Volume 9 Number 23 1981 Nucleic Acids Research

Human DNA sequences complementary to the small nuclear RNA U2

Gunnar Westin, Hans-Jürg Monstein, Jan Zabielski, Lennart Philipson and Ulf Pettersson

Department of Microbiology, The Biomedical Center, Box 581, S-751 23 Uppsala, Sweden

Received 6 October 1981

ABSTRACT

Clones containing sequences complementary to the small nuclear RNA U2 were isolated from a human DNA library (1). Three clones, designated U2/4, U2/6 and U2/7 were purified and characterized by restriction enzyme cleavage, hybridization and heteroduplex analysis. Hybridization showed that the three clones each contained one single region which is complementary to U2 RNA. Restriction enzyme cleavage revealed furthermore that the inserted fragments in the three recombinants are different. Heteroduplex analysis identified a 240-380 bp long duplex region in each heteroduplex which includes sequences complementary to U2 RNA. Heteroduplexes between clones U2/4 and U2/7 as well as between U2/4 and U2/6 revealed two additional approximately 200 bp long homologies. The remainder of the inserts were found to lack measurable sequence homology.

Two fragments from clone U2/4 were subcloned in the pBR322 vector and the subclones were used to determine the nucleotide sequence of a region in clone U2/4 which is complementary to U2 RNA. A comparison between the established sequence and the sequence for rat U2 RNA (2) reveals several discrepancies.

INTRODUCTION

The nucleus of animal cells contains several low molecular weight RNA species (snRNA) which can be separated according to their size. They are designated U1, U2, U3 etc (for a review see ref. 2). Recently a great deal of attention has been focused on the snRNAs since they may play a role in the splicing reaction. Lerner et al. (3) and Rogers and Wall (4) have suggested that the U1 RNA serves its function by aligning the exon sequences during the splicing reaction. This hypothesis is based on the observation that sequences near the 5' end of the U1 RNA can base-pair with sequences at exon/intron junctions. Since the primary sequence at intervening sequence junctions cannot by itself explain the specificity of the splicing reaction the existence of RNA linker molecules is an attractive model. Ohshima et al. (5) have proposed a modified version of this model which also involves U2 RNA.

Several laboratories have performed structural studies on the snRNAs and the primary sequence is known for several species, both of human and animal origin. Little is known, however, about the genes for the snRNAs and their transcription. Conflicting reports have been published as to whether they are transcribed as a precursor molecule or not. Some investigators have reported that the transcription unit for the U1 RNA is very large whereas others failed to recognize precursors to mature U1 RNA (6,7).

We have studied the low molecular weight RNAs which are produced as a result of adenovirus infection and recently turned our attention to their cellular counterparts. The present study concerns the organization of sequences in the human genome which are complementary to the U2 RNA. Several clones from a human DNA library, containing sequences related to U2 RNA were isolated and their structures examined. The results show that the human genome contains multiple copies of sequences which are related to U2 RNA. Furthermore we show that the sequences which flank the U2 related sequences are strikingly different in all clones which have been analyzed.

MATERIALS AND METHODS

Isolation of ^{32}P-labeled RNA:

HeLa cells in suspension ($5x10^5$/ml) were maintained in Eagle's spinner medium containing 2% of the usual concentration of phosphate and 8% dialysed calf serum. The cells were labeled with ^{32}P-phosphate at a concentration of 0.2 mCi/ml for 24 hrs. Nuclear and cytoplasmic fractions were prepared at 4°C by the following procedure: cells were harvested by centrifugation and washed with 50 ml phosphate buffered saline (PBS). The cells were then disrupted by ten strokes in a tight fitting Dounce homogenizer in 4 ml of homogenizing buffer (100 mM Tris HCl pH 7.9, 100 mM NaCl, 4 mM $MgCl_2$, 1 mM $MnCl_2$, 1 mM PMSF, 6 mM β-mercaptoethanol, and 0.4% Nonidet P-40). The nuclear fraction was isolated by low speed centrifugation and washed once more with the homogenizing buffer. The purified nuclei were resuspended in 4 ml sonication buffer (the same as the homogenizing buffer without NP-40 and with only 1 mM $MgCl_2$) and disintegrated by sonication for 2x15 sec. In some experiments it was necessary to disperse the nuclei with a wide bore pipette prior to sonication. The extracts from the sonicated nuclei were purified by centrifugation over a 30% sucrose cushion in sonication buffer in order to remove the chromatin.

Purification of ^{32}P-labeled RNA:

The nuclear and cytoplasmic extracts were incubated in the presence of 1

mg/ml proteinase K and 0.5% SDS at 37°C for 1-2 hrs. After incubation, the supernatants were diluted 1:10 with 10xHSB buffer (0.5 M Tris-HCl, pH 7.5, 0.1 M EDTA and 3.5 M NaCl) and extracted three times with phenol, saturated with HSB buffer and finally extracted once with chloroform isoamylalcohol (24:1). The aqueous phases were precipitated with ethanol and dissolved in a minimal amount of a buffer which contained 10 mM Tris HCl pH 7.5, 1 mM EDTA and 0.2% SDS. The nuclear and cytoplasmic RNA fractions were then separated by electrophoresis on an 8% denaturing urea-polyacrylamide gel (8). The radioactive bands were eluted with a buffer which contained 10 mM Tris HCl pH 7.5, 1 mM EDTA 0.2% SDS and concentrated by ethanol precipitation. The purity of the RNA preparations was tested by two-dimensional finger print analysis (9).

