Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 1768–1770, March 1984
Genetics

Are DNA spacers relics of gene amplification events?

(eukaryotes/genome/evolution/ribosomal RNA)

WILLIAM M. WONG, JOHN L. A. ABRAHAMSON, AND ROSS N. NAZAR

Department of Botany and Genetics, University of Guelph, Guelph, ON, Canada N1G 2W1

Communicated by J. Tuzo Wilson, November 28, 1983

ABSTRACT **The genome of a thermophilic fungus, *Thermomyces lanuginosus*, contains a tandemly arranged cluster of sequences that are ≈50% homologous with the cytoplasmic 5S RNA and that are selectable by hybridization techniques. Unlike typical pseudogenes, these sequences are not truncated; rather, they bear a limited sequence homology with the entire length of the 5S RNA and are oriented end to end without significant intervening sequences. We suggest that these are gene relics that were duplicated by a rolling circle-like mechanism and that have evolutionarily drifted to become gene spacers. Accordingly, we raise the possibility that this offers a fortuitous glimpse at the origins for many of the gene spacers in the eukaryotic genome.**

In recent years the eukaryotic genome has revealed a number of intriguing and unexpected features, such as intervening sequences, pseudogenes, and extended spacer regions (see refs. 1–4). The pseudogenes, for example, are DNA segments that do not code for functional products but bear striking sequence homology to DNA of known function. Although the role and origins of these sequences are still the subject of considerable controversy, at least three mechanisms for the origin of pseudogenes have been advanced: unequal crossover events followed by mutations and passive amplification (5), imperfect transposition events followed by subsequent mutations (4), or, most recently, incomplete copying of RNA transcripts into DNA and the subsequent integration of the complementary DNA (cDNA) back into the genome (6, 7). The origins of intervening sequences or gene spacer regions are equally unclear and a number of similar alternate mechanisms can also be advanced.

Several studies on the DNAs encoding rRNAs (rDNAs) from fungi (8, 9) have revealed striking differences in the distribution of the 5S rRNA genes in this branch of the eukaryotic kingdom. Although a simple, highly repeated and tandemly arranged cluster is observed in *Saccharomyces* (8), at least seven different 5S rRNA genes are found to be dispersed in the genome of *Neurospora* (9). As a further step to explaining these intriguing evolutionary changes, we have been examining the 5S rRNA of the thermophilic fungus *Thermomyces lanuginosus*, an organism that might be expected to be phylogenetically intermediate to the other two fungi. Indeed, RNA sequence studies have indicated that there are at least two slightly different cytoplasmic RNAs (10) and preliminary hybridization studies show that these RNAs are complementary to three different *Eco*RI fragments of the genomic DNA (11).

In the present study, we characterized one of these complementary fragments, expecting it to contain a single copy or a cluster of 5S rRNA genes. Instead, we found an unusual cluster of tandemly arranged sequences that bear a limited homology with the cytoplasmic 5S rRNAs. These apparent gene relics raise the possibility that gene spacer regions, at least for the rDNA, may have been derived from the structural genes themselves.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

DNA Preparation and Cloning. *T. lanuginosus* (ATCC 16455) cells were grown with aeration at 52–55°C in aqueous medium and the mycelia were collected by filtration on Whatman no. 1 paper (10). Genomic DNA was extracted by using the methods (steps 1–12) of Cryer *et al.* (12) with a crude β-glucuronidase preparation (G0876; Sigma) substituted for Glusulase. The DNA was then repurified by centrifugation to equilibrium in a cesium chloride/ethidium bromide gradient (density: 1.54 g/ml). A genomic library of *Eco*RI fragments was prepared by using the λ Charon 3A vector (13) and this was screened for fragments complementary to the cytoplasmic 5S rRNAs by plaque hybridization (see ref. 14). The fragments were then subcloned into pBR322 and colony hybridization was used to select complementary clones (see ref. 15).

