MOLECULAR AND CELLULAR BIOLOGY, Sept. 1985, p. 2172–2180
0270-7306/85/092172-09$02.00/0

Vol. 5, No. 9

Human U1 Small Nuclear RNA Pseudogenes Do Not Map to the Site of the U1 Genes in 1p36 but Are Clustered in 1q12-q22

VALERIE LINDGREN,[1] LAUREL B. BERNSTEIN,[2]† ALAN M. WEINER,[2] AND UTA FRANCKE[1]*

Departments of Human Genetics,[1] and Molecular Biophysics and Biochemistry,[2] Yale University, New Haven, Connecticut 06510

Received 28 February 1985/Accepted 5 June 1985

Human U1 small nuclear RNA is encoded by approximately 30 gene copies. All of the U1 genes share several kilobases of essentially perfect flanking homology both upstream and downstream from the U1 coding region, but remarkably, for many U1 genes excellent flanking homology extends at least 24 kilobases upstream and 20 kilobases downstream. Class I U1 RNA pseudogenes are abundant in the human genome. These pseudogenes contain a complete but imperfect U1 coding region and possess extensive flanking homology to the true U1 genes. We mapped four class I pseudogenes by in situ hybridization to the long arm of chromosome 1, bands q12-q22, a region distinct from the site on the distal short arm of chromosome 1 to which the U1 genes have been previously mapped (Lund et al., Mol. Cell. Biol. 3:2211–2220, 1983; Naylor et al., Somat. Cell Mol. Genet. 10:307–313, 1984). We confirmed our in situ hybridization results by genomic blotting experiments with somatic cell hybrid lines with translocation products of human chromosome 1. These experiments provide further evidence that class I U1 pseudogenes and the true U1 genes are not interspersed. The results, along with those published elsewhere (Bernstein et al., Mol. Cell. Biol. 5:2159–2171, 1985), suggest that gene amplification may be responsible for the sequence homogeneity of the human U1 gene family.

U1 small nuclear RNA (snRNA), as part of a small nuclear ribonucleoprotein particle (snRNP), is an essential component of the cellular apparatus for splicing the precursors of mRNA. The 5′ end of U1 snRNA is capable of base pairing with the 5′ end of introns (24), and enzymatic removal of this region of the U1 molecule from an otherwise intact U1-snRNP specifically destroys the ability of the particle to participate in splicing (16). U1 is a very abundant RNA species in mammals, presumably because introns are common in mammalian genes, and cells meet the demand for U1 RNA by having multiple U1 genes. The human haploid genome contains about 30 U1 genes (19), which together produce about 10^6 molecules of U1 per cell per generation (29). The coding regions in each of the sequenced U1 genes are identical (20), as expected for a homogeneous RNA species (26), but analysis of the flanking sequences has proved that they also are highly conserved for several kilobases (kb) both upstream (20) and downstream (13). In fact, as shown elsewhere (1), excellent flanking homology between many human U1 genes extends even farther, for a remarkable 24 kb upstream and 20 kb downstream, so that each of these U1 genes is embedded in at least 44 kb of nearly identical sequence environment. Although most or all of the 30 U1 genes are known to be clustered in band p36 of chromosome 1 (18, 25), the exact extent of the sequence homology and the distance between the genes remain to be established.

In addition to the active genes, the U1 multigene family contains at least 500 to 1,000 unexpressed pseudogenes with imperfect coding sequences (6). The pseudogenes can be divided into at least three classes based on their structures. Pseudogenes of classes II and III appear to have arisen by transposition mechanisms involving the RNA molecule as an intermediate, and in this respect, they resemble many other processed pseudogenes whose flanking regions do not resemble those of the active genes. In contrast, the U1 pseudogenes of class I share both 5′ and 3′ flanking sequence homology with the true U1 genes and thus appear to be derived from the true genes through DNA-mediated mechanisms that do not distinguish between flanking and RNA coding regions. However, when we began these studies we did not know whether class I U1 pseudogenes were the aging ancestors or the degenerate descendents of the contemporary family of true U1 genes.

The mechanism that maintains the homogeneity of multiple U1 genes is unknown. As first noted by Edelman and Gally (7), natural selection alone cannot act forcefully on a single mutated member of a multigene family, because the other members of the family should be sufficient to meet the needs of the organism. Many mechanisms have been proposed to explain the homogenization of multigene families including gene conversion, unequal recombination, and gene amplification. Whatever mechanism is responsible for maintaining the homogeneity of the U1 coding regions, it must also account for the homogeneity of at least 44 kb of DNA sequence flanking many U1 genes, even though the relatively small size of control sequences such as TATA boxes, enhancers, and cap sites would not appear to require such extensive conservation. We reasoned that the extraordinary abundance of class I U1 pseudogenes compared with U1 genes might be a byproduct of the homogenization process (30). Therefore, we wished to determine whether the class I U1 pseudogenes are interspersed with the U1 genes at 1p36, clustered at a different chromosomal location, or dispersed throughout the genome. Knowledge of the chromosomal map position of the genes and pseudogenes might allow us to distinguish between various theories regarding the homogenization of multigene families.

We describe here the mapping of four different class I U1 pseudogenes to chromosomal region 1q12-q22 by in situ hybridization. These results were confirmed by Southern blotting experiments in which one of the pseudogene probes

* Corresponding author.
† Present address: Institute of Molecular Biology, University of Oregon, Eugene, OR 97403.