Screening of recombinant phages:

A human DNA library was kindly supplied by Dr. T. Maniatis. This library consists of fragments from human fetal DNA, generated by partial digestion with a mixture of endonucleases AluI and HaeIII and inserted into the Charon 4A vector (1). The library was screened according to the method of Benton and Davis (10) using in vivo labeled U2 RNA as a probe. Approximately 100.000 phages were screened with 10.000 on each nitrocellulose filter. Hybridizations were carried out in 6xSSC (SSC is 0.15 M NaCl, 0.015 M tri-sodiumcitrate), 3xDenhardt's solution (11) and 0.5% SDS for 12-15 hrs at 65°C. The filters were subsequently washed in 2xSSC with 0.5% SDS for 2 to 3 hrs at 65°C and autoradiographed for 2-3 days, using intensifying screens.

Plaque purification of recombinant phages and preparation of DNA:

Phage stocks were prepared by at least three cycles of plaque purification and then propagated in NZY medium, using the PDS method (12). Virus particles were precipitated from the medium by polyethylene glycol 6000 and purified by two cycles of CsCl equilibrium gradient centrifugation. DNA was prepared by Proteinase K treatment and phenol extraction.

Restriction enzyme analysis:

The restriction endonucleases were purchased from New England Biolabs or Boehringer and used according to the specifications of the manufacturers. Fragments were separated on 1% agarose gels and were in selected experiments transferred to nitrocellulose filters (13).

Subcloning of restriction fragments from clone U2/4 in pBR322:

Twenty µg of U2/4 DNA was cleaved with endonuclease TaqI and the resulting fragments were separated in a 10-40% linear sucrose gradient at 30.000 rpm and 20°C for 20 hrs using the SW40 rotor (14). DNA from the individual

gradient fractions was analyzed by gel electrophoresis and fractions containing fragments of the appropriate size were pooled. The pooled fractions were concentrated by ethanol precipitation and subsequently ligated into the ClaI site of the vector pBR322. The ligated DNA was used to transform the E.coli K12 strain HB101, and transformants were selected on ampicillin containing plates. Recombinant clones were finally screened for U2 sequences, using ^{32}P-labeled U2 RNA as a probe. Plasmids were amplified and purified as described before (15).

An approximately 2 kb long HpaII fragment of DNA from clone U2/4 which contains sequences complementary to U2 RNA was also cloned in the ClaI site of the pBR322 vector, using the same protocol as described above.

Electron microscopy:

Three clones, hybridizing to U2 RNA were heteroduplexed against each other essentially as described by Davis et al. (16). Isodenaturing conditions were used for spreading (17). The hyperphase contained 45% formamide, 0.1 M Tris pH 8.5, 0.01 M EDTA, 0.4 µg DNA/ml and 50 µg/ml cytochrome C, layered upon a hypophase of 15% formamide, 0.01 M Tris pH 8.5, 0.001 M EDTA, and at an ambient temperature of 25°C. This corresponds to Tm -25°C for a DNA with 50% G+C (18) and under these spreading conditions a segment of heteroduplex DNA with fewer than 35% of bases mismatched should appear as duplex DNA. PM2 DNA was added as a double stranded DNA marker (9950 base pairs) and ØX174 DNA as the single stranded DNA marker (5383 bases).

In order to minimize lateral aggregation between the substitution arms, cytochrome C, prepared as described by Ferguson and Davis (19) was used, as well as three times recrystallized formamide.

DNA sequence analysis:

The protocol of Maxam and Gilbert (8) was followed for end-labeling of fragments and subsequent sequence analysis.

RESULTS

Cloning of sequences complementary to U2 RNA. To isolate genes for U2 RNA, the recombinant DNA library of Lawn et al. (1) was screened for the presence of sequences complementary to U2 RNA. Approximately 100.000 recombinants were screened and among those 10 gave a positive signal. Three of the positive recombinants were purified by repeated cycles of plaque purification and phage was produced and purified as described by Maniatis et al. (20). DNA was then extracted and studied by restriction enzyme cleavage, hybridization and heteroduplex analysis.

Characterization of DNA from three recombinant phages by restriction enzyme cleavage and hybridization. The three isolated recombinants were designated U2/4, U2/6 and U2/7. DNA from each of them was cleaved with several restriction enzymes. The results showed that all recombinants gave rise to a unique restriction enzyme cleavage pattern. This was particularly notable when the DNA was cleaved with a mixture of restriction endonucleases EcoRI and XhoI which fail to cleave in the vector arms (Fig. 1). In order to establish whether the recombinants contain one or more copies of the presumptive gene for U2 RNA, DNA from each recombinant was cleaved with several different restriction enzymes and the resulting fragments were separated on 1% agarose gels. The fragments were then transferred to a nitrocellulose sheet by the method of Southern (13) and hybridized with ^{32}P-labeled U2 RNA. Fig. 2 shows that even when the phages were cleaved with endonucleases such as TaqI which generate small fragments, only one band was found to hybridize to U2 RNA. The size of the fragment which gave positive hybridization was unique for each recombinant, again suggesting that the three recombinants are different.

From those results we conclude that the three recombinants U2/4, U2/6 and U2/7 each contain one copy of sequences which are complementary to U2 RNA and that they have different structures.

