recovered by ethanol precipitation and dissolved in hybridization buffer containing 75% formamide, 20 mM Tris-HCl (pH 7.4), 0.4 M NaCl, 1 mM EDTA, and 0.1% SDS. To hybridize the probe to the RNA, 40 μl of the probe were added to 10 μg of precipitated RNA in a 1.5-ml microfuge tube; the tube was heated for 15 min at 75°C and incubated overnight at 50°C. For the S1 nuclease reaction, 350 μl of S1 buffer containing 1.5 M NaCl, 16.6 mM $ZnSO_4$, 0.3 M sodium acetate (pH 4.5), 30 μg/ml denatured MRE 600 DNA, and 1000 U/ml S1 nuclease were added to the tube and incubated for 1 h at 37°C. The resistant DNA was extracted, precipitated, and dried. The dried pellet was redissolved in sequencing gel loading buffer, heat denatured, and loaded on a sequencing gel.

Results

S1 Mapping of G2 and S Histone H4 mRNA

The two histone H4 genes have been cloned and sequenced and were found to code for identical H4 protein molecules (Wilhelm and Wilhelm 1987). There is 87% similarity of the nucleotide sequences of the two genes, but the 5′ and 3′ nontranslated sequences are divergent (less than 30% similarity, see Fig. 3). The transcripts of H41 and H42 can

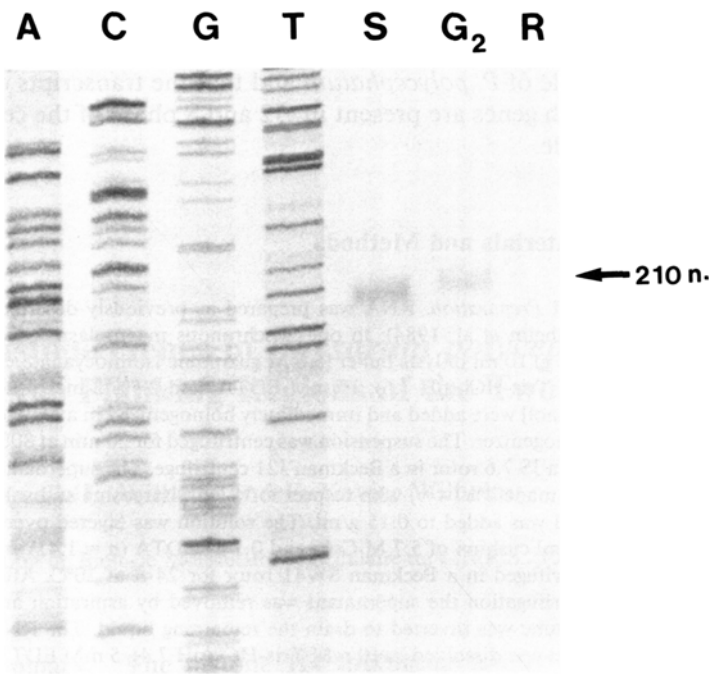

Fig. 2. Mapping of the 3′ end of H42 transcripts by S1 nuclease analysis. The ^{32}P-labeled probe described in the text was hybridized with 10 μg of total RNA isolated from macroplasmodia in S or G2. The hybridization conditions described in Fig. 1 were used. The reaction mixture to which RNA species were added was analyzed in lanes S and G2. Lane S: 10 μg of S-phase RNA was hybridized to the probe. Lane G2: 10 μg of G2-phase RNA was hybridized to the probe. No band was observed when 10 μg of *E. coli* tRNA was hybridized to the probe (lane R).