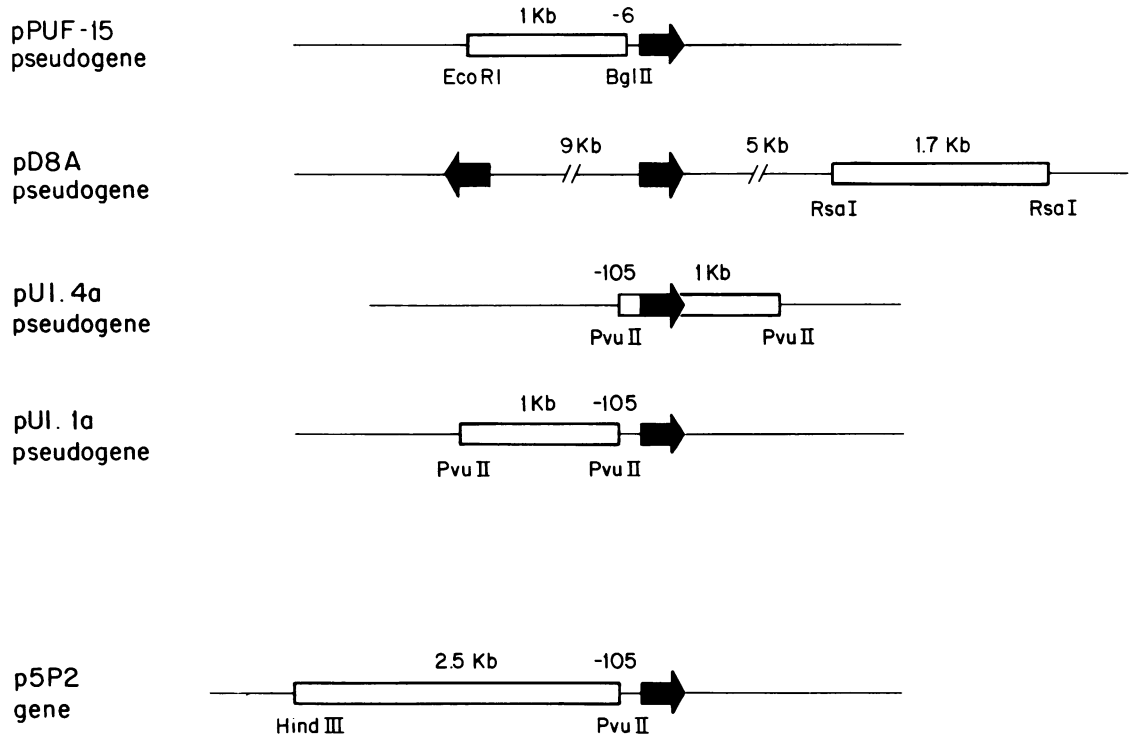

FIG. 1. Origins of the four class I U1 pseudogene probes (top four drawings) and one U1 gene probe (bottom drawing). The black arrows represent the regions corresponding to sequences that code for U1 RNA. Boxed regions are those nucleotides that were subcloned into pBR322 for use as hybridization probes. Distances are not shown to scale.

and a U1 gene probe were hybridized to DNA from several somatic cell hybrid lines containing translocations that separated the U1 genes in 1p36 from the pseudogenes in 1q12-q22. Therefore, the class I pseudogenes are not interspersed with the true genes but are clustered at a different chromosomal site. These results, along with the restriction mapping, sequence analysis, and Southern blot analysis of U1 genes and class I U1 pseudogenes presented elsewhere (1), lead us to suggest that both the genes and pseudogenes are arranged in large, irregular tandem arrays. We also argue that repeated cycles of gene amplification best account for the extraordinary sequence homogeneity of the contemporary U1 gene family.

MATERIALS AND METHODS

Cell culture and production of somatic cell hybrids. Chromosomes were prepared from phytohemagglutinin-stimulated, methotrexate-synchronized cultures of lymphocytes from normal human donors as previously described (8).

Somatic cell hybrids were formed between Chinese hamster cells (380-6) and human fibroblasts with a balanced reciprocal translocation between the short arms of chromosomes 1 and 6 [46,XY,t(1;6)(p3200;p2100)] (see Fig. 4) (28). The chromosomal constitution of the hybrids was determined by trypsin-Giemsa banding and confirmed by isozyme analysis at the same passage as DNA extraction. Cell line XV-18A-10a-D4 (hereafter referred to as XV-D4) contained one copy of the derivative 1 chromosome [der(1)] per cell but lacked the derivative 6 chromosome [der(6)] and normal chromosome 1. Cell line XV-16A-F4 (XV-F4) contained one copy of the der(6) per cell and no other chromosomes with parts of chromosome 1. No part of chromosome 1 was apparent in the karyotype of line XV-18A-8b-G1 (XV-G1) (but see below), and line XV-18B-7a-N4 (XV-N4) had one copy each of the der(1) and normal chromosome 1.

Plasmid construction. From four cloned class I U1 pseudogenes, we subcloned fragments ranging in size from 1 to 1.7 kb for use as in situ hybridization probes. Probes pU1.1a, pU1.4a, and pPUF-15 were derived from clones λU1.1, λU1.4, and λU1.15, respectively; the isolation of these clones from a lambda bacteriophage library and their characterization have been described previously (6). The pD8A probe is a subclone of cosD8A, a clone isolated from a cosmid library and described elsewhere (1). The four original clones λU1.1, λU1.4, λU1.15, and cosD8A represent distinct genomic loci, a conclusion based on both restriction mapping and direct DNA sequence analysis of the U1 RNA coding regions (1, 6).

Figure 1 diagrams the derivations of the four class I U1 pseudogene probes. Subclone pPUF-15 was constructed by insertion of the *Bgl*II-*Eco*RI fragment, spanning the 5′ flanking sequences of λU1.15 from nucleotide −6 and extending 1 kb upstream, between the *Eco*RI and *Bam*HI sites of pBR322. The other three subclones were constructed by insertion into the *Pvu*II site of pBR322. Plasmid pU1.1a contains the *Pvu*II fragment spanning nucleotides −105 to −1100 of λU1.1. Plasmid pU1.4a contains a 1.0-kb *Pvu*II fragment that spans the U1 coding region of λU1.4, starting at position −105. Plasmid pD8A contains a 1.7-kb *Rsa*I fragment of cosD8A, located 5 kb downstream from one of the two class I pseudogene coding regions of cosD8A. Surprisingly, pD8A cross-hybridizes with sequences in the immediate 5′ flanking regions of all U1 genes (1).

We used the true U1 gene subclone p5P2 (provided by T. Manser) as a control for cross-hybridization between true U1 genes and class I U1 pseudogenes. This plasmid consists of a *Hin*dIII-*Pvu*II fragment spanning nucleotides −105 to

TABLE 1. In situ hybridizations with U1 pseudogene and U1 gene probes

Probe	Concn (ng/ml)	Emulsion exposure (days)	No. of cells	Total no. of grains	Grains over 1q12-q22 (% of total)	Grains over 1p35-p36 (% of total)	Grains over q/ grains over p
pPUF-15[a]	50	14	104	474	67 (14.1)	24 (5.1)	2.79
pPUF-15[1a,b]	50	14	34	129	26 (20.2)	17 (13.2)	1.53
pD8A[a]	25	10	27	76	27 (35.5)	9 (11.8)	3.00
pD8A[a]	50	10	47	205	32 (15.6)	24 (11.7)	1.33
pU1.4a[a]	25	14	75	184	41 (22.3)	15 (8.1)	2.70
pU1.1a[a]	25	14	49	113	32 (28.3)	18 (15.9)	1.78
p5P2[c]	25	5	94	581	42 (7.2)	128 (22.0)	0.33

[a] Class I pseudogene probe.
[b] The same chromosome donor was used in all experiments except this one.
[c] Gene probe.