Characterization and DNA Sequence Analysis of Complementary Fragments. Fragments that were complementary to the cytoplasmic 5S RNAs were characterized and identified for DNA sequence analysis by electroblot-hybridization techniques (16). The complementary DNA was prepared by digesting the hybrid plasmid with *Eco*RI restriction endonuclease and purifying the inserted DNA on a 0.8% agarose slab gel (17). The fragment was then further digested with restriction endonuclease *Taq* I or *Hinf*I, and the products were separated on an 8% polyacrylamide sequencing gel (18) and transferred electrophoretically onto nitrocellulose membranes for hybridization studies. The 5S RNA probe was then prepared from whole cell RNA and end-labeled with [γ-^{32}P]ATP using polynucleotide kinase (19). The nucleotide sequence of the complementary fragments was determined by using the chemical degradation procedures of Maxam and Gilbert (18) and the fragments were ordered by using sequence overlaps in the two types of restriction enzyme digests.

RESULTS AND DISCUSSION

In our initial studies on the 5S rRNA genes in a thermophilic fungus, *T. lanuginosus*, hybridization experiments indicated that an *Eco*RI digest of the genomic DNA contained three fragments, 6000, 3000, and 900 base pairs, respectively (Fig. 1), which were complementary to the cytoplasmic 5S rRNA. Each of these fragments was isolated by molecular cloning (see ref. 14) in the λ Charon 3A vector using plaque hybridization to screen for the complementary regions in a *T. lanuginosus* gene library. The fragments were then subcloned into pBR322 and colony hybridization was used to select the three complementary clones (see ref. 15) designated pTL560, pTL530, and pTL509. To further characterize pTL530, the insert was isolated after *Eco*RI digestion and digested with *Taq* I or *Hinf*I restriction endonuclease for hy-

Abbreviation: rDNA, DNA encoding rRNA.

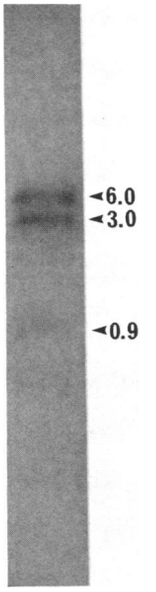

FIG. 1. Hybridization of *T. lanuginosus* 5S RNA to a restriction endonuclease *Eco*RI digest of the genomic DNA. The DNA (2 μg) was digested with 5 units of *Eco*RI, fractionated on a 0.8% agarose gel (14), and electrophoretically transferred to a nitrocellulose membrane (16). The filter was then hybridized with ^{32}P-labeled 5S RNA at 50°C overnight in 50% formamide/0.75 M sodium chloride/0.075 M sodium citrate, washed extensively with the same solution at 50°C, and washed twice with 0.30 M sodium chloride/0.030 M sodium citrate at room temperature prior to autoradiography. The molecular sizes (shown in kilobase pairs) of the fragments were determined from marker fragments (*Hin*dIII digest of λ phage DNA).

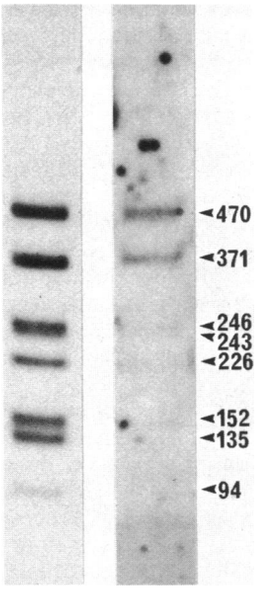

FIG. 2. Hybridization of *T. lanuginosus* 5S RNA to a restriction endonuclease *Hin*fI digest of the 3000-base-pair insert in pTL530. The DNA (1 μg) was digested with 2 units of *Hin*fI, fractionated on an 8% polyacrylamide gel (18), and subjected to hybridization analyses as described in the legend to Fig. 1. A ^{32}P-labeled digest is included (*Left*) to indicate the positions of all of the digestion products. Sizes are shown in base pairs.

bridization studies or DNA sequence analysis. As indicated in the examples shown in Fig. 2, two of either the *Hin*fI or *Taq* I fragments were complementary to the cytoplasmic 5S rRNAs. However, sequence analyses (Fig. 3) indicated that the fragments did not contain the 5S rRNA gene and were not truncated sequences of striking homology, characteristic of the typical pseudogene. Instead, more careful sequence analysis indicated that pTL530 contains a cluster of tandemly arranged sequences that are ≈50% homologous (49–51%) with the full length of the 5S rRNA gene. Still more important is the fact that these 5S rDNA-like sequences are essentially end to end without intervening spacer regions.