therefore be distinguished by S1 nuclease analysis. The 5′ mapping of isolated RNA was carried out as described by Davis et al. (1986). A 1.2-kb HindIII–EcoRI fragment containing the H42 gene and a 0.6-kb HindIII fragment containing the H41 gene were subcloned in M13 mp19, and single-stranded templates were prepared. A primer, 5′CACACCACCACG GCGAG3′, complementary to 17 nucleotides of the coding sequence (codons 38–43) was synthesized in our Institute by J. Colin. The primer was hybridized to the single-stranded templates and a uniformly labeled probe was synthesized using the Klenow fragment of DNA polymerase I in the presence of α^{32}P-dATP as the radioactive isotope. To recover the ^{32}P-labeled single-stranded probe, the M13 DNA was digested with HindIII, denatured, and the products were separated on a sequencing gel. A radioactive fragment of 288 nucleotides was recovered as a probe for H41. This probe is complementary to the first 129 nucleotides of the coding region and 159 nucleotides of the 5′ flanking region. A radioactive fragment of 305 nucleotides containing the first 129 nucleotides of the coding region and 176 nucleotides of the 5′ flanking region was recovered as a probe for H42. A sample of the probe for H41 or H42 was hybridized with total RNA isolated from S- or G2-phase macroplasmodia and then treated with S1 nuclease. The size of the radioactive fragments protected from enzymatic hydrolysis was measured by electrophoresis through a sequencing gel. Figure 1 shows the results of this experiment. The fragments protected by S- or G2-phase H41 transcripts and S- or G2-phase H42 transcripts are shown in lanes 5, 6 and 2, 3, respectively. A major protected fragment of 150 nucleotides is detected for H41 and 152 nucleotides for H42. The strength of the signal is identical for H41 and H42 at both stages of the cell cycle. This result demonstrates that transcripts corresponding to the two H4 genes are present in approximately equimolecular amount and that there is no differential transcription of the genes during the cell cycle. The extent to which the probes are protected establishes that the 5′ nontranslated sequences of the H4 mRNA comprises 18 (H41) and 20 (H42) nucleotides and allows us to propose a cap site sequence CATCA for H41 and TATTC for H42. These sequences compare relatively well with the consensus sequence found for other histone genes, because three (H41) or four (H42) out of five nucleotides are similar to the consensus sequence CATTC (Hentschel and Birnstiel 1981).

We have confirmed that histone H42 gene transcripts are present in S and G2 by 3′ mapping of isolated RNA (Fig. 2). The nontranslated sequences at the 3′ end of mRNA transcripts were investigated using a probe complementary to 163 nucleotides of the coding sequence, 107 nucleotides of the 3′ flanking sequences, and 67 nucleotides of M13. One major protected fragment of 210 nucleotides was detected after the probe was hybridized with S or G2 RNA and subsequently treated with S1 nuclease. The length of the protected fragment establishes that the 3′ nontranscribed sequence of the H42 mRNA comprises 47 nucleotides and maps the end of the H42 mRNA 3′ to a dyad symmetry motif hyphenated by four bases.

Shown in Fig. 3 are the parts of the H4 gene sequences relevant to the 5′ and 3′ end mapping and