−2600 from the U1 gene HSD2 (20) that was cloned into pBR322 between the *Hin*dIII and *Pvu*II sites (Fig. 1).

In situ hybridization. Plasmids were labeled by nick translation with tritiated dATP, dCTP, and dTTP to specific activities of 2×10^7 to 3.3×10^7 cpm/μg. The probes were hybridized to chromosomes overnight at 37°C by the method of Harper and Saunders (11). After completion of the washes, the chromosomes were stained with quinacrine dihydrochloride (17). The slides then were coated with Kodak NTB2 photographic emulsion, which was exposed for 5 to 14 days at 4°C. After development of the emulsion in Kodak Dektol at 15°C for 3 min, the chromosomes were stained with quinacrine mustard dihydrochloride (15) and then with 0.06% Wright stain in 0.06 M phosphate buffer (pH 6.8). This double-staining method permitted unequivocal identification of all chromosomes, and grains could be easily scored on the Wright-stained preparations.

Southern blotting experiments. DNA was isolated from each of the somatic cell hybrid lines described above, from the parental Chinese hamster cell line (380-6), and from the placenta of a human individual unrelated to the donor of the hybrid parental fibroblasts. DNA samples (10 μg) were digested with *Pst*I, resolved by agarose gel electrophoresis, and transferred to nitrocellulose as described previously (2). As hybridization probes we used either plasmid p5P2 labeled by nick translation or mp9-PUF15 (the pPUF-15 insert in the M13 mp9 vector; 23) labeled by primer extension (2). The blot first was hybridized to mp9-PUF15, autoradiographed, and then stripped of probe DNA by brief boiling in 0.1× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) (27). After another exposure of the blot to verify that the first probe had been removed, it was rehybridized with p5P2 and autoradiographed. Prehybridization and hybridization were performed as described previously (6) but at 54 to 55°C.

RESULTS

The mapping of class I U1 RNA pseudogenes is complicated by their considerable homology to the U1 coding and flanking sequences. Under the relatively nonstringent conditions of in situ hybridization, pseudogene probes would be expected to cross-react with the true U1 genes. Because the flanking homologies are less extensive (about 80%) than those of the coding regions (about 95%) (6), we constructed plasmids containing predominantly pseudogene flanking regions for use in the mapping (Fig. 1). Probes pPUF-15 and pU1.1a are from the 5′ flank of pseudogenes U1.15 and U1.1, respectively. Probe pD8A is from the 3′ flanking region of the cosD8A-II pseudogene but is homologous to the sequences found at the 5′ end of both U1 genes and other class I U1 pseudogenes (see reference 1 for a description of the probable evolution of the cosD8A locus). Although probe pU1.4a contains the entire but defective U1 coding sequence of the U1.4 pseudogene, most of the insert sequences are from the 3′ flank of the U1.4 pseudogene. Plasmid p5P2, with sequences derived from the 5′ flank of the true U1 gene HSD2, was used as a control.

In situ hybridizations. The four pseudogene probes and the U1 gene probe were used in a total of seven in situ hybridizations to normal human chromosome spreads (Table 1). For each hybridization with a pseudogene probe, the greatest concentration of silver grains was over bands q12-q22 of chromosome 1 (Fig. 2 and 3). In each case, a smaller percentage of the total grains was over bands p35-p36 of chromosome 1, where the U1 genes are known to map (18, 25). When the hybridization of the U1 gene probe (p5P2) was analyzed, the opposite result was found. That is, the greatest concentration of grains was over 1p35-p36 as expected, and a smaller proportion was over 1q12-q22 (Table 1).

Because the probes for class I U1 pseudogenes cross-hybridize with the U1 genes, the results of these experiments may be interpreted more easily by comparing, in each case, the number of grains over the long-arm site (1q12-q22) to the number of grains over the short-arm site (1p35-p36) in the form of a ratio (last column of Table 1). This ratio can be manipulated experimentally, as seen in the two pD8A hybridizations. At the lower probe concentration of 25 ng/ml, the ratio was larger and the 1q site appeared to be the major site of hybridization. The different ratios in the two pPUF-15 experiments at the same probe concentration of 50 ng/ml may reflect individual differences in the number of U1 genes and class I pseudogenes, since different chromosome donors were used. In any case, the q/p ratios for the pseudogene experiments were all greater than 1 (1.33 to 3.00), but with the true U1 gene probe (p5P2) the ratio was much less than 1 (0.33). The results indicate that all four of these class I U1 pseudogenes map to 1q12-q22.

The statistical significance of the data was determined by chi-square analysis, in which the number of grains observed over a chromosome arm was compared to the number expected based on the length of the arm. For every experi-