A similar comparison of sequences immediately adjacent to this cluster (e.g., region shown in Fig. 3) or in several of the other fragments showed the much lower degree of sequence homology (≈25%) expected in random unrelated sequences.

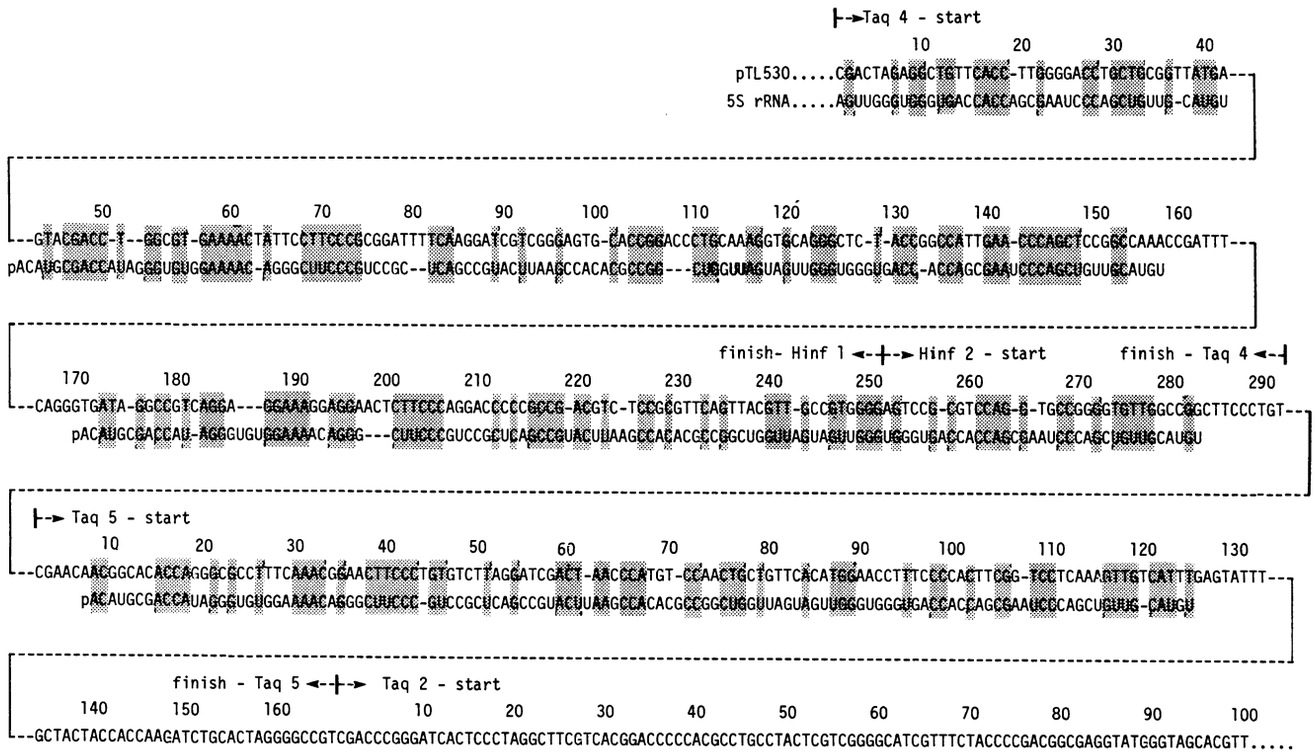

FIG. 3. A comparison between the nucleotide sequence of the cytoplasmic 5S rRNA from *T. lanuginosus* and complementary fragments in pTL530. *Hin*fI and *Taq* I digestion products were analyzed for complementarity as described in the legend to Fig. 2 and the nucleotide sequences of complementary fragments (*Hin*fI fragments 1 and 2 and *Taq* I fragments 4 and 5) were determined by the chemical degradation techniques of Maxam and Gilbert (18). The fragments are joined by overlaps in their sequences; the sequence that is shown includes the complete sequences for the two *Taq* I fragments and an adjacent nonhomologous region for comparison.