Heteroduplex analysis: In order to compare the sequences which are present in clones U2/4, U2/6, and U2/7 heteroduplex molecules were formed between the DNA from the three recombinants and λ Charon 4A DNA. Examination of the heteroduplexes, formed between each of the clones and the Charon 4A DNA, revealed that the size of the inserted fragments in the three recombinants ranged between 15.9 and 16.9 kilobases (kb). By studying under-renatured

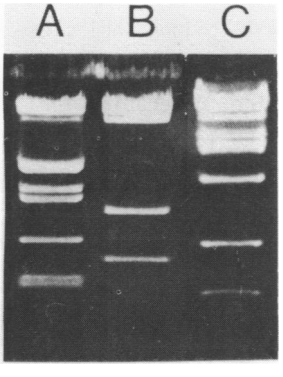

Figure 1: Cleavage of DNA from the three recombinants U2/4 (A), U2/7 (B) and U2/6 (C) with a mixture of restriction endonucleases EcoRI and XhoI. The fragments were separated on a 1% agarose gel.

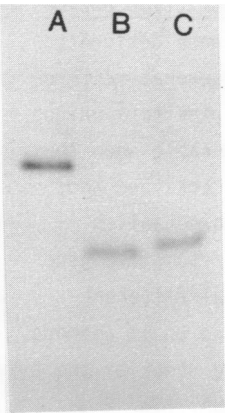

Figure 2: Hybridization between ^{32}p-labeled U2 RNA and TaqI fragments of clones U2/4 (A), U2/6 (B) and U2/7 (C). The fragments were separated on a 1% agarose gel and transferred to a nitrocellulose membrane (13).

molecules of each cloned DNA it was usually possible to assign each sustitution arm to the corresponding DNA in the heteroduplex (see, for instance Fig. 3A).

a. Heteroduplexes between DNA from clones U2/4 and U2/7: A total of 19 heteroduplex molecules were examined. The inserts showed in 17 of the 19 heteroduplex molecules one homology area which was 230±48 base pairs long (bp) (Fig. 3A and 4). Some molecules contained one or two additional homology regions which also were approximately 200 bp long (Fig. 3B). The distribution of the different homology regions in individual heteroduplex molecules is summarized in Table I. Since the hybridization results showed that each recombinant only contains one copy of the sequence which are complementary to U2 RNA the most likely interpretation of the results is that the most frequently occurring homology region, designated "a" in Fig. 4, corresponds to U2 sequences. In order to prove that this region, indeed, corresponds to U2 sequences, we made use of the finding that the U2-related sequences in clone U2/7 are located in an approximately 1 kb long XhoI cleavage fragment (Fig. 1). Our results from restriction enzyme cleavage showed that endonuclease XhoI only cuts DNA from clone U2/7 twice, generating the 1.09 kb fragment and two large fragments, 20.1 and 23.5 kb long (histograms not shown). The latter fragments contain arms "A" and "B" of the vector. In order to correlate the 1.09 kb fragment to the observed homology region heteroduplexes were formed between the XhoI cleaved clone U2/7 DNA and DNA from the Charon 4A vector. Because of the typical secondary structures which are present in the single strands from clone U2/7 (Fig. 3A) it was possible to show that the small XhoI fragment is located between 20.1 and 21.2 kb

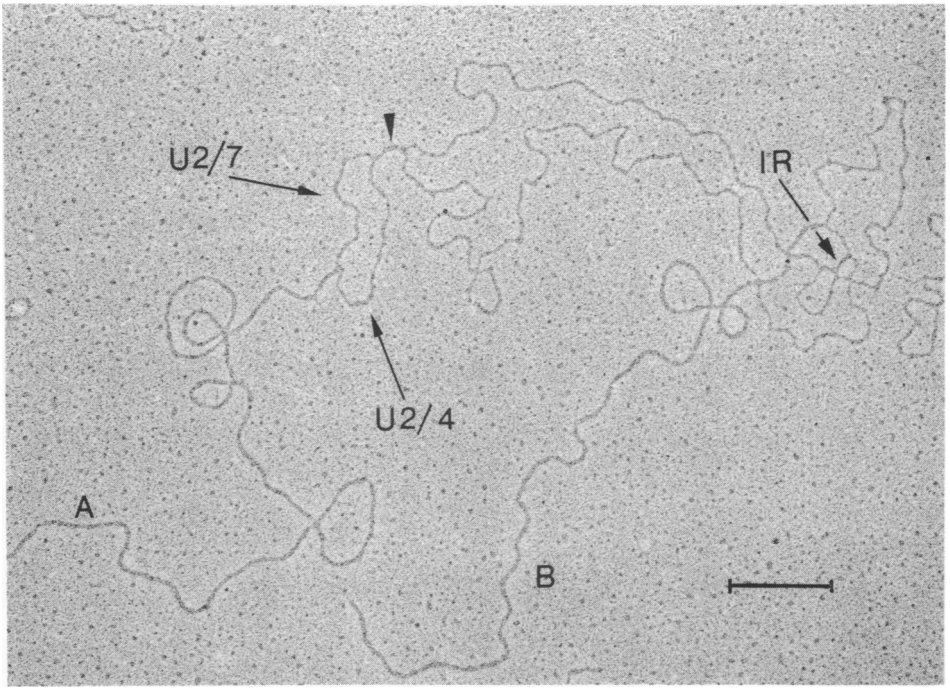

Figure 3: Electronmicrographs showing heteroduplex molecules formed between the three clones, which contain U2-like sequences. Panel A and B show heteroduplexes between DNA from clones U2/4 and U2/7, and panel C a heteroduplex between clones U2/4 and U2/6. Arms "A" and "B" of the vector are indicated and the bar represents 1000 bp of duplex DNA. In panels A and C only homology region "a" (▲) is visible. In panel B all three homology regions "a", "b" and "c" (▲), are visible (homology region "a" is closest to arm "A" of the vector). Several prominent secondary structures are present in the U2/7 insert. A frequently occurring inverted repeat structure (IR) is illustrated in panel A. Thick lines represent duplex DNA and thin lines simplex DNA in the enlargements.