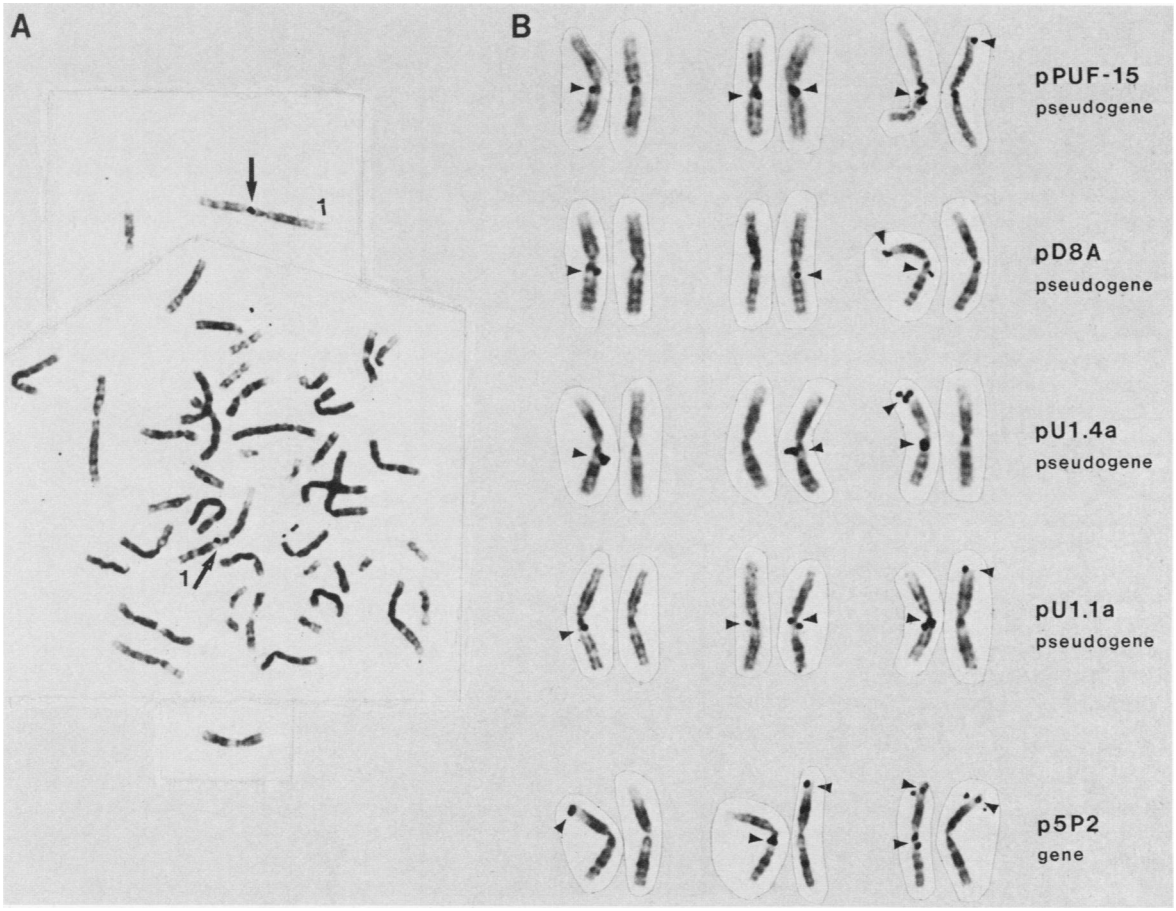

FIG. 2. Normal human metaphase chromosomes stained with Wright stain after in situ hybridization. (A) Representative cell after hybridization with probe pD8A. Arrows point to silver grains just below the heterochromatic regions of the long arms of both chromosomes 1. (B) Pairs of chromosomes 1 after hybridization with the pseudogene probes (upper four rows) and the U1 gene probe (bottom row). Arrowheads mark the locations of grains. For each of the pseudogene probes, most of the grains were seen just below the heterochromatic region of 1q, but some grains were over region 1p35-p36. For the p5P2 gene probe, most of the grains were over the 1p site, but some were over region 1q12-q22.

ment the amount of label over 1q and the amount over 1p were highly significant ($P < 0.0001$). In only one case did we observe a higher number of grains than expected over a chromosomal arm: in the pU1.4a experiment, eight grains were counted over the short arm of chromosome 19, whereas only two were expected. The possibility of a third site of hybridization was evaluated by scoring the grains over chromosome 19 in 50 additional cells from the same experiment. In these cells, only one grain was over 19p, exactly the number expected by chance. Therefore, the excess of grains over 19p in the first group of 75 cells may be explained as a sampling error. In addition, no label above background was observed over 19p with the other three pseudogene probes or with the U1 gene probe. Although single copy sequences are routinely detected with this technique with probes even smaller than 1 kb, none of the four pseudogene probes hybridized significantly to sites other than 1q12-q22 and 1p35-p36. It is unlikely that class I U1 pseudogenes are present elsewhere in the human genome.

Genomic blotting with hybrid cell lines. The results of the in situ hybridizations suggest that most or all class I U1 pseudogenes are clustered in 1q12-q22 and that most or all true U1 genes are clustered in 1p35-p36. However, using this approach we could not exclude the possibility that some U1 genes are interspersed with class I pseudogenes or vice versa. Therefore, we analyzed the human U1 gene family by genomic blotting with somatic cell hybrid lines to look for a low level of interspersion of U1 genes and class I pseudogenes.

Our analysis was based on the observation that U1 genes lie on restriction fragments of characteristic size and can be readily distinguished from class I pseudogenes by genomic blotting. In particular, digestion of human DNA with *Pst*I divides the 30 true U1 genes into two approximately equal groups; the p5P2 U1 gene probe reacts with *Pst*I fragments of 3.9 and 2.6 kb, each present in approximately 15 copies per haploid genome (1, 20). This restriction-fragment-length polymorphism reflects minor sequence heterogeneity and is completely consistent with our conclusion that many true U1 genes share at least 44 kb of nearly perfect flanking homology (1). In contrast, sequences flanking the many class I U1 pseudogenes are much more diverse, and *Pst*I digestion probably divides them into fragments of many different lengths. However, at high stringency, the mp9-PUF15 probe, which contains a class I U1 pseudogene on a 2.3-kb *Pst*I fragment, hybridizes only with fragments that are 3.9, 2.3, and 1.8 (doublet) kb in length; the major 2.3-kb fragment is present at approximately one copy per haploid genome. Apparently, the mp9-PUF15 pseudogene probe cross-hybridizes weakly with the 10 to 15 copies of the 3.9-kb *Pst*I