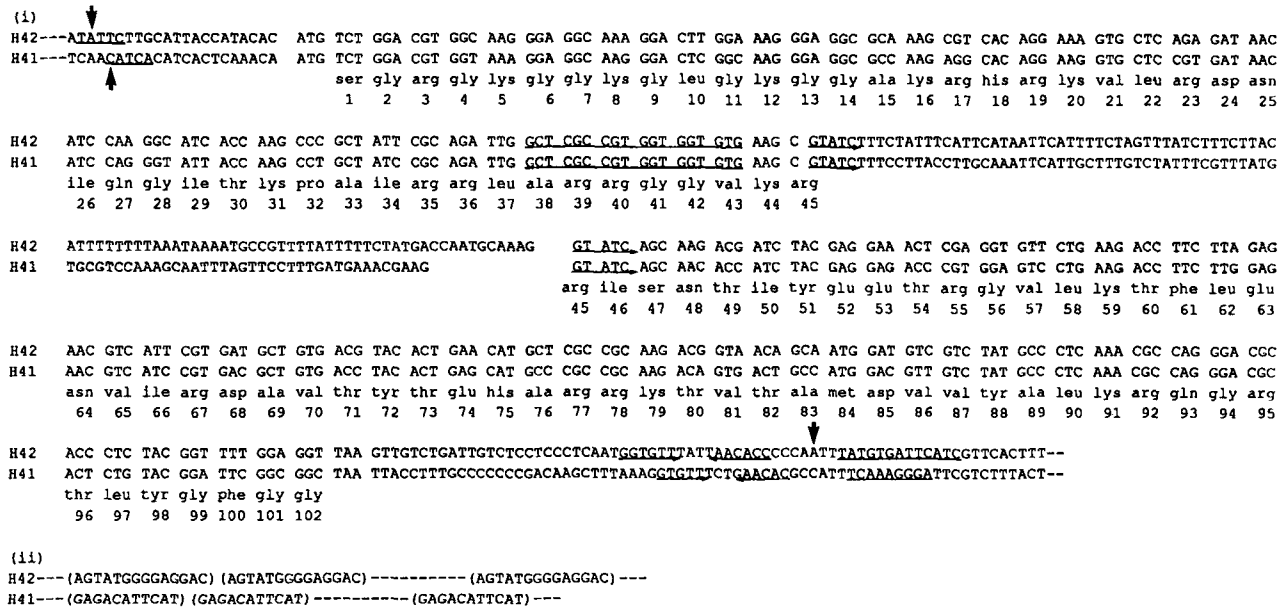

Fig. 3. Coding sequences of the 5′ and 3′ flanking sequences of H41 and H42 genes. Sequence of the two repetitive elements 3′ to the two H4 genes. **i** The coding region is shown separated from the noncoding sequences. In the 5′ flanking region the cap sites are underlined and the last nucleotides of the RNA deduced from the S1 mapping data are indicated with an arrow. The region complementary to the 17-mer used to prepare the 5′ probes is underlined. The 5-bp direct repeat flanking the intron–exon splice junction is indicated. In the 3′ flanking region the arrow indicates the position of the 3′ end of the H42 mRNA deduced from the S1 nuclease analysis. The hyphenated dyad symmetry motifs and the putative CAGA sequences are underlined. **ii** Sequence of the repetitive element 3′ to H41 (this element is directly repeated 13 times) and 3′ to H42 (this element is directly repeated 11 times).

the result of this mapping. A dyad symmetry motif is found in the 3′ noncoding region of both H4 genes and its sequence is very similar between the two genes. By analogy to its role in other species, this observation suggests that the stem-loop structure may be important in processing RNA. The sequence similarities of the dyad motif of the two genes point to the importance of the nucleotide sequence for specific recognition events. In contrast to the sequence similarities of the dyad symmetry motif, there is very little similarity between the 3′ noncoding sequences of the H41 and H42. This suggests that if the two genes are the result of the duplication of a common ancestor gene, their divergence is not recent.

Characterization of a Repeated Motif in the 3′ Region of the H4 Genes

We have found an unusual structure in the 3′ region of both H4 genes. This structure is shown in Fig. 3. Approximately 200 bp from the termination codon of H41 and H42 a short oligonucleotide is directly repeated 13 and 11 times, respectively. The function, if any, of these repeats is not known. As shown in Fig. 3, the sequences of the two repeated elements are different, GAGACATTCAT for H41 and AGTATGGGAGGAC for H42. This suggests that, whatever the exact function of this structure, the fact that the oligonucleotide is repeated is more important than its sequence. By analogy with the observation that in a number of bacterial strains (Kornberg 1982) an oligonucleotide GATC is repeated nine times within a 150-nucleotide stretch near the origin of replication, a role in the replication process can be suggested. It is also noteworthy that we have found (Jalouzot et al. 1985) that the histone H4 genes are replicated very early in S phase. Hence, a replication origin should be found very close to the histone genes. On the other hand, a series of repetitive DNA sequences have been found to be associated with human histone genes and it has been suggested that they play a role in recombination and gene duplication (Collart et al. 1985).

Conclusion

The two histone H4 genes of *P. polycephalum* have an unusual structure. Both genes are interrupted by a small intron (Wilhelm and Wilhelm 1987 and Fig. 3) that splits the codon for amino acid 45 of the predicted protein. The coding region of both genes is interrupted at precisely the same position and an identical 5-bp direct repeat flanks the intron–exon splice junction of the two genes. Recently, direct repeats flanking intron–exon splice borders have also been found in human H3.3 genes (Wells et al. 1987)

and it was argued that their presence is not fortuitous although their function is still unclear. A few histone genes containing introns have now been characterized, e.g., the H3 and H4 genes of *Neurospora crassa* (Woudt et al. 1983) or the unique H2a gene of *Aspergillus nidulans* (May and Morris 1987). In higher eukaryotes, intervening sequences have been found in the genes coding for the replacement variant H3 genes of chickens and humans (Brush et al. 1985; Wells et al. 1987). The transcripts of the replacement variant genes have characteristic features. They are posttranscriptionally polyadenylated, contain lengthy 5′ and 3′ leader and trailer sequences, and lack a 3′ hyphenated dyad symmetry element.