from the end of arm "A" in the vector (pictures not shown). This region covers the frequently appearing homology region "a" but not the other two homology regions (Fig. 4).

From these results we conclude that each clone carries an approximately 230±48 bp long homology region which includes sequences that are related to U2 RNA. In addition there are two less frequently appearing homology regions, which are designated "b" and "c" in Fig. 4 and Table 1. These do not appear to contain U2 sequences since the hybridization studies, described in the previous section, reveal that only one fragment in each recombinant

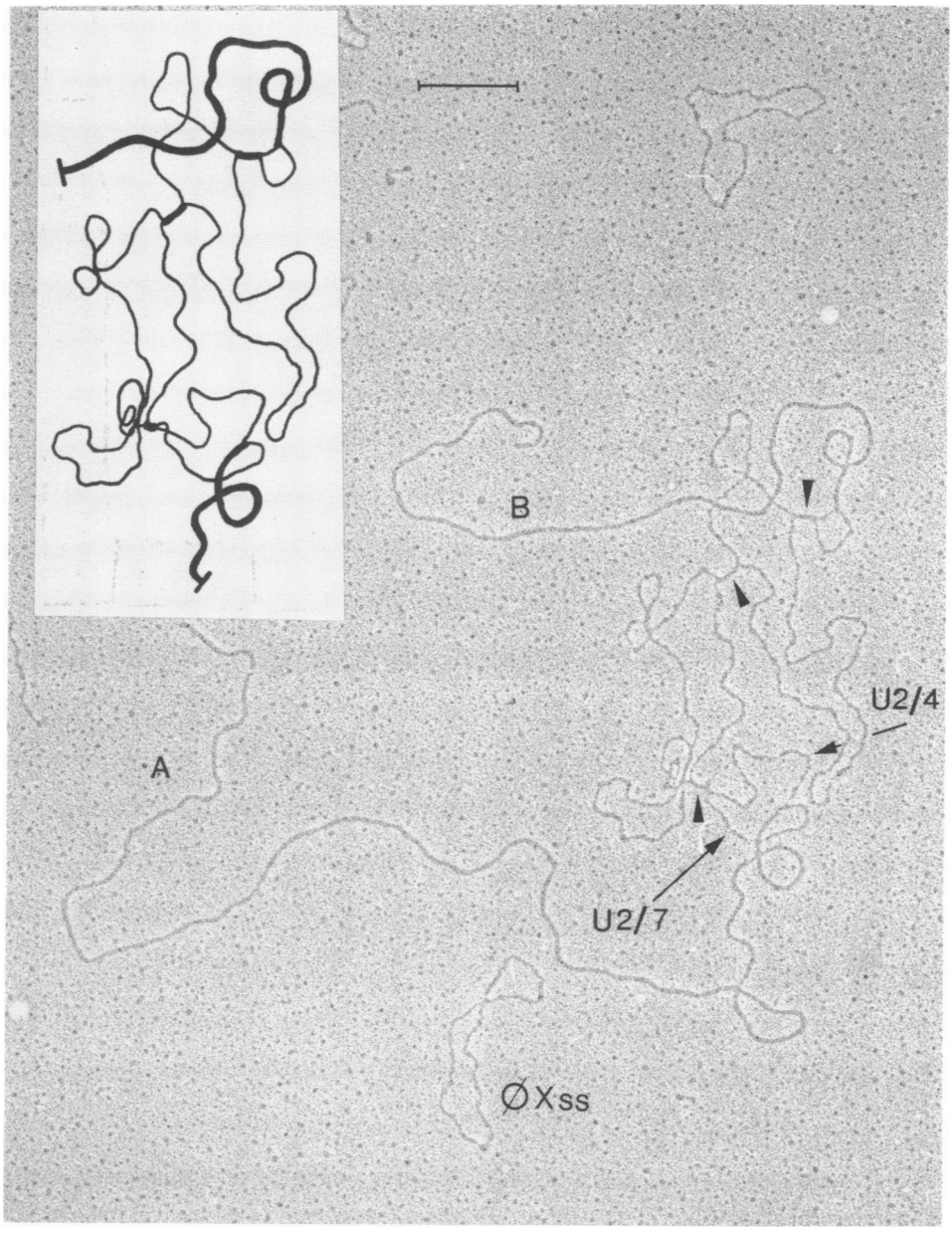

Figure 3B

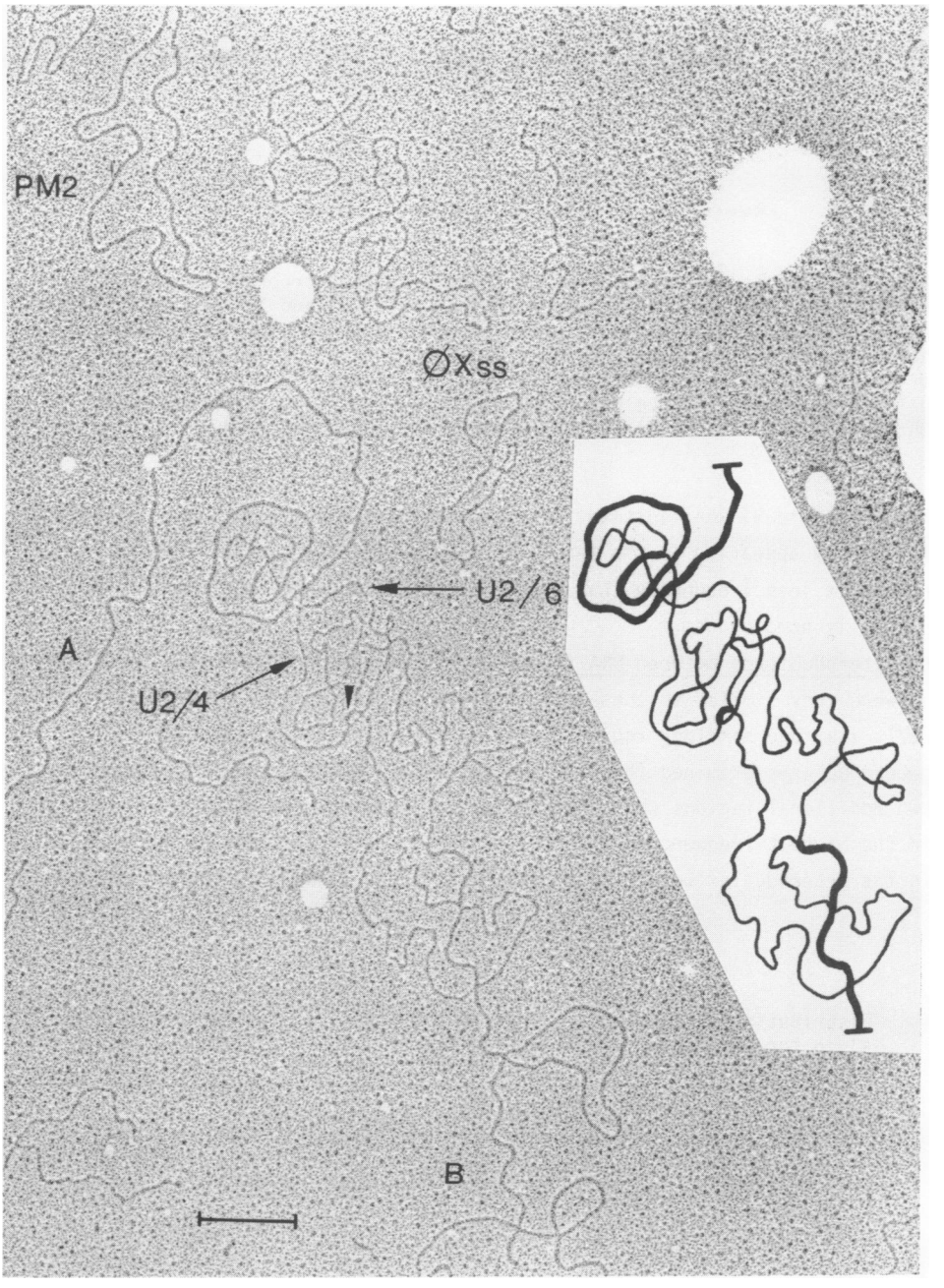

Figure 3C

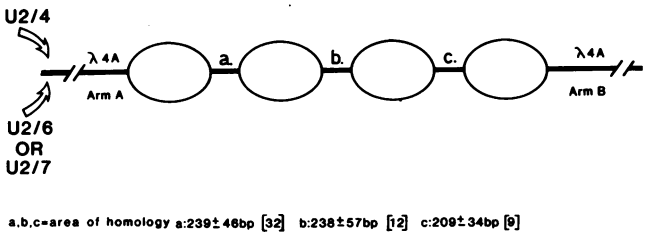

Figure 4: A schematic drawing which summarizes the structure of the three heteroduplex pairs. Three homology regions "a", "b" and "c" are shown. The measurements for the single stranded regions which are located between the homology regions are given in Table 2. At the bottom of the figure the sizes of thee homology regions are indicated. The numbers represent averages calculated from all three heteroduplex pairs. The number of measurements which were used to calculate each average is also shown ( ).

hybridizes to U2 RNA. Instead, they probably represent repetitive sequences, frequently appearing in the mammalian genome. The observation that they hybridize less frequently than the U2 sequences suggest that they are only partially homologous.

b. Heteroduplexes between DNA from clones U2/4 and U2/6: The results obtained were very similar to those obtained with heteroduplexes between U2/4 and U2/7. i.e. one homology region 239±42 bp long was present in nearly all heteroduplexes examined (Figs.3C and 4,Table I) whereas some contain one or two additional regions which hybridize. The homology region which is present in the highest frequency maps precisely in the same position as region "a" in the heteroduplex between U2/4 and U2/7 (Table 2) and must hence include

TABLE I

Distribution of homology regions "a", "b" and "c" (as defined in Fig. 4) in three different heteroduplex pairs.