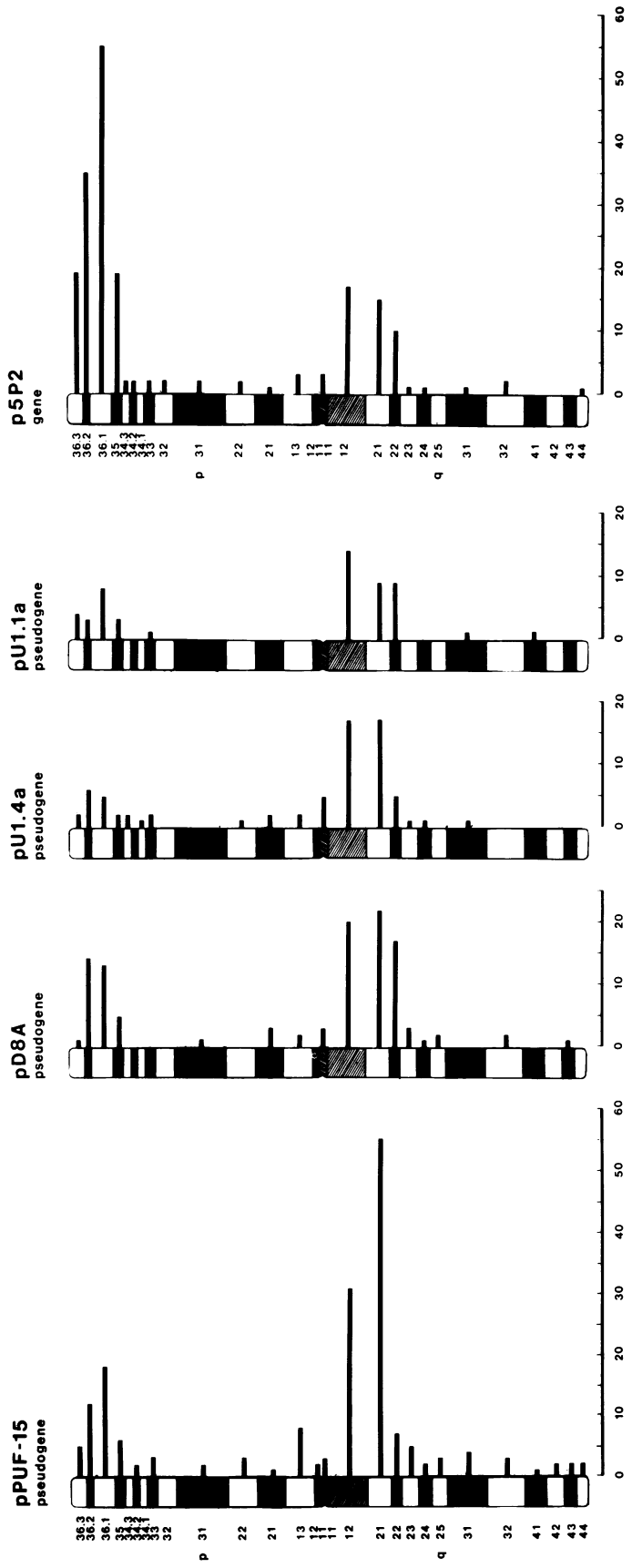

FIG. 3. Ideograms showing the distribution of silver grains over chromosomes 1 after hybridization with each of the four pseudogene probes (four ideograms on the left) and with the U1 gene probe (ideogram on the right). Bars indicate the numbers of grains observed over the entire width of the adjacent bands.

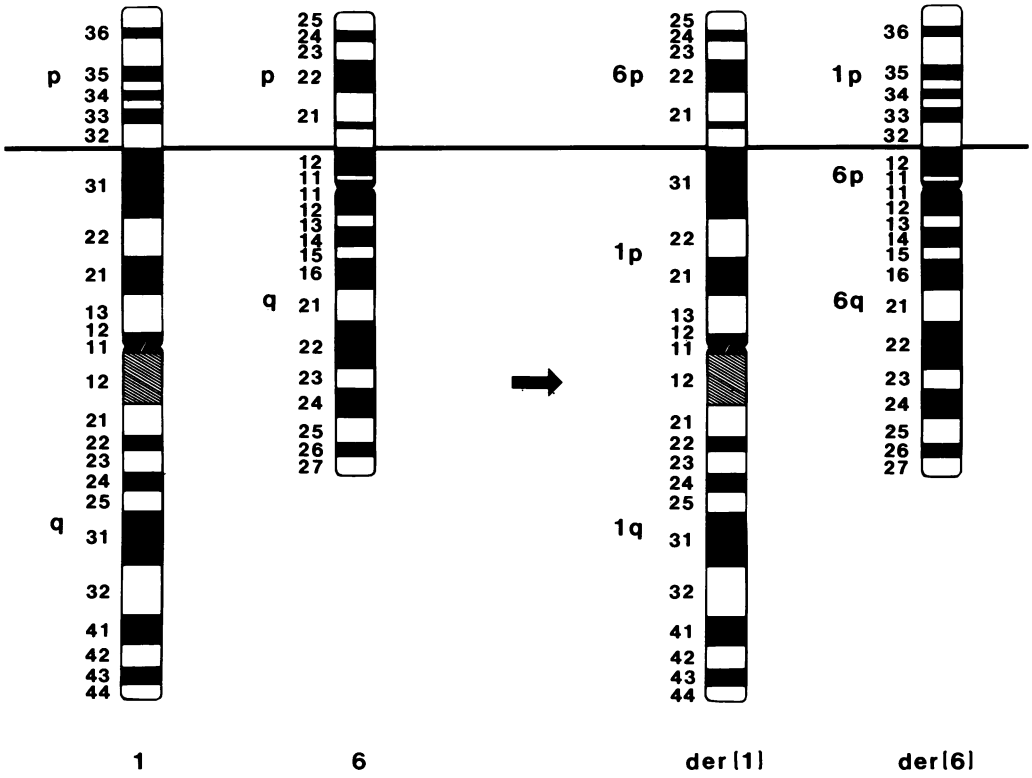

FIG. 4. Ideograms demonstrating the derivation of the balanced reciprocal translocation in the human parental fibroblasts [46,XY,t(1;6)(p3200;p2100)] used to form the series XV cell hybrids. The black line indicates the sites of breaks in chromosomes 1 and 6 that generated the der(1) and der(6) by reciprocal exchange.

true gene fragment or with another class I pseudogene which lies within a 3.9-kb *Pst*I fragment (see below).

The DNA used in this analysis was obtained from somatic cell hybrids constructed from human parental cells with a balanced reciprocal translocation between the short arms of chromosomes 1 and 6 [46,XY,t(1;6)(p3200;p2100)] (28) (Fig. 4). Thus, this translocation should separate the site in 1q assumed to carry the class I U1 pseudogenes from the site in 1p assumed to carry the true genes. One hybrid cell line, XV-D4, isolated the der(1) away from the der(6) and normal chromosome 1. Another line, XV-F4, isolated the der(6) away from the der (1) and chromosome 1. Cell line XV-G1, lacking both the 1q and the 1p site, served as a negative control, whereas cell line XV-N4, with an intact chromosome 1 as well as the der(1), was a positive control.