We have mapped the 5′ and 3′ ends of the histone H4 transcripts of *P. polycephalum*. Our results show that, although the genes are interrupted by introns, the transcripts have none of the properties characteristic of replacement variant histones, but in contrast have short 5′ and 3′ untranslated regions and possess a stem-loop structure found at the ends of most of the replication variant transcripts. Furthermore, a putative CAGA sequence (Hentschel and Birnstiel 1981) clearly is found 3′ to the hyphenated hairpin structure of histone H41 and, in a less well-conserved form, 3′ to the stem-loop structure of histone H42 gene (Fig. 3). Finally, we have recently shown (Wilhelm et al. 1988) that the transcripts lack a poly-A tail at the 3′ end.

In conclusion, the histone H4 genes of *P. polycephalum* have a hybrid structure. They are interrupted by an intervening sequence, similar to the replacement variant histone genes of higher eukaryotes, but their 5′ and 3′ noncoding regions have the properties of replication-dependent histone genes. This report also provides evidence that the two H4 genes are both expressed and that their expression is temporally coordinated and quantitatively similar during the cell cycle. Therefore both genes are required, as gene dosage considerations would suggest, in order to accumulate the necessary amount of gene products during the mitotic cycle of *P. polycephalum*.

References

Artishevsky A, Delegeane AM, Lee AS (1984) Use of a cell cycle mutant to delineate the critical period for the control of histone mRNA levels in the mammalian cell cycle. Mol Cell Biol 4:2364–2369

Brush D, Dodgson JB, Choi OR, Wilkins-Stevens P, Engel JD (1985) Replacement variant histone genes contain intervening sequences. Mol Cell Biol 5:1307–1317

Carrino JJ, Laffler TG (1986) Transcription of α-tubulin and histone H4 genes begins at the same point in the *Physarum* cell cycle. J Cell Biol 102:1666–1670

Collart D, Stein GS, Stein JL (1985) A series of repetitive DNA sequences are associated with human core and H1 histone genes. Mol Cell Biochem 67:161–170

Davis LG, Dibner MD, Battey JF (1986) In: Basic methods in molecular biology. Elsevier, New York

Hentschel CC, Birnstiel ML (1981) The organization and expression of histone gene families. Cell 25:301–313

Hereford LS, Bromley S, Osley MA (1982) Periodic transcription of yeast histone genes. Cell 30:305–310

Jalouzot R, Toublan B, Wilhelm ML, Wilhelm F-X (1985) Replication timing of the H4 histone genes in *Physarum polycephalum*. Proc Natl Acad Sci USA 82:6475–6479

Kornberg A (1982) In: Supplement to DNA replication. WH Freeman, San Francisco

May GS, Morris NR (1987) The unique histone H2A gene of *Aspergillus nidulans* contains three introns. Gene 58:59–66

Sittman DB, Graves RA, Marzluff W (1983) Histone mRNA concentrations are regulated at the level of transcription and mRNA degradation. Proc Natl Acad Sci USA 80:1849–1853

Wells D, Hoffman D, Kedes L (1987) Unusual structure, evolutionary conservation of non-coding sequences and numerous pseudogenes characterize the human H3.3 histone multigene family. Nucleic Acids Res 15:2871–2889

Wilhelm ML, Wilhelm F-X (1987) Both histone H4 genes of *Physarum polycephalum* are interrupted by an intervening sequence. Nucleic Acids Res 15:5478

Wilhelm ML, Toublan B, Jalouzot R, Wilhelm F-X (1984) Histone H4 gene is transcribed in S phase but also late in G2 phase in *Physarum polycephalum*. EMBO J 3:2659–2662

Wilhelm ML, Toublan B, Fujita RA, Wilhelm F-X (1988) Histone H4 mRNA is stored as a small cytoplasmic RNP during the G2 phase in *Physarum polycephalum*. Biochem Biophys Res Comm 153:162–171

Woudt LP, Pastink A, Kempers-Veenstra AE, Jansen AEM, Mayer WH, Planta RJ (1983) The genes coding for histone H3 and H4 in *Neurospora crassa* are unique and contain intervening sequences. Nucleic Acids Res 11:5347–5360

Received April 28, 1988/Accepted September 2, 1988