	Number of molecules which contains the indicated homology regions						
Heteroduplex	a	b	c	a+b	b+c	a+c	a+b+c
U2/4 vs U2/6	11	1	0	5	1	0	1
U2/4 vs U2/7	15	1	1	1	0	0	1
U2/6 vs U2/7	11	0	0	0	0	0	0

TABLE 2
Measurements (kb) of the single stranded segments present in different heteroduplex pairs.

Heteroduplex pair	Arm A-a	a-b	b-c	c-Arm B
U2/4(vs.U2/6)	2.50±0.29 (17)[a]	3.13±0.8 (5)	7.95±0.1 (2)	1.43±0.15 (7)
U2/4(vs.U2/7)	2.76±0.29 (19)	2.62±0.1 (2)	8.84 (1)	1.41±0.49 (2)
U2/6(vs.U2/4)	5.38±0.35 (17)	1.53±1.3 (5)	0.44±0.1 (2)	9.19±0.55 (6)
U2/6(vs.U2/7)	5.14±0.26 (11)			
U2/7(vs.U2/4)	1.74±0.19 (24)	10.0±0.70(2)	3.77 (1)	0.56±0.45 (3)
U2/7(vs.U2/6)	1.73±0.20 (11)			

a) The numbers in parenthesis refer to the number of molecules which were studied for each particular homology region.

U2 sequences. The other two homology regions "b" and "c" are located in the same position in the U2/4 DNA when heteroduplexes were formed with U2/7 and U2/6 DNA (Table 2). Thus, they most likely represent the same sequences in both cases.

c. Heteroduplexes between DNA from clones U2/6 and U2/7

These heteroduplexes exhibited only one homology region which is located in the same position as segment "a", in the previously described heteroduplexes (Table 1 and 2). The duplex segment was in this case significantly longer, 382±90 bp.

Typical heteroduplex molecules are demonstrated in Fig. 3A,B,C and Tables 1 and 2 show the interpretation of the data.

Subcloning and sequencing of the segment from U2/4 DNA which is complementary to U2 RNA

In an attempt to obtain more detailed structural information about clone U2/4 which appeared to contain sequences that are closely related to U2 RNA we sequenced the appropriate part of the U2/4 DNA. To facilitate the sequence analysis a fragment was subcloned in the pBR322 vector. Hybridization between U2 RNA and TaqI fragments of the U2/4 recombinant showed that most of the U2 RNA sequences were present in an approximately 1500 bp long TaqI fragment (Fig. 2). This fragment was inserted into the ClaI site of the pBR322 vector and a recombinant plasmid, designated U2/118, was isolated. The nucleotide sequence was determined for part of the inserted fragment,

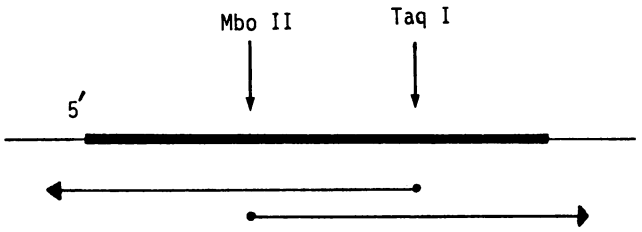

Figure 5: A schematic drawing which shows our sequencing strategy. The U2-related sequence is indicated by a thick bar. Selected restriction enzyme cleavage sites are indicated and the arrows show how the sequence in Fig. 6 was established.

according to the strategy which is outlined in Fig. 5. The established sequence revealed, as expected, homology with the known sequence for rat U2 RNA (2). When the sequence was established it became apparent that a small part of the sequence which is complementary to U2 RNA is absent from clone U2/118. Therefore a longer, approximately 2 kb long HpaII fragment, from clone U2/4 was inserted into the ClaI site of the pBR322 vector. Part of the inserted fragment was sequenced as outlined in Fig. 5. The combined sequence covers the entire sequence for U2 RNA as well as some flanking sequences and is shown in Fig. 6. The results show that the sequenced region matches the

```
                -40       -30       -20       -10
(A)          5'-GAGCAACAAATTTTCAAACACCGTGGAGGGGAGAGAGAAGAGA

             1         10        20        30        40        50
(A)          A T CACTTCTTG G CCTTTTA GCTAAG ATCAA GTGTAGTATCTGTTTTT ATCA
(B)  m3 2,2,7 GpppAmUmCGCΨΨCUCGmGmCCΨUUUGmGCUAAGmAUCAAmGUGΨAGΨAΨCΨGΨΨCUUmAUCA
                *    *         *                              *
                60        70        80        90        100
(A)          GTTTAATATA TGATACATCTTCTATCCAAGGACAATATATTAAATGGATTTTTGGAGC
(B)          GUΨUAAΨAUCmUGAUACGUCCUCUAUCCGAGGACAAUAΨAΨUAAAUGGAUUUUUGGAAC
                      *       * *       *                          *
             110       120       130       140       150       160
(A)          AGAGAGATGGAATAGAAGCTTGCTCCATCCACTCCATGCATCGACCTGGTATTGCAGTA
(B)          UAGGAGUUGGAAUAGGAGCUUGCUCCGUCCACUCCACGCAUCGACCUGGUAUUGCAGUA
             ***   *         *         *         *
              170       180
(A)          CCTCPyAGGAACAGTGCACC C CC-3'
(B)          CCUCC AGGAACGGUGCACC(A)OH
                       *
```

Figure 6: Partial sequence of the insert present in clone U2/4 (A). The sequence for rat U2 RNA (B) is also shown. The positive numbers refer to nucleotides in the sequence for rat U2 RNA (2). Discrepancies between the rat U2 RNA sequence and the established DNA sequence are indicated (*).

sequence of rat U2 RNA (2) although there are several discrepancies between the two sequences which cannot be ascribed to sequencing or cloning artefacts.