Even at the high stringency attainable in Southern blotting, cross-hybridization between genes and pseudogenes cannot be completely avoided. However, the results are always interpretable because of the restriction-fragment-length polymorphisms. If the class I pseudogenes all lie within 1q12-q22 and the U1 genes are all within 1p35-p36, the characteristic pseudogene bands will be present only when the cell line contains chromosome 1 or the der(1) (as in XV-N4 and XV-D4, respectively); the characteristic gene bands will be present only when chromosome 1 or the der(6) is present in the hybrid (as in XV-N4 and XV-F4, respectively).

When pseudogene probe mp9-PUF15 was hybridized to DNA from line XV-D4 with the der(1), *Pst*I fragments characteristic of pseudogenes (2.3 kb and a 1.8-kb doublet) were observed (Fig. 5A); a 3.9-kb fragment found with both gene and pseudogene probes also was present. In hybrid XV-F4 containing the der(6), fragments characteristic of U1 genes (3.9 and 2.6 kb) were detected. The positive control (hybrid XV-N4), with an intact chromosome 1, had all gene and pseudogene bands. The negative control (cell line XV-G1) had a single, very faint band at 2.3 kb. Although no part of chromosome 1 is apparent in the karyotype of this hybrid, sequences on the chromosome 6-derived portion of the der(6) also have been detected at a low level in the absence of cytologically detectable fragments of chromosome 6 (B. de Martinville and U. Francke, unpublished data). Therefore, the der(6) or a rearranged form of it may be present in this hybrid at a very low copy number. In any case, genomic blotting with a pseudogene probe demonstrates that the reciprocal translocation does separate the true U1 genes from the class I U1 pseudogenes.

We confirmed this result by rehybridizing the same blot of hybrid cell DNA with the U1 gene probe p5P2 (Fig. 5B). Although this probe cross-hybridized with the class I U1 pseudogenes at the relatively low stringency of in situ hybridizations, it did not cross-hybridize at the higher stringency of genomic blotting. No bands were seen in the DNA containing only the 18 site (XV-D4) or in the negative control DNA (XV-G1). The hybrid containing only the 1p site (XV-F4) had the 3.9- and 2.6-kb bands characteristic of the U1 gene, but it also had a 2.3-kb band. This band is not the 2.3-kb pseudogene band: since the relatively long insert of the p5P2 probe spans a *Pst*I site, we evidently detected the 2.3-kb *Pst*I fragment lying immediately upstream of the 2.6- or 3.9-kb gene fragment. This upstream fragment is also present in the reconstruction lane (lane R of Fig. 5) and only coincidentally overlaps the pseudogene band. The 1.8-kb pseudogene band was not detected in any of the hybrid lines.

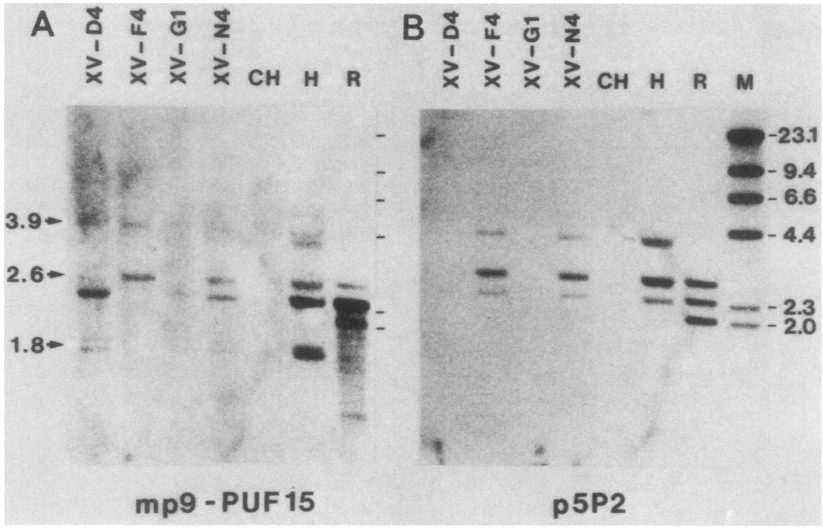

FIG. 5. Southern blot of DNAs digested with *Pst*I and hybridized first with pseudogene probe mp9-PUF15 (A) and then with true gene probe p5P2 (B). The first four lanes contain DNA from hybrid cell lines XV-D4, XV-F4, XV-G1, and XV-N4. The remaining lanes contain: CH, DNA from the Chinese hamster parental cells (380-6) used to construct the somatic cell hybrids; H, human placental DNA; R, genomic reconstruction containing five copy equivalents each of HSD2 (a true U1 gene) and U1.15 (a class I U1 pseudogene); and M, lambda DNA digested with *Hin*dIII as markers.

The fact that the pseudogene probe cross-hybridizes to true U1 genes (XV-F4), whereas the gene probe does not cross-react with pseudogenes (XV-D4), can be readily explained by reconstruction experiments. When a mixture of *Pst*I-digested cloned U1 gene and U1.15 DNA was hybridized with pseudogene probe mp9-PUF15 (Fig. 5A, lane R), the five copy equivalents of the pseudogene (2.3-kb band) gave a much stronger signal than did the five copy equivalents of the gene DNA (2.6-kb band). Thus, the bands in XV-F4 DNA probably represent many copies of weakly hybridizing U1 genes. Although the true U1 genes with highly conserved flanking sequences are all found on 2.6- or 3.9-kb *Pst*I fragments, it is likely that U1 pseudogenes, which have much more divergent flanking sequences than do the true genes, reside on *Pst*I fragments of many different sizes. Therefore, too few copies of U1 pseudogenes may be present within a given restriction fragment length in *Pst*I-digested XV-D4 DNA to produce a detectable hybridization signal with the p5P2 gene probe.

Despite the complexities presented by the genomic blots (Fig. 5), the data indicate that the U1 genes segregate with the 1p chromosomal site and the class I pseudogenes segregate with the 1q site. We conclude that none or very few of the class I pseudogenes are interspersed with the U1 genes and none or very few of the U1 genes are interspersed with the pseudogenes.