Although several explanations can be raised, the most compelling is the possibility that these sequences are gene relics, which in some distant past arose by an amplification process using a rolling circle-like mechanism. This type of mechanism is postulated because other means of gene duplication, such as unequal crossing-over, are not likely to result in essentially exact gene copies arranged end to end. Because the copies are probably all inactive genes (i.e., no 5S RNA of this sequence was isolated), it would further appear that the original amplified sequence was already a gene relic. Because they bear no greater sequence homology with each other, it would also appear that they are evolutionarily drifting, gradually becoming unrelated sequences. The fact that they are not identical or highly homologous in sequence is also suggestive of their being inactive.

The nature of these sequences raises a very interesting possibility about the origins of gene spacer sequences in general. We suggest that the observation, described here, represents a fortuitous look at the genesis of a spacer sequence and perhaps highly repetitive sequences as well. Similar mechanisms have been postulated before (see ref. 20), but the experimental evidence presented here makes the argument considerably more compelling.

We thank Ingrid Rasquinha and Michael Walter for technical assistance and Dr. Vern Seligy for supplying the cloning vectors used in these experiments. This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

1. Abelson, J. (1979) *Annu. Rev. Biochem.* **48,** 1035–1069.
2. Long, E. O. & Dawid, I. B. (1980) *Annu. Rev. Biochem.* **49,** 727–764.
3. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50,** 349–383.
4. Selker, E. U., Free, S. J., Metzenberg, R. L. & Yanofsky, C. (1981) *Nature (London)* **294,** 576–578.
5. Miller, J. R., Cartwright, E. M., Brownlee, G. G., Federoff, N. V. & Brown, D. D. (1978) *Cell* **13,** 717–725.
6. Jagadeeswaran, P., Forget, B. G. & Weissman, S. M. (1981) *Cell* **26,** 141–142.
7. Bernstein, L. B., Mount, S. M. & Weiner, A. M. (1983) *Cell* **32,** 461–472.
8. Maxam, A. M., Tizard, R., Skryabin, K. G. & Gilbert, W. (1977) *Nature (London)* **267,** 643–645.
9. Selker, E. U. & Yanofsky, C. (1981) *Cell* **24,** 819–828.
10. Wildeman, A. G. & Nazar, R. N. (1982) *J. Biol. Chem.* **257,** 11395–11404.
11. Wong, Wm. M., Abrahamson, J. L. A., Wildeman, A. G. & Nazar, R. N. (1983) *Proc. Can. Fed. Biol. Soc.* **26,** 227.
12. Cryer, D. R., Eccleshall, R. & Marmur, J. (1975) *Methods Cell Biol.* **12,** 39–44.
13. Blattner, F. R., Williams, B. G., Blechl, A. E., Denniston-Thompson, K., Faber, H. E., Furlong, L.-A., Grunwald, D. J., Kiefer, D. O., Moore, D. D., Schumm, J. W., Sheldon, E. L. & Smithies, O. (1977) *Science* **196,** 161–169.
14. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
15. Davis, R. W., Botstein, D. & Roth, J. R. (1980) *Advanced Bacterial Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
16. Bittner, M., Kupferer, P. & Morris, C. F. (1980) *Anal. Biochem.* **102,** 459–471.
17. Studier, F. W. (1973) *J. Mol. Biol.* **79,** 237–248.
18. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
19. Donis-Keller, H., Maxam, A. & Gilbert, W. (1977) *Nucleic Acids Res.* **4,** 2527–2538.
20. Federoff, N. V. (1979) *Cell* **16,** 697–710.