Sequences which are related to different snRNAs do not seem to be clustered.

In order to study whether sequences for more than one class of snRNA are present in the recombinants, hybridizations were carried out. DNA from three recombinants U2/4, U2/6 and U2/7 was cleaved with a mixture of restriction endonucleases EcoRI and XhoI. The resulting fragments were separated on a 1% agarose gel and subsequently transferred to a nitrocellulose sheet. Purified ^{32}P-labeled U1 or U6 RNA was used as probes for hybridization. The results showed no distinct hybridization of U1 or U6 RNA to either of three clones (Fig.7). DNA from clones containing U1- or U6 related sequences (unpublished data) was included as a positive control in the experiment (Fig. 7).

DISCUSSION

The purpose of this investigation was to study regions in the human genome which are complementary to the small nuclear RNA U2. Available reports on U2 genes are conflicting both with regard to their number and their mode of transcription. From the results described in this report it is apparent that

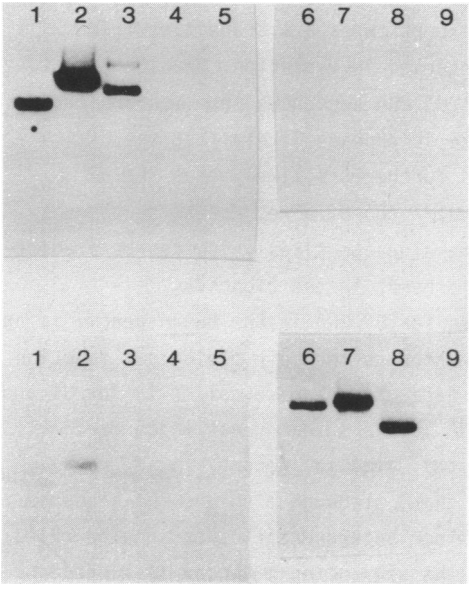

Figure 7: Hybridization of ^{32}P-labeled U1 (top) and U6 (bottom) RNA to different snRNA clones. DNA from the different clones was cleaved with a mixture of restriction endonucleases XhoI and EcoRI, separated on a 1% agarose gel and transferred to a nitrocellulose membrane (13). Clones U2/4 (4), U2/7 (5) and U2/6 (9) were analyzed as well as three different clones containing U1 related sequences (1,2,3) and three different clones containing U6 related sequences (6,7,8).

the human genome contains multiple copies of sequences which are complementary to U2 RNA. Ten out of 100,000 recombinants were found to carry sequences related to U2 RNA which is much higher number than expected for a single copy gene. Extrapolation from the observed frequency of U2 sequences in the DNA library would suggest at least 30 copies per haploid genome equivalent. This is, however, a minimal number since the DNA library probably does not represent the entire human genome.

The second conclusion that can be drawn from our results is that the U2 sequences are not clustered in the human genome. Each of the three recombinants appear to contain one single copy of the sequences which hybridize to U2 RNA, as demonstrated by hybridization of U2 RNA to fragments of the three clones. Hybridization studies with other snRNAs like U1 and U6 (Fig. 7) show moreover that sequences corresponding to different snRNAs are not grouped together, at least not within a distance of 15 kb.

Denison et al. (21) have recently reported similar results and they also found multiple copies of sequences which are complementary to both U1 and U2 RNA. Their clones had sequences which did not match the known snRNA sequences perfectly and were hence designated as "pseudogenes". The clone which was sequenced in the present study is different from that reported by Denison et al. (21). It is for the moment impossible to decide whether the U2/4 clone represents a pseudogene since the sequence of human U2 RNA is not yet known. A comparison between the U2/4 sequence and the sequence of rat U2 RNA reveals 17 differences which all except three are transitions. The snRNAs appear, however, to be well conserved in evolution. Branlant et al. (22) have, for instance, reported that U1 RNA sequences from man and rat only differ in two positions. Therefore it appears likely that the U2/4 clone carries a pseudogene. It appears furthermore likely that the U2 sequences which are present in clones U2/7 and U2/6 represent pseudogenes since not all of the restriction enzyme cleavage sites which can be predicted from the RNA sequence appear to be present in the clones.

The presence of multiple pseudogenes for U2 RNA in the human genome is an enigmatic discovery and raises the question as to their biological function, if any. From our results and from the data of Denison et al. (21) for U1 and U2 RNA and from Hayashi (23) for U6 RNA it seems that genes which perfectly match the snRNA sequences are rare in the mammalian genome.

It should, however, be pointed out that, although a unique RNA-sequence has been established for most snRNAs minor heterogeneities can not be ruled out due to the limited sensitivity of the sequencing methods. If represent-

ing only a small fraction, a transcript from a pseudogene would propably have escaped detection. If the snRNAs have a role in the splicing reaction it would be an intriguing possibility to have multiple forms of the RNA which could splice different mRNA precursors. A more careful search for alternative forms of U2 RNA in vivo should clearly be conducted.