DISCUSSION

Very little is understood about the maintenance of homogeneity among members of large mammalian gene families such as those encoding rRNA, 5S RNA, or U1 and U2 snRNAs. For gene families with only a few members, it is conceivable that natural selection alone maintains the identity of the gene copies, because a defect in a single gene could have an effect on the fitness of the organism. In contrast, the existence of many wild-type genes in a large gene family would shield each individual gene copy from the full force of natural selection, and we might have expected that each member of the family would accumulate a different set of mutations. This, however, is not the case. Large gene families display as much, and often more, homogeneity than do small gene families. We studied the genomic organization of the U1 snRNA family of genes and pseudogenes in the hope of reconstructing its evolutionary history, and thus providing a factual basis for speculation regarding the mechanisms that might be responsible for maintaining the extraordinary homogeneity of the contemporary family of true U1 genes.

To determine the chromosomal organization of the human U1 RNA multigene family, we used probes for four different class I U1 pseudogenes in hybridizations to metaphase chromosomes. Although U1 pseudogenes of this class are closely related to each other and to the true U1 genes, we were able to reduce the expected amount of cross-hybridization by using probes from flanking regions that are more divergent than the actual coding regions. The results of in situ hybridizations with the pseudogene probes were compared with results obtained from hybridization with a true gene probe. For each pseudogene probe, the major site of hybridization was the proximal long arm of chromosome 1 (region 1q12-q22); however, a minor signal was also detected on the distal short arm of chromosome 1 (region 1p35-p36) at the location of the bona fide U1 genes (18, 25). These data strongly imply that all four pseudogenes map to 1q12-q22. This assignment was confirmed by Southern blotting experiments with DNA from rodent-human hybrid cell lines containing reciprocal translocation chromosomes that separated the 1q site of the class I pseudogenes from the 1p site of the true U1 genes. Restriction fragments of lengths characteristic of the class I U1 pseudogenes segregated with the chromosome containing 1q, whereas fragments characteristic of the U1 genes segregated with the chromosome containing distal 1p. Thus, U1 pseudogenes do not appear to be interspersed with the U1 genes, or vice versa.

Genes cannot be assigned reliably to a single chromosomal band by in situ hybridization to normal chromosomes, because the silver grains seen after autoradiography are the result of random radioactive disintegrations in three dimen-

sions (12). Consequently, we assign the class I U1 pseudogenes to a three-band region, 1q12-q22. However, this region includes the large heterochromatic region (1q12) on the long arm of chromosome 1. In fact, most of the grains over this band were over the distal portion. Heterochromatin is thought to be composed largely of relatively simple repeated sequences and is generally assumed to be condensed, transcriptionally inactive chromatin (4, 5). Thus, the class I U1 pseudogenes, which presumably were active genes at one time, are unlikely to lie within even the distal portion of the heterochromatin; however, we have not ruled out this possibility.

Our in situ and genomic blotting experiments with the true U1 gene probe p5P2 confirm the U1 gene assignment to 1p36.1 by Naylor and colleagues (25), which they made by in situ hybridization with the same gene probe. (Band 1p36.1 of the standard ISCN 1981 nomenclature (14) corresponds to band 1p36.3 in the nomenclature of Naylor et al.) Band 1p36.1 would appear to be the most likely location of the U1 genes (Fig.3), but we cannot rigorously exclude band 1p36.2 on the basis of our in situ hybridization data. Our results also strengthen and extend the conclusion that most true U1 genes are within this region of 1p (25). In their experiments, Naylor et al. (25) detected a second, minor site of hybridization in 1q and suggested that this might represent one or a few U1 genes. Based on our results, this signal was probably the result of cross-hybridization of the U1 gene probe with class I U1 pseudogenes in 1q12-q22.

Because all four of the pseudogenes we examined map to region 1q12-q22, the class I pseudogenes must be at least loosely clustered. In addition, two different pseudogenes have been isolated on the same recombinant cosmid (cosD8A; see reference 1). Although these two pseudogenes are closer together than most of the other class I pseudogenes, the relative homogeneity of all characterized class I pseudogenes suggests that these pseudogenes may be organized in an irregular tandem array. The bona fide U1 genes must be at least loosely clustered, because most or all of them map to 1p36. However, no two U1 genes have ever been found together on the same cosmid, nor has the distance between U1 gene copies been established by chromosomal walking. Nevertheless, the extensive conservation of sequences flanking U1 genes (over 44 kb in many cases) and the common chromosomal location of all true U1 genes suggest that the genes are also organized in an irregular tandem array. The simplest explanation of the available data is that human U1 genes were once located in the region 1q12-q22. One of the copies was amplified, and the new copies were moved to 1p36 either as part of the amplification process or subsequent to it. The genes in 1p then became the predominant cluster of active genes, and those in 1q continued to diverge. A more complete discussion of this theory and the arguments in favor of it are contained in reference 1.

As noted previously (17), the location of the tandemly repeated human U2 genes in 17q21-q22 correlates exactly with one of the three major chromosomal modification sites of adenovirus type 12. The other two major modification sites 1p36 and 1q21 correspond surprisingly well with the location of the U1 genes and class I U1 pseudogenes. These three major sites and several minor sites (one near the 5S RNA genes) often appear decondensed or fragile in metaphase chromosomes examined shortly after infection with highly oncogenic adenoviruses (21, 22, 31). Because of the coincidence of the adenovirus modification sites with snRNA gene loci, it was proposed that the highly transcribed snRNA genes are the major targets of chromosome modification by adenovirus 12 (17). It also was suggested that band 1q21 might be modified because the class I U1 pseudogenes retain functional regulatory sequences. We note, however, that a histone gene cluster has subsequently been mapped to proximal 1q (10), and it is possible that these highly transcribed genes rather than the U1 RNA pseudogenes might be the targets of modification. Activation or repression of snRNA genes by an adenoviral gene product (3) could account for the chromosomal modification. Alternatively, viral integration may occur at cellular sites that are actively transcribed; recently, adenovirus types 2 and 12 were found to integrate at chromosomal sites encoding snRNA species (9). Further experiments will be required to determine the significance of the correlation of adenovirus modification sites and snRNA genes.

ACKNOWLEDGMENTS

We are grateful to Catherine Joyce for supplying Klenow polymerase and to Tim Manser for the p5P2 clone.

This work was supported by institutional postdoctoral fellowship GM 07439 from the National Institutes of Health to V.L., by Public Health Service grants GM 31073 and GM 31335 from the National Institutes of Health to A.M.W., and by Public Health Service grant GM 26105 from the National Institutes of Health to U.F.

LITERATURE CITED

1. **Bernstein, L. B., T. Manser, and A. M. Weiner.** 1985. Human U1 small nuclear RNA genes: extensive conservation of flanking sequences suggests cycles of gene amplification and transposition. Mol. Cell. Biol. **5**:2159–2171.
2. **Bernstein, L. B., S. M. Mount, and A. M. Weiner.** 1983. Pseudogenes for human small nuclear RNA U3 appear to arise by integration of self-primed reverse transcripts of the RNA into new chromosomal sites. Cell **32**:461–472.
3. **Borrelli, E., R. Hen, and P. Chambon.** 1984. Adenovirus-2 E1A products repress enhancer-induced stimulation of transcription. Nature (London) **312**:608–612.
4. **Comings, D. E.** 1972. The structure and function of chromatin. Adv. Hum. Genet. **3**:237–431.
5. **Comings, D. E.** 1978. Mechanisms of chromosome banding and implications for chromosome structure. Annu. Rev. Genet. **12**:25–46.
6. **Denison, R. A., and A. M. Weiner.** 1982. Human U1 RNA pseudogenes may be generated by both DNA- and RNA-mediated mechanisms. Mol. Cell. Biol. **2**:815–828.
7. **Edelman, G. M., and J. A. Gally.** 1970. Arrangement and evolution of eukaryotic genes, p. 962–972. *In* F. O. Schmitt (ed.), Neurosciences: second study program. Rockefeller University Press, New York.
8. **Francke, U., and N. Oliver.** 1978. Quantitative analysis of high-resolution trypsin-Giemsa bands on human prometaphase chromosomes. Hum. Genet. **45**:137–165.
9. **Gahlmann, R., M. Schulz, and W. Doefler.** 1984. Low molecular weight RNAs with homologies to cellular DNA at sites of adenovirus DNA insertion in hamster or mouse cells. EMBO J. **3**:3263–3269.
10. **Green, L., R. Van Antwerpen, J. Stein, G. Stein, P. Tripputi, B. Emanuel, J. Selden, and C. Croce.** 1984. A major human histone gene cluster on the long arm of chromosome 1. Science **226**:838–840.
11. **Harper, M. E., and G. F. Saunders.** 1981. Localization of single copy DNA sequences on G-banded human chromosomes by in situ hybridization. Chromosoma **83**:431–439.
12. **Henderson, A. S.** 1982. Cytological hybridization to mammalian chromosomes. Int. Rev. Cytol. **76**:1–46.
13. **Htun, H., E. Lund, and J. E. Dahlberg.** 1984. Human U1 RNA genes contain an unusually sensitive S1 cleavage site within the conserved 3′ flanking region. Proc. Natl. Acad. Sci. U.S.A. **81**:7288–7292.

14. **ISCN.** 1981. An international system for human cytogenetic nomenclature—high resolution banding. 1981. Cytogenet. Cell Genet. **31**:1–23.
15. **Kirsch, I. R., C. C. Morton, K. Nakahara, and P. Leder.** 1982. Human immunoglobulin heavy chain genes map to a region of translocations in malignant B lymphocytes. Science **216**:301–303.
16. **Krämer, A., W. Keller, B. Appel, and R. Lührmann.** 1984. The 5′ terminus of the RNA moiety of U1 small nuclear riboprotein particles is required for the splicing of messenger RNA precursors. Cell **38**:299–307.
17. **Lindgren, V., M. Ares, Jr., A. M. Weiner, and U. Francke.** 1985. Human genes for U2 small nuclear RNA map to a major adenovirus 12 modification site on chromosome 17. Nature (London) **314**:115–116.
18. **Lund, E., C. Bostock, M. Robertson, S. Christie, J. L. Mitchen, and J. E. Dahlberg.** 1983. U1 small nuclear RNA genes are located on human chromosome 1 and are expressed in mouse-human hybrid cells. Mol. Cell. Biol. **3**:2211–2220.
19. **Lund, E., and J. E. Dahlberg.** 1984. True genes for human U1 small nuclear RNA: copy number, polymorphism, and methylation. J. Biol. Chem. **259**:2013–2021.
20. **Manser, T., and R. F. Gesteland.** 1982. Human U1 loci: genes for human U1 RNA have dramatically similar genomic environments. Cell **29**:257–264.
21. **McDougall, J. K.** 1971. Adenovirus-induced chromosome aberrations in human cells. J. Gen. Virol. **12**:43–51.
22. **McDougall, J. K.** 1979. The interactions of adenovirus with host cell gene loci. Cytogenet. Cell Genet. **25**:183.
23. **Messing, J., and J. Vieira.** 1982. A new pair of M13 vectors for selecting either DNA strand of double-digested restriction fragments. Gene **19**:269–276.
24. **Mount, S. M., I. Pettersson, M. Hinterberger, A. Karmas, and J. A. Steitz.** 1983. The U1 small nuclear RNA-protein complex selectively binds a 5′ splice site *in vitro*. Cell **33**:509–518.
25. **Naylor, S. L., B. U. Zabel, T. Manser, R. Gesteland, and A. Y. Sakaguchi.** 1984. Localization of human U1 small nuclear RNA genes to band p36.3 of chromosome 1 by in situ hybridizations. Somat. Cell Mol. Genet. **10**:307–313.
26. **Nohga, K., R. Reddy, and H. Busch.** 1981. Comparison of RNase T_1 fingerprints of U1, U2, and U3 small nuclear RNA's of HeLa cells, human normal fibroblasts, and Novikoff hepatoma cells. Cancer Res. **41**:2215–2220.
27. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
28. **Taggart, R. T., and U. Francke.** 1982. Mapping of polypeptide genes by two-dimensional gel electrophoresis of hybrid cell extracts. Cytogenet. Cell Genet. **32**:99–110.
29. **Weinberg, R. A., and S. Penman.** 1968. Small molecular weight monodisperse nuclear RNA. J. Mol. Biol. **38**:289–304.
30. **Weiner, A. M., and R. A. Denison.** 1983. Either gene amplification or gene conversion may maintain the homogeneity of the multigene family encoding human U1 small nuclear RNA. Cold Spring Harbor Symp. Quant. Biol. **47**:1141–1149.
31. **zur Hausen, H.** 1967. Inductions of specific chromosomal aberrations by adenovirus type 12 in human embryonic kidney cells. J. Virol. **1**:1174–1185.