Little is known about the mechanism by which snRNAs are transcribed. Recent studies (7) in an avian system imply that U1 RNA is a product of mammalian RNA polymerase II. Approximately 30 bp upstream from most RNA polymerase II start sites an AT-rich, so called TATA-box, is found (for a review see ref 24). In the sequence shown in Fig. 6 no sequence resembling the TATA-box can be found in the appropriate position. The finding should, however, be interpreted with caution since it is not yet established whether the U2 like sequence in clone U2/4 is transcribed.

The most unexpected finding in the present study is the observation that the sequences which flank the U2-like sequences are highly diverged. Under conditions which would allow hybridization of 35% mismatched sequences only 600-800 bp of the 16 kb long inserts were found to hybridize. One possible explanation for this finding is that multiple copies of U2 genes arose accidentally during mammalian evolution through repeated duplications. Being superfluous, they may then have rapidly degenerated. There is, however, a striking discrepancy between the degree of conservation of the "pseudogenes" themselves and their immediate surroundings. Therefore, if correct, this hypothesis demands that the pseudogenes at some time during evolution served a function which has conserved part of their structure as opposed to the flanking sequences. A more intriguing possibility is, however, that the multiple copies were created through other mechanisms than duplication of large chromosomal segments, i.e. transposition, and that they still have a biological function.

The nature of the two additional segments, "b" and "c" which are present in all three recombinants is obscure. These sequences are apparently more diverged than the pseudogenes themselves. In heteroduplexes between U2/6 and U2/7 they are not even able to form hybrids which were stable under the hybridization conditions used. These regions are likely to represent repetitive sequences which are known to be scattered in the mammalian genome. We are currently trying to correlate them to other repeated sequences, like those which are located adjacent to the globin genes.

ACKNOWLEDGEMENTS

We thank Marianne Gustafson for excellent secretarial help and Michael Katze for linguistic revision of the manuscript. The expert photographic work of Hannu Ukkonen and Uno Skatt is also gratefully acknowledged. This investigation was supported by grants fromt he Swedish Medical Research Council and the Swedish Cancer Society.

REFERENCES

1. Lawn, R.M., Fritsch, E.F., Parker, R.C., Blake, G. and Maniatis, T. (1978) Cell 15, 1157-1174.
2. Reddy, R. and Busch, H. (1981) in The Cell Nucleus, Busch, H. Ed., Vol. VIII, pp. 261-306, Academic Press, New York.
3. Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L. and Steitz, J.A. (1980) Nature 283, 220-224.
4. Rogers, J. and Wall, R. (1980) Proc.Natl.Acad.Sci. U.S.A. 77, 1877-1879.
5. Ohshima, Y., Itoh, M., Okada, N. and Miyata, T. (1981) Proc. Natl.Acad.Sci. U.S.A. 78, 4471-4474.
6. Eliceiri, G.L. (1979) Nature 279, 80-81.
7. Roop, D.R., Kristo, P., Stumph, W.E., Tsai, M.J. and O'Malley, B.W. (1981) Cell 23, 671-680.
8. Maxam, A.M. and Gilbert, W. (1980) in Methods in Enzymology, Grossman, L. and Moldave, K. Eds., Vol. 65, pp 499-560, Academic Press, New York.
9. Vennström, B., Pettersson, U. and Philipson, L. (1978) Nucleic Acids Res. 5, 205-219.
10. Benton, W.D. and Davis, R.W. (1977) Science 196, 180-182.
11. Denhardt, D. (1966) Biochem.Biophys. Res. Commun. 23, 641-646.
12. 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. and Smithies, O. (1977) Science 196, 161-169.
13. Southern, E.M. (1975) J.Mol.Biol. 18, 503-517.
14. Westin, G., Visser, L., Zabielski, J., van Mansfeld, A.D.M., Pettersson, U. and Rozijn, T.H. (1981) Gene, submitted.
15. Aleström, P., Akusjärvi, G., Perricaudet, M., Mathews, M.B., Klessig, D.F. and Pettersson, U. (1980) Cell 19, 671-681.
16. Davis, R., Simon, R. and Davidson, N. (1971) in Methods in Enzymology, Grossman, L. and Moldave, K., Eds. Vol. 21, pp. 413-428, Academic Press, New York.
17. Davis, R.W. and Hyman, R.W. (1971) J.Mol.Biol. 62, 287-301.
18. McConnaughty, B.L., Laird, C.D. and McCarthy, B.J. (1969) Biochemistry 8, 3289-3295.
19. Ferguson, J. and Davis, R.W. (1978) in Advanced techniques in biological electron microscopy, Koehler, J.K. Ed., vol. II, pp. 123-171, Springer, Berlin.
20. Maniatis, T., Hardison, R.C., Lazy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G.K. and Efstratiadis, A. (1978) Cell 15, 678-701.
21. Denison, R.A., Van Arsdell, S.W., Bernstein, L.B. and Weiner, A.M. (1981) Proc.Natl.Acad.Sci. U.S.A. 78, 810-814.
22. Branlant, C., Krol, A., Ebel, J-P., Lazar, E., Gallinaro, H., Jacob, M., Sri-Widada, J. and Jeanteur, P. (1980) Nucleic Acids Res. 8, 4143-4154.
23. Hayashi, K. (1981) Nucleic Acids Res. 9, 3379-3388.
24. Breathnach, R. and Chambon, P. (1981) Ann. Rev. Biochem. 50, 349